Study on Data-sharing with Countries of Origin

Contract no. GBIFS/2003/04

February - 2004

Team:

Dora Ann Lange Canhos (dora@cria.org.br)

Arthur Chapman (Arthur@cria.org.br)

Vanderlei Perez Canhos (vcanhos@cria.org.br)

Executive Summary

The present report is the result of a commissioned study by GBIF (OCB Work Programme - contract no. GBIFS/2003/04), to analyze experiences on data sharing with countries of origin (a.k.a. repatriation). This study was carried out by the Reference Centre on Environmental Information (CRIA). Work involved selecting and contacting institutions, developing a questionnaire, tabulating answers, analyzing the responses, and writing this report.

The questionnaire was formulated in consultation with the GBIF secretariat and sent to 27 institutions. A total of 27 institutions were contacted and we received 18 filled out forms. An attempt was made to cover all continents and to include both biological collections and institutions responsible for structuring online information systems. Of the 18 institutions, 4 are from the USA, 8 from Europe, 1 from India, 2 from Australia, and 3 from Latin America. Included are important herbaria and natural history museums that curate specimens from practically the whole world; initiatives concerned with digitizing data for their own institutional purposes; and explicit bilateral agreements directly dealing with data sharing with countries of origin. From a technological view there are institutions concerned in bringing data into the country and maintaining it in a centralized database and others developing distributed systems that will harvest data at remote sites.

The following points were analyzed by this report: data sharing; criteria for the selection of taxonomic groups and collections, priorities and user groups contractual arrangements and intellectual property rights receiving institution/country contribution; process; data format and transfer; exchange formats and standards; funding, time scale, and costs; results and products; problems and hurdles addressed; and recommendations and advice.

In general terms, for the group of institutions that answered the questionnaire it was found that:

- most projects, regardless of whether the main purpose is to share data with the country of origin or not, are making information freely available on the internet;
- data-sharing is carried out as a collaborative effort through informal agreements;
- as to IPR, important issues are proper attribution or credit to all partners involved, custodianship and ownership (i.e. each contributing museum retains ownership of its records) and acknowledgement;
- digitizing and data basing collections are fundamental for the day-to-day operation of a collection. Regardless whether linked to a data-sharing program or not collections must database their collections and document their activities.
- international collaboration, which has always existed in taxonomy, is largely enhanced by on-line dissemination of data and information;
- with the evolution of information and communication technology it is becoming possible to develop inter-disciplinary applications, disseminate information to a wider public and promote the use of scientific data for other purposes;
- financial constraints, followed closely by technological problems and human resources constraints were the three problems and hurdles most mentioned;
- digitization is a trend and will continue to happen. It depends on proper policies and technologies to accelerate the process, and to ensure the accuracy and quality of the resultant product.
- international collaboration is largely enhanced by on-line dissemination of data and information;

Recommendations are that:

- all countries are considered users and providers of data and information. Not only is it important to promote digitization of the largest holdings, but it is fundamental to promote digitization of smaller and perhaps more specialized collections and promote capacity building in countries with rich biological diversity, also in the fields of curatorial practices (to guarantee quality information) and in information and communication technology;

- GBIF promote the concept that information be made freely and quickly available for educational purposes and for scientific research;
- GBIF explore/support the development of tools and standards that are open to different architectures and that guarantee that the ultimate control over data remains in the hands of each provider;
- GBIF explore/support the development of standards and tools to aid in the improvement of the quality and accuracy of the data;
- GBIF must acknowledge, respect and consider different levels of technological development. We recommend that GBIF promote regular meetings with developers of different countries to exchange experience and expertise;
- GBIF could consider a small on-line questionnaire that would be open to any collection to respond to. In this way, once the GBIF Discovery system is available, an automated email could be sent to each of the respondents with details on the Discovery system and on registration;
- GBIF support demonstrative projects that clearly show the benefits of distributed systems, data sharing and collaborative efforts (applications).

Index

Tables

1. Introduction

This report is the result of a commissioned study by GBIF (OCB Work Programme - contract no. GBIFS/2003/04) to analyze experiences on data sharing with countries of origin (a.k.a. repatriation) carried out by the Reference Centre on Environmental Information (CRIA). Activities included: preparation of a questionnaire, selection of institutions/initiatives, survey, analysis and synthesis of replies, preparation of a draft report submitted to the GBIF Secretariat and preparation of the final report based on the comments and suggestions received.

The questionnaire for the study on data sharing with countries of origin was formulated in close collaboration with the GBIF Secretariat, in order to ensure the development of a simple form with a good coverage on content, technology and political framework. This questionnaire was sent to a preliminary list of institutions that was also prepared in consultation with the GBIF Secretariat. This list was further extended in order to include relevant ongoing experiences in different regions of the world. Table 1 has the list of the 27 institutions that the questionnaire was sent to.

Table 1. Institutions that were contacted, their website, and status.

Institution	Web site	Status
1. Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia. **Initiative:** Biodiversity Information System	www.humboldt.org.co/sib	replied
2. Australian National Insect Collection, CSIRO Entomology, Canberra AUSTRALIA	www.ento.csiro.au/	NA*
3. BioMap Project	www.biomap.net/english/whatisbiomap.htm	no reply
4. Botanic Garden and Botanical Museum, Berlin-Dahlem, Germany	ww2.bgbm.org/Herbarium/Default2.cfm	replied
5. CABI Bioscience, UK **Initiative:** The Herb IMI Database – online Website: in preparation	www.cabi-bioscience.org/	replied
6. Centre of Plant Biodiversity Research; Rapid Assessment of Biodiversity Project – World Bank, Department Entomology, Cromwell Road, London,		no reply
7. The National Council for Scientific and Technological Development (CNPq), International Cooperation, Brasília, Brazil **Initiative:** Plants of the Northeast Program.	www.cnpq.br/areas/cooperacaointernacional/programas/pne.htm	NA*
8. Department of Mycology and Phycology, Botanische Staatssamlung Munich, Germany **Initiative:** projects GLOPP and InfoComp funded by the German Ministry for Education and Sciences	www.botanischestaatssammlung.de/projects/coll_online.html	replied
9. Forest Herbarium, Oxford, UK		no reply
10. Herbarium, National Botanic Garden of Belgium	www.br.fgov.be	replied
11. Herbarium, Royal Botanic Gardens, Kew, UK **Initiative:** Repatriation of Herbarium Data for the Flora of Northeastern Brazil	www.rbgkew.org.uk/data/repatbr/homepage.html - in preparation	replied
12. Information Division, National Chemical Laboratory **Initiative:** Repatriated Biological Collections of India, RBCI	www.ncbi.org.in/rbci/	replied
13. Nacional Institute of Biodiversity, INBio, Costa Rica	www.inbio.ac.cr/	NA*
14. Missouri Botanical Garden, USA **Initiative:** TROPICOS & W3TROPICOS	www.mobot.org and www.tropicos.org	replied

Institution	Web site	Status
15. Museum für Naturkunde, Berlin	www.museum.hu-berlin.de	no reply
16. National Commission on Biodiversity, CONABIO, Mexico	www.conabio.gob.mx/	replied
17. National Herbarium Netherlands - Leiden University Branch **Initiative:** South East Asian Botanical Collections Information Network SEABCIN Web site in development	www.nationaalherbarium.nl/rhb/	replied
18. National Museum of Natural History, Smithsonian Institution, USA	www.nmnh.si.edu/	replied
19. New York Botanical Garden	www.nybg.org/	replied
20. Reference Center on Environmental Information CRIA, Brazil **Initiative:** Biota/Fapesp Program	www.cria.org.br/	replied
21. Royal Belgian Institute for Natural Sciences RBINS, Brussels, Belgium		no reply
22. Royal Botanic Gardens Sydney **Initiative:** Repatriation of data to Papua New Guinea	www.rbgsyd.gov.au/	replied
23. Royal Museum for Central Africa, Belgium **Initiative:** African Biodiversity Information Centre, ABIC	www.africamuseum.be	replied
24. The Natural History Museum NHM, London, UK	www.nhm.ac.uk/	replied
25. University of Kansas Biodiversity Research Center, USA **Initiatives:** Specify, Lifemapper, Species Analyst, and NABIN	lifemapper.org	replied
26. Western Australian Museum **Initiative:** Western Australian Museum's FaunaBase and FaunaList	www.museum.wa.gov.au/faunabase	replied
27. Wildlife Trade & Sustainable Fisheries Branch, Department of the Environment and Heritage	www.deh.gov.au/	no reply

* Note: NA = not applicable, refers to institutions that acknowledged having received the questionnaire but stated that they could not fill out the questionnaire as they weren't directly involved in any activity concerning data sharing with countries of origin.

The three institutions that did not fill out the questionnaire were:

- The Instituto Nacional de Biodiversidad (INBio), Costa Rica, answered: "We have not been repatriating information from museums or herbaria, as we have focused on generating information from our own collections. In that sense we try to obtain cooperation from foreign institution via training, identification, etc."
- The Australian National Insect Collection, CSIRO Entomology, Canberra that answered: "we at the Australian National Insect Collection have not been involved in efforts at repatriating data to countries of origin. We are not opposed to doing this, it has just never come up" (sic). We are mainly a collection of Australian insects, and only a small proportion of our holdings is from outside the country (then mostly SE Asian). We only rarely receive any requests for data or information, and the few that we do receive are generally about small amounts of specific information. We do our best to answer these requests when we get them, but we do not get them often."; and,
- The National Council for Scientific and Technological Development (CNPq), International Cooperation. They were contacted because of the repatriation program developed with Great Britain. Their answer about this program was included in the brief profile of the project "Repatriation of Herbarium Data for the Flora of Northeastern Brazil".

The following 18 institutions answered the questionnaire:

1. Missouri Botanical Garden, USA
2. Natural History Museum & Biodiversity Research Center, University of Kansas, USA
3. New York Botanical Garden Herbarium, USA
4. National Museum of Natural History, Smithsonian Institution, USA
5. Herbarium, Royal Botanic Gardens, Kew, UK
6. The Natural History Museum, London
7. CABI Bioscience, UK
8. Botanische Staatssammlung Munich, Germany
9. Botanic Garden and Botanical Museum, Berlin Dahlem, Germany
10. African Biodiversity Information Centre, Belgium
11. National Botanic Garden of Belgium
12. South East Asian Botanical Collections, Information Network - National Herbarium Netherlands-Leiden University Branch, Netherlands
13. Repatriated Biological Collections of India, Information Division, National Chemical Laboratory, India
14. FaunaBase: Western Australian Museum, Australia
15. Royal Botanic Gardens Sydney, Repatriation of data to Papua New Guinea
16. National Commission on Biodiversity, CONABIO, Mexico
17. Alexander von Biological Resources Research Institute, Bogotá, Colombia.
18. Reference Center on Environmental Information - CRIA, Brazil

2. Answers and Analysis

2.1. Brief Profile

Within this group we have four institutions from the US, eight from Europe, one from India, two from Australia, and three from Latin America. The surveyed institutions include important herbaria and natural history museums that curate specimens from practically the whole world. We have initiatives concerned in digitizing their data to help with their own research and to make their curatorial work easier and more efficient, and we have explicit bilateral agreements directly dealing with data sharing with countries of origin. There are institutions concerned in bringing data into the country and maintaining it in a centralized database and others developing distributed systems that will harvest data at remote sites. We believe that this set of institutions and initiatives form a good base of experiences where good examples can be further studied and lessons can be learned.

A brief profile of each institution, based on each initiative's objectives and aims, is presented in table 2 For a more complete assessment please refer to each individual questionnaire (annex) or to the web sites indicated.

Table 2. A brief profile of each institution based on answers from the questionnaire.

Institution	Brief Profile
1. Missouri Botanical Garden, USA	The botanical databases at the Missouri Botanical Garden are maintained to support the research of the institution and it's on going research and training activities in Latin America, Africa, Madagascar and Asia. The main purpose of TROPICOS was education and training to promote scientific research in the countries. Data sharing has always been a primary objective to promote research, conservation, and sustainable use of resources. Countries providing data: Mexico, Honduras, Costa Rica, Panama, Colombia, Ecuador, Peru, Bolivia, Argentina, Chile, Suriname, Venezuela, Paraguay, South Africa, Tanzania, Gabon, Zambia, Madagascar, Vietnam, China.
2. Natural History Museum & Biodiversity Research Center, University of Kansas, USA	The initiatives at the University of Kansas dealing with data-sharing include Lifemapper, Species Analyst, Specify, and NABIN. This is part of a larger international framework, involving capacity building (biodiversity informatics architecture/network), training of students and collection professionals and distributed sharing and integration of biodiversity data. The main purpose of these initiatives is integration, access to, serving, sharing of collection-based biodiversity information distributed in museums and herbaria worldwide, and predictive modeling of biodiversity phenomena based on this biodiversity data. The Species Analyst Network involves more then 15 countries and more then 50 institutions worldwide
3. New York Botanical Garden Herbarium, USA	The New York Botanical Garden Virtual Herbarium's objective is to provide data on-line to improve access to its collections by the worldwide scientific community, to aid its own research programs and to aid the management of the herbarium. Data-sharing is the main purpose of the project.
4. National Museum of Natural History, Smithsonian Institution, USA	The Smithsonian National Museum of Natural History holds one of the largest collections of biological specimens in the world. For many years it has been working to database the collections for its internal uses for science and management, yet only about 10% of the specimens have been databased. With limited exceptions for security and sensitivities, museum information has always been freely available to scientists and policy makers who have visited and worked at the museum Internally the Smithsonian has invested substantial funds and time in moving the databases into an integrated multimedia catalog that will make the data freely available on the internet and accessible through the search processes being developed by GBIF. The main purpose of the project is to make its data freely and widely available, expanding its research capacities and from Congressional mandate.

5. Herbarium, Royal Botanic Gardens, Kew, UK	The name of the project concerned with data-sharing with countries of origin is "Repatriation of Herbarium Data for the Flora of Northeastern Brazil" which is part of the Biodiversity subprogram of the Plants of the Northeast Program (PNE) of Brazil. The present phase of the project aims to repatriate herbarium data on 50% of the Northeastern Brazilian collections found at Kew herbarium in 3 years, in order to facilitate the access of Brazilian scientists to the information contained in the collections. This is a bilateral collaboration between RBG Kew and the *Associação Plantas do Nordeste* (Association Plants of the Northeast) supported by the Brazilian and UK governments, among other partners. It is important to understand that the project on repatriation of herbarium data is part of a broader program that is structuring databases, carrying out surveys and taxonomic studies, and is also carrying out community programs. Information from the Brazilian government (CNPq) states that between 1995 and 2003 R$1.590.897,44 was spent on scholarships and 10 different projects were supported, including a specific project on "Information, Dissemination, and Training".
6. The Natural History Museum, London	The Museum is a major international repository, holding collections of approximately 70 million specimens. Projects undertaken by NHM staff include capacity building, training and information sharing, although the major institutional aim is to carry out research, under the institutions mission statement: "to maintain and develop the collections and use them to promote the discovery, understanding, responsible use and enjoyment of the natural world." The major institutional aim is not data sharing with countries of origin. Projects include: Checklist of amphibians and reptiles of Belize (www.nhm.ac.uk/botany/lascuevas/belize_herplist.html); Birds of Colombia (www.biomap.net/); Plant diversity in Paraguay (internt.nhm.ac.uk/cgi-bin/botany/paraguay/); and Trees of the shade coffee farms of El Salvador (internt.nhm.ac.uk/cgi-bin/botany/estrees/) among many others. The NHM is a leading partner in ENHSIN – European Natural History Specimen Information Network (www.nhm.ac.uk/science/rco/enhsin/index.html) which developed standards for data exchange, and is also involved with BioCISE – Resource Identification for a Biological Collection Information Service in Europe (www.bgbm.fu-berlin.de/biocise).
7. CABI Bioscience, UK	The project is IMI digitization (fungi). Countries providing data are mainly Commonwealth countries and the data has been accumulated over a period of more than 60 years. Data-sharing is the main purpose of the project.
8. Botanische Staatssammlung, Munich, Germany	One of the aims of the projects GLOPP and InfoComp funded by the German Ministry for Education and Sciences was the online presentation of specimen data of large German Natural History Collections. The main purpose is to present this data to support research and services.
9. Botanic Garden and Botanical Museum, Berlin Dahlem, Germany	The initiative is "Digital specimen images at the Herbarium Berolinense" which offers on-line access to high resolution images of its herbarium holdings. Digitizing specimen information is seen as a curatorial activity.

10. National Botanic Garden of Belgium	The initiative is "Prototype Image server to integrate the Martius Herbarium and the Digital Flora Brasiliensis. This project is part of ongoing networking efforts between Brazilian herbaria, North American (especially St. Louis, Madison and New York Botanical Garden), and European institutions (particularly National Botanic Garden of Belgium in Meise, National Herbarium of the Netherlands in Leiden, Herbarium Munich and Royal Botanic Gardens in Kew). Its ultimate goal is to expand the digitization of Martius' collection, already begun at Munich and Leiden, to cover all relevant collections and build transparent links to a number of key illustrated works including Martius' Flora Brasiliensis. The project will contribute to the development of a prototype for structured data management of modern on-line treatments of the revised Flora brasiliensis. The prototype will give access in an electronic form to a selection of reference and type specimens from the Martius's collection. In addition to the specimen and taxonomic databases, access will also be given to digitized herbarium sheets, images of the related plates in Flora brasiliensis, electronic copies of the taxonomic literature and archives of Martius's collection. Brazilian post-doctorate students could participate of the selection and digitization of specimens in Europe. Stakeholders include: ? Countries providing data: Belgium, Germany, The Netherlands, The United States of America ? Institutions providing data: National Botanic Garden of Belgium, Herbarium Munich, National Herbarium of the Netherlands, Missouri Botanical Garden ? Country receiving data: Brazil ? Institutions receiving data: State University of Campinas, Reference Center on Environmental Information (CRIA) ? Country of origin: Herbarium specimens (~ 550 type specimens) were all collected in Brazil in the 1800's. Flora brasiliensis was edited and printed in Germany (1840-1906).

11. African Biodiversity Information Centre, Belgium	The African Biodiversity Information Centre, department of African Zoology, is affiliated to the Royal Museum for Central Africa and in located in Belgium. This is part of a framework agreement with the belgian ministry of development cooperation. The aims and objectives are to provide a structural contribution to the national obligations under the Convention on Biological Diversity concerning co-operation with developing countries. It gives particular emphasis on key activities directly committed by the CBD relating to co-operation in identification, conservation and sustainable use of biodiversity, such as: ? exchange and repatriation of specialised information ? technical and specialised scientific education and training ? human capacity building and institutional strengthening ? research co-operation, exchange of experts ? support for in-situ and ex-situ conservation. Data-sharing is an important component, but not the main purpose; training in the use of the data is a more important component.
12. South East Asian Botanical Collections, Information Network - National Herbarium Netherlands-Leiden University Branch, Netherlands	The network's main aim is to share label information on Southeast Asian plants as part of Flora Malesiana and Flora of Thailand projects. Its goals include collection maintenance (improvement identifications of duplicates) and research (revisions, distribution maps, biodiversity analyses). Its activities are: ? development of a central database, ? fixed data format, ? capacity building through short software courses, ? development of software tools, ? internet publication, ? scale enlargement to include species information, ? information exchange via workshops. The project is being financed by the European Commission (SE Asia IT&C programme).

13. Repatriated Biological Collections of India, NCL Centre for Biodiversity Informatics (NCBI), Information Division, National Chemical Laboratory (NCL), India	The mission of this initiative is "to develop tools and standards and to help improve infrastructure and capacity building to accelerate national progress in collection, collation, analysis, prediction and dissemination of knowledge about Indian biotic resources and its environment to make their sustainable use". Repatriated Biological Collections of India (RBCI) is an effort to form a link between taxonomists dealing with Indian biota and museums that hold the specimens (mostly in Europe and USA). RBCI plans to link information with the other datasets being developed by NCBI and other agencies (e.g. ECATs on Indian Biota, Database on Sacred Groves of India, Conservation Sites in India, data collated using SAMPADA, ecological and genetics datasets, etc.) Acquisition of data on specimens originated from India from the collection facilities / museums abroad is the main objective of this project.
14. FaunaBase: Western Australian Museum, Australia	FaunaBase is a search engine originally developed to make information derived from the Western Australian Museum's vertebrate collections available online. Primarily designed with the general public in mind, FaunaBase also caters for specialists by providing restricted access to records contained in the underlying database. FaunaBase was expanded to include the vertebrate records of the Queensland and Northern Territory Museums on a trail basis to test the ability of the technology to handle larger data sets and still deliver an acceptably fast response. FaunaBase will soon include invertebrate records. It was not developed with being part of a larger framework in mind.
15. Royal Botanic Gardens Sydney, Repatriation of data to Papua New Guinea	This activity is based on the Australia's Virtual Herbarium framework, and involves capacity building for the botanists of the Papua New Guinea National Herbarium (LAE) - providing a tool for readily providing data for conservation and environmental management decisions within their region. The staff will be trained in database, record and content management. Data will be shared between the Australian National Herbarium (CANB), National Herbarium of New South Wales (NSW) National Herbarium of Victoria (MEL), and the Queensland Herbarium (BRI). Initially, the database will be housed off-site (at NSW) and will be access via Internet technology. This will minimise the need for expensive hardware and software. It will also minimise the risks associated with database management for the LAE because the management will be done at NSW.
16. Sistema Nacional de Información sobre Biodiversidad (SNIB), National Commission on Biodiversity, CONABIO, Mexico	SNIB is a law mandated initiative of the Mexican Government that requires CONABIO to create a permanently updated biodiversity information system to support research for its growth and to use it to provide advice to all sectors of Mexican society. Work includes supporting projects, promoting and creating standards, capacity building, data compilation, data sharing and data distribution.

17. Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia	The initiative is "Sistema de Información sobre Biodiversidad" or The Biodiversity Information System (BIS). It was established and co-ordinated in Colombia by the Humboldt Institute, in association with other organizations from public and private sector (including research institutes, universities, NGOs, biological collections, etc.). The BIS is part of the Environmental Information System of Colombia, a big initiative managed directly by the Ministry of Environment, Housing and Territorial Development. The BIS promotes data and information sharing, by providing technical guidelines and tools, as well as capacity building and trainning among participants. Information sharing process under BIS scheme are not completely implemented at this moment, because designing of tools and policies are still in progress. This initiative has been established to facilitate information sharing among national institutions. Exchange process with other countries are not a main purpose of the BIS, but it is considered in terms of interoperability tools under implementation. As to stake-holders, the institution that is providing data is the Missouri Botanical Garden, USA.
18. Reference Center on Environmental Information - CRIA, Brazil	CRIA is a not-for-profit private organization that aims at disseminating electronic information as a tool for the organization of the scientific and technological community of the country. It disseminates biological information of environmental and industrial interest and through this, hopes to contribute directly to the conservation and sustainable use of Brazil's biological resources. The initiatives included in this survey are part of the São Paulo State Program "Biota/Fapesp The Virtual Biodiversity Institute". They are speciesLink, a distributed information network of biological collections (splink.cria.org.br/) and SinBiota (sinbiota.cria.org.br/atlas), an information system for the program's field surveys. CRIA works with biodiversity informatics.

2.2. Comment on Aims

The questionnaire asked whether data-sharing with the originating country was the main purpose of the project. Answers are presented in Table 3.

Table 3. Aims: questions and answers

Institution	Was data-sharing with the original country the main purpose of the project?		If not, what was the main purpose of the project?
	yes	no	
Missouri Botanical Garden, USA		x	The main purpose was education and training to promote scientific research in the countries. Data sharing has always been a primary objective to promote research, conservation, and sustainable use of resources.
Natural History Museum & Biodiversity Research Center, University of Kansas, USA	x	x	Integration, access to, serving, sharing of collection-based biodiversity information distributed in museums and herbaria worldwide; predictive modeling of biodiversity phenomena based on this biodiversity data.
New York Botanical Garden Herbarium, USA	x		
National Museum of Natural History, Smithsonian Institution, USA		x	Making our data freely and widely available, expanding our research capacities and from Congressional mandate.
Herbarium, Royal Botanic Gardens, Kew, UK	x		
The Natural History Museum, London		x	The major institutional aim is not data sharing with countries of origin, although this can be a, or the, major aim of projects carried out by our staff…
CABI Bioscience, UK	x		
Botanische Staatssammlung Munich, Germany		x	Present collection data of our large natural history collection to support research and services
Botanic Garden and Botanical Museum, Berlin Dahlem, Germany		x	Digitizing specimen information is seen as a curatorial activity.
National Botanic Garden of Belgium	x		
African Biodiversity Information Centre, Belgium	x	x	Data-sharing is an important component, but not the main purpose; training in the use of the data is a more important component (i.e. no data dumps without further training or follow-up.
South East Asian Botanical Collections, Information Network - National Herbarium Netherlands-Leiden University Branch, Netherlands	x		
Repatriated Biological Collections of India, NCL Centre for Biodiversity Informatics (NCBI), Information Division, National Chemical Laboratory (NCL), India	x		Acquisition of data on specimens of Indian origin from the collection facilities / museums abroad is the main objective of this project.
FaunaBase: Western Australian Museum, Australia			
Royal Botanic Gardens Sydney, Repatriation of data to Papua New Guinea	x		Yes but capacity building was also a major driver of this project
Sistema Nacional de Información sobre Biodiversidad (SNIB), National Commission on Biodiversity, CONABIO, Mexico	x		
Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia		x	This initiative has been established to facilitate information sharing among national institutions. Exchange process with other countries is not a main purpose of the BIS, but it is considered in terms of interoperability tools under implementation.

	Was data-sharing with the original country the main purpose of the project?		
Reference Center on Environmental Information - CRIA, Brazil	x		Data sharing is the main purpose of the project. But more then data sharing we hope to contribute to promote collaborative research. The system is also integrated with other networks such as Species Analyst.
Total	**11**	**8**	

Only one institution did not answer. Of the total of 8 institutions that answered that data-sharing was not the main purpose of the project, 2 answered both "yes" and "no", and of the 6 that just answered "no", 5 specifically state data-sharing as an important component. Basically, most projects are making information generally available or, in other words are actually sharing their data on the Internet. Data sharing with country of origin is seen as a valuable spin-off, and the rational is that by making the information freely available, it becomes available not only to the countries of origin, but to anyone else who needs or can benefit from such access.

2.3. Strategy

The following aspects related to data sharing strategies were analyzed:

- criteria for selection of taxonomic group;
- criteria for selection of collection;
- priorities addressed; and,
- user groups for shared data. Tables 4-7 present a summary of the answers given.

The criteria "selection of taxonomic groups" shows interesting answers from institutions working with information systems. While herbaria and natural history museums set as a priority recent taxonomic revision, recently curated, and as "other", the presence of research groups, institutions such as CONABIO, CRIA, and the Humboldt Institute in Colombia, concerned with developing information systems, include as "other" the willingness to share data and information and the availability of data in an electronic format.

The most important criteria for the selection of a collection, based on the answers given are taxonomic and regional scope answered respectively by 72% and 78%. Here again, the institutions closely involved with the development of information systems, University of Kansas, CRIA and CONABIO all include willingness to participate as a key factor.

As to priorities addressed, once again taxonomic, included by 15 of the 18 institutions and conservation, 12 out of 18, were the most important criteria. Policy-making is included by 7 institutions which may indicate the increasing concern to make quality information a basis for the elaboration of policies.

Finally, as to user groups, as was expected, **all** institutions included the scientific community as a target user. It may have been interesting to have requested that a priority list be made in order to be able to analyze in which order these users are placed for each institution. Policy makers and the "general public" were included by 61% of the institutions followed closely by Educators with 56%. Only 5 of the 18 institutions (28%) included private companies as target users.

Table 4. Criteria for selection of taxonomic group

Institution	Economic Importance	Recent Taxonomic Revision	Biogeographic Study	Conservation Priority	Recently Curated	Phylogenetic Study	Endemism	Historic sources of data	other
Missouri Botanical Garden		1	1	1	1	1		1	
Natural History Museum & Biodiversity Research Center, University of Kansas	1		1	1	1				disease, invasive species, pest
New York Botanical Garden Herbarium							1		groups/areas currently the focus of research programs at NYBG and elsewhere
National Museum of Natural History, Smithsonian Institution		1	1		1	1		1	
Herbarium, Royal Botanic Gardens, Kew	1		1	1	1		1	1	Presence of specialist on staff or group of collaborating specialists
The Natural History Museum	1	1	1	1	1	1	1	1	Because of the wide variety of staff and reasons for projects, all of the alternatives apply to activities undertaken by the NHM. There is no institutional policy that would necessarily prioritize these
CABI Bioscience								1	
Botanische Staatssammlung München	1	1		1					
Botanic Garden and Botanical Museum Berlin Dahlem					1				Loans
National Botanic Garden of Belgium	1	1							The herbarium material will be selected within 8 pilot plant groups that are currently the focus of research of taxonomists at the universities of São Paulo and Rio de Janeiro. These groups are the most feasible ones for a prototype image server and will result in data of immediate interest for the users.
African Biodiversity Information Centre	1	1	1	1		1	1	1	

Institution	Economic Importance	Recent Taxonomic Revision	Biogeographic Study	Conservation Priority	Recently Curated	Phylogenetic Study	Endemism	Historic sources of data	other
South East Asian Botanical Collections Information Network - National Herbarium Netherlands-Leiden University Branch		1	1	1					
Repatriated Biological Collections of India Information Division, National Chemical Laboratory								1	All taxonomic groups are selected
FaunaBase: Western Australian Museum		1					1	1	Fully databased and most complete taxonomy
National Commission on Biodiversity, CONABIO	1			1	1				Willingness of curators to allow access to their data
Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia			1		1				Data electronically available
Centro de Referência em Informação Ambiental - CRIA									Willingness to database information and to share data
Total	7	8	8	8	8	4	5	8	
% of possible total	39	44	44	44	44	22	28	44	

Table 5. Criteria for Selection of Collection

Institution	Types	Ecological	Collectors	Taxonomic	Expeditions	Regional Scope	Other
Missouri Botanical Garden			1	1	1	1	
Natural History Museum & Biodiversity Research Center, University of Kansas				1		1	All collections willing to participate in species analyst network
New York Botanical Garden Herbarium					1	1	
National Museum of Natural History, Smithsonian Institution	1			1		1	
Herbarium, Royal Botanic Gardens, Kew	1		1	1	1	1	
The Natural History Museum, London	1	1	1	1	1	1	All of these apply. Institutionally we favor data basing types and particular collections if significant
CABI Bioscience						1	

Institution	Types	Ecological	Collectors	Taxonomic	Expeditions	Regional Scope	Other
Botanische Staatssammlung München	1	1		1			
Botanic Garden and Botanical Museum Berlin Dahlem							
National Botanic Garden of Belgium	1					1	Martius' specimens collected in Brazil by himself (1817-1821) and others purchased by him are historic sources of data.
African Biodiversity Information Centre				1		1	
South East Asian Botanical Collections Information Network - National Herbarium Netherlands-Leiden University Branch	1			1		1	
Repatriated Biological Collections of India Information Division, National Chemical Laboratory	1	1	1	1	1	1	
FaunaBase: Western Australian Museum				1			
Royal Botanic Gardens Sydney, Repatriation of data to Papua New Guinea						1	
National Commission on Biodiversity, CONABIO				1			Good quality and willingness to participate
Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia				1	1	1	
Centro de Referência em Informação Ambiental - CRIA	1			1	1	1	willingness to participate
Total	8	3	4	13	7	14	
% of possible total	44	17	22	72	39	78	

Table 6. Priorities Addressed

Institution	Taxonomic	Conservation	Policy-making	socio-economic	political	specify other
Missouri Botanical Garden	1	1	1			
Natural History Museum & Biodiversity Research Center, University of Kansas	1	1	1			disease; invasive species; pests
New York Botanical Garden Herbarium	1	1				
National Museum of Natural History, Smithsonian Institution	1					
Herbarium, Royal Botanic Gardens, Kew	1	1	1			
The Natural History Museum, London	1	1				
CABI Bioscience			1	1		invasives, quarantine, natural enemies
Botanische Staatssammlung München	1					
Botanic Garden and Botanical Museum Berlin Dahlem						administrative
National Botanic Garden of Belgium	1					Returning biodiversity data to their country of origin
African Biodiversity Information Centre	1	1	1	1		
South East Asian Botanical Collections Information Network - National Herbarium Netherlands-Leiden University Branch	1	1				
Repatriated Biological Collections of India Information Division, National Chemical Laboratory	1	1				
FaunaBase: Western Australian Museum	1					
Royal Botanic Gardens Sydney, Repatriation of data to Papua New Guinea	1	1	1			
National Commission on Biodiversity, CONABIO		1				Vertebrates and Plants tend to be useful
Alexander von Humboldt Biological Resources Research Institute, Bogotá,	1	1				

Institution	Taxonomic	Conservation	Policy-making	socio-economic	political	specify other
Colombia						
Centro de Referência em Informação Ambiental - CRIA	1	1	1			
Total	15	12	7	2	0	
% of possible total	83	67	39	11	0	

Table 7. User groups for shared data

Institution	Policy-makers	Scientific Community	Educators	General Public	Private Companies	Specify Other
Missouri Botanical Garden	1	1	1		1	
Natural History Museum & Biodiversity Research Center, University of Kansas	1	1	1	1		
New York Botanical Garden Herbarium		1				
National Museum of Natural History, Smithsonian Institution	1	1	1	1		
Herbarium, Royal Botanic Gardens, Kew		1				
The Natural History Museum		1		1		One initiative -Worldmap- is specifically targeted at conservation issues for policy-makers (www.nhm.ac.uk/science/projects/worldmap/index.html)
CABI Bioscience	1	1			1	
Botanische Staatssammlung München		1		1		
Botanic Garden and Botanical Museum Berlin Dahlem		1				
National Botanic Garden of Belgium		1		1		wide public interested in historical collections and archives
African Biodiversity Information Centre		1	1			Conservation NGOs
South East Asian Botanical Collections Information Network - National Herbarium Netherlands-Leiden University Branch	1	1	1	1	1	
Repatriated Biological Collections of India Information Division, National Chemical Laboratory	1	1	1	1		
FaunaBase: Western Australian Museum	1	1	1	1	1	
Royal Botanic Gardens Sydney, Repatriation of data to Papua New Guinea	1	1	1	1		
National Commission on Biodiversity, CONABIO	1	1	1	1	1	
Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia	1	1				
Centro de Referência em Informação Ambiental - CRIA	1	1	1	1		
Total	**11**	**18**	**10**	**11**	**5**	
% of possible total	**61**	**100**	**56**	**61**	**28**	

2.4. Contractual Arrangements

Table 8 shows the result for the type of contractual arrangement used to facilitate data-sharing procedures. It is interesting to note that not one institution included a formal contract, making it clear that data-sharing is still very informal and most initiatives are probably carried out as a collaborative effort through informal agreements.

Table 8. Type of contractual arrangement established to facilitate the data-sharing procedures

Institution	Formal Contract	MoU	Letter of Agreement	Verbal Agreement	Specify Other
Missouri Botanical Garden		1	1	1	
Natural History Museum & Biodiversity Research Center, University of Kansas			1	1	
New York Botanical Garden Herbarium					informal agreements
National Museum of Natural History,		1		1	
Smithsonian Institution					
Herbarium, Royal Botanic Gardens, Kew		1		1	
The Natural History Museum		1	1	1	
CABI Bioscience					
Botanische Staatssammlung München					
Botanic Garden and Botanical Museum Berlin Dahlem			1		
National Botanic Garden of Belgium				1	Flora Brasiliensis volumes are no more subject to copyright
African Biodiversity Information Centre			1	1	
South East Asian Botanical Collections Information Network - National Herbarium Netherlands-Leiden University Branch		1			
Repatriated Biological Collections of India Information Division, National Chemical Laboratory			1		We usually do email correspondence and get formal consent to incorporate their datasets
FaunaBase: Western Australian Museum			1	1	
Royal Botanic Gardens Sydney, Repatriation of data to Papua New Guinea					
National Commission on Biodiversity, CONABIO		1			
Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia			1	1	
Reference Center on Environmental Information - CRIA			1	1	informal agreements
Total	0	6	9	10	
% of possible total	0	33	50	55	

Table 9 shows the answers as to how Intellectual Property Rights are being dealt with. Although this was a free text, we structured a table with the most frequent answers and it is clear that for these institutions ownership is retained with the data custodian who should be acknowledged.

The control of data custodians over their own data, thanks to the evolution of information and communication technology enables them to withhold sensitive or not validated data. This concept is very important. For many years, the fact that sensitive data does exist and that validating is fundamental would be the main reason for custodians to hold complete datasets. Now it is possible/feasible to select what will be or not made publicly, fully or partially available.

Intellectual Property Rights issues didn't come across as the 'number 1' issue of concern for most institutions. It is important to bear in mind that only 30% of the institutions that answered this questionnaire indicated private companies as their target users. So the answers we received may be reflecting this. For those institutions that answered the questionnaire, it appears that issues that are important are proper attribution or credit, custodianship (i.e. each contributing museum retains ownership of its records), acknowledgement and control over data.

Table 9. How IPR issues are dealt with.

Institution	not formally addressed	free flow for scientific purposes	data access policy / disclaimer, info online	clearly marking with logo	acknowledgement provider, provider retains ownership	respect for Mexican Law	Time is given for publishing before data becomes public	Providers may withhold fields, records	at an ad hoc basis
Missouri Botanical Garden	1	1							
Natural History Museum & Biodiversity Research Center, University of Kansas		1							
New York Botanical Garden Herbarium	1								
National Museum of Natural History, Smithsonian Institution			1				1		
Herbarium, Royal Botanic Gardens, Kew				1					
The Natural History Museum, London									1
CABI Bioscience	1								
Botanische Staatssammlung München			1						
Botanic Garden and Botanical Museum Berlin Dahlem			1						
National Botanic Garden of Belgium									
African Biodiversity Information Centre			1		1				
South East Asian Botanical Collections Information Network - National Herbarium Netherlands - Leiden University Branch					1				
Repatriated Biological Collections of India Information Division, National Chemical Laboratory					1				
FaunaBase: Western Australian Museum					1				
National Commission on Biodiversity, CONABIO					1	1	1	1	
Alexander von Humboldt Biological Resources Research Institution		1			1				
Centro de Referência em Informação Ambiental - CRIA					1		1	1	
Total	3	3	4	1	7	1	3	2	1
% of possible total	17	17	22	6	39	6	17	11	6

2.5. Receiving Institution / Country Contribution

To the question on how the receiving institution or country contributed with the definition of the strategy, selection of data and format, transfer of methodologies, and in training. It was made clear that this depended on the project and could extend to limited or no involvement to full partnership.

The answers were:

- Missouri Botanical Garden indicated that it works closely with scientists to determine what information they require and what format is appropriate for their use. It also provides access to data transfer on its web site, allowing both input of new data and output to the users of stored information.
- Natural History Museum & Biodiversity Research Center, University of Kansas stated that some countries (e.g., Brazil; Mexico; Canada) play a critical role in the design of informatics architecture for data access and predictive modeling tools.
- New York Botanical Garden Herbarium: didn't answer
- National Museum of Natural History, Smithsonian Institution stated that there are a set of agreements focusing on taxa of particular interest to the country
- Herbarium, Royal Botanic Gardens, Kew indicated that it is a bilateral agreement, which means that the receiving country took part in writing and revising the proposal and in evaluating the results
- The Natural History Museum, London indicates that this will depend on the project under consideration and can extend for limited or no involvement to full partnership, with the latter increasingly being the norm, and partner countries having a great say in project outputs.
- CABI Bioscience: didn't answer
- Botanische Staatssammlung München: didn't answer
- Botanic Garden and Botanical Museum Berlin Dahlem: didn't answer
- National Botanic Garden of Belgium answered in separate blocks:

  Selection of herbarium material: 8 pilot plant groups that are currently the focus of research of taxonomists at the universities of São Paulo and Rio de Janeiro were selected by Brazilian partners.

  Data quality control and updating of taxonomy: expertise of Brazilian scientific staff will guarantee that the taxonomic requirements are met and taxonomy is updated. Brazilian partners will bring concepts up to date with current taxonomy.

  Digitization of plates in Flora brasiliensis: methodology of Brazilian partners (in collaboration with Missouri Botanical Garden and University of Wisconsin) in carrying out “Flora brasiliensis on-line" will be adopted.
- African Biodiversity Information Centre indicated that it is demand driven by requests from receiving institutions
- South East Asian Botanical Collections Information Network - National Herbarium Netherlands-Leiden University Branch indicated that various herbaria of various countries cooperate in this project. Decisions such as selection of test taxa and data format were decided at a workshop.
- Repatriated Biological Collections of India Information Division, National Chemical Laboratory indicated that they are keeping the database flexible to accommodate all types of data.
- FaunaBase: Western Australian Museum indicated that they promote consultation with partners.
- Royal Botanic Gardens Sydney, Repatriation of data to Papua New Guinea indicated that the Papua New Guinea National Herbarium's staff is actively involved with the design of the Papua New Guinea database and printed outputs. Training program is yet to be developed.

- National Commission on Biodiversity, CONABIO indicated that this varies.
- Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia stated that the main contribution proposed to the Missouri Botanical Garden was the completion of geographic coordinates for individual records.
- Reference Center on Environmental Information - CRIA: indicated that the institution carries out collaborative work with developers (tools, protocols, standards) within the country and from abroad. International developments today include DiGIR and Predictive Modeling. With data providers, collaborative work includes defining data field standards and web outputs.

It is important to bear in mind that in taxonomy, collaboration among specialists is a necessity. The proximity of nations due to the advance of communication technology is a fact. Therefore we believe that the tendency is, aside from political and biosafety reasons, towards the enhancement of collaborative efforts.

2.6. Process

The answers to the questions about the process used for data-sharing are indicated in table 10 – type of data exchanged; table 11 – data format; table 12 – data transfer method; and table 13 – exchange format and standards.

As to the type of data (table 10) practically all institutions, with the exception of one, included "label data". All, with the exception of two, included "location information" and 15 out of 18 included "Geocoding". It is important to note that these represent the fundamental data necessary for predictive distribution modeling. It is also significant that "image" appears in 61% of the institutions.

Table 10. Type of data exchanged/transferred

Institution	label data	location information	Geo-coding	Collecting details	repository and storage	image	video	sound	specify other
Missouri Botanical Garden	1	1	1	1	1	1			
Natural History Museum & Biodiversity Research Center, University of Kansas	1	1	1		1				
New York Botanical Garden Herbarium	1	1	1			1			
National Museum of Natural History, Smithsonian Institution	1	1	1	1		1		1	
Herbarium, Royal Botanic Gardens, Kew	1		1		1	1			additional taxonomic data are provided by specialist and incorporated in the database
The Natural History Museum, London	1	1	1	1	1	1			data exchanged depend on the project goals, but all forms of data have been exchanged to some extent.
CABI Bioscience	1	1	1						
Botanische Staatssammlung München									
Botanic Garden and Botanical Museum Berlin Dahlem	1	1		1		1			Most data given within picture (label, annotations)
National Botanic Garden of Belgium	1	1	1	1	1	1			
African Biodiversity Information Centre	1	1	1	1	1	1		1	
South East Asian Botanical Collections Information Network - National Herbarium Netherlands-Leiden University Branch	1	1	1	1		1			
Repatriated Biological Collections of India Information Division, National Chemical Laboratory	1	1	1	1	1	1			
FaunaBase: Western Australian Museum	1	1	1	1					
Royal Botanic Gardens Sydney, Repatriation of data to Papua New Guinea	1	1	1	1					
National Commission on Biodiversity, CONABIO	1	1	1	1		1			
Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia	1	1		1					
Centro de Referência em Informação Ambiental - CRIA	1	1	1	1					
Total	17	16	15	13	7	11	0	2	
% of possible total	94	89	83	72	39	61	0	11	

Table 11 shows that there is no predominant format for data exchange, whereas table 12 shows that "interactive web search" was answered by over 70% (13 institutions) followed by CD-ROM 50% (9 out of 18 institutions).

Table 11. Data format

Institution	txt	dbf	mdb	xls	xml	spreadsheet	word	html	jpg	as required
Missouri Botanical Garden	1	1			1					
Natural History Museum & Biodiversity Research Center, University of Kansas										
New York Botanical Garden Herbarium	1	1	1	1						
National Museum of Natural History, Smithsonian Institution						1				
Herbarium, Royal Botanic Gardens, Kew		1								
The Natural History Museum, London		1								
CABI Bioscience										1
Botanische Staatssammlung München			1							
Botanic Garden and Botanical Museum Berlin Dahlem									1	
National Botanic Garden of Belgium		1			1					
African Biodiversity Information Centre										1
South East Asian Botanical Collections Information Network - National Herbarium Netherlands-Leiden University Branch		1								
Repatriated Biological Collections of India Information Division, National Chemical Laboratory	1		1	1			1	1		
FaunaBase: Western Australian Museum			1	1						
Royal Botanic Gardens Sydney, Repatriation of data to Papua New Guinea										
National Commission on Biodiversity, CONABIO										1
Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia	1			1						
Centro de Referência em Informação Ambiental - CRIA	1	1			1	1	1	1	1	1
Total	5	7	4	4	3	2	2	2	2	5
% of possible total	28	39	22	22	16	11	11	11	11	28

Table 12. Data Transfer Method

Institution	cd	email	internet	interactive web search	ftp	printouts	as appropriate
Missouri Botanical Garden	1	1			1		
Natural History Museum & Biodiversity Research Center, University of Kansas				1			
New York Botanical Garden Herbarium	1	1	1	1			
National Museum of Natural History, Smithsonian Institution	1	1	1	1			
Herbarium, Royal Botanic Gardens, Kew	1			1		1	
The Natural History Museum, London	1			1			
CABI Bioscience							1
Botanische Staatssammlung München							
Botanic Garden and Botanical Museum Berlin Dahlem	1	1	1	1			
National Botanic Garden of Belgium	1		1	1			
African Biodiversity Information Centre	1	1	1				
South East Asian Botanical Collections Information Network - National Herbarium Netherlands-Leiden University Branch				1			
Repatriated Biological Collections of India Information Division, National Chemical Laboratory		1	1	1			
FaunaBase: Western Australian Museum				1			
Royal Botanic Gardens Sydney, Repatriation of data to Papua New Guinea				1			
National Commission on Biodiversity, CONABIO	1			1			
Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia		1					
Centro de Referência em Informação Ambiental - CRIA				1			
Total	9	7	6	13	1	1	1
% of possible total	50	39	33	72	6	6	6

Table 13 shows exchange formats and standards. XML seems to have become a standard as 9 out of 18 (50%) indicated it. It is interesting to see DiGIR, which unfortunately was not included as an option by the questionnaire, being mentioned by 4 institutions, and this way becoming the second most used standard, ahead of Z39.50.

Table 13. Exchange formats and standards

Institution	xml	Z39.50	hispid	corba	DiGIR	Remib	Brahms	TDWG	user defined	TCP/IP sockets
Missouri Botanical Garden	1				1	1			1	
Natural History Museum & Biodiversity Research Center, University of Kansas	1	1			1					
New York Botanical Garden Herbarium	1		1							
National Museum of Natural History, Smithsonian Institution	1									
Herbarium, Royal Botanic Gardens, Kew			1					1		
The Natural History Museum, London										
CABI Bioscience										
Botanische Staatssammlung München	1									
Botanic Garden and Botanical Museum Berlin Dahlem	1									
National Botanic Garden of Belgium	1						1	1		
African Biodiversity Information Centre										
South East Asian Botanical Collections Information Network - National Herbarium Netherlands-Leiden University Branch					1		1			
Repatriated Biological Collections of India Information Division, National Chemical Laboratory	1	1	1							
FaunaBase: Western Australian Museum										
Royal Botanic Gardens Sydney, Repatriation of data to Papua New Guinea			1							
National Commission on Biodiversity, CONABIO										1
Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia										
Centro de Referência em Informação Ambiental - CRIA	1				1					
Total	9	2	4	0	4	1	2	2	1	1
% of possible total	50	11	22	0	22	6	11	11	6	6

The last aspect analyzed within the process framework was transfer of technology specifically requesting information on software and expertise (table 14) If one excludes the three institutions that did not answer, we have 8 institutions indicating that they transfer software (44% of those that answered) and 7 institutions indicating that they offer training (39%).

Table 14. Technology Transfer

Institution	software	training	other
Missouri Botanical Garden		bring overseas scientists and students for training (databases and GIS). Our researchers and in-country personnel are also available to provide expertise and training as needed	pcs and other computer equipment, help with proposals
Natural History Museum & Biodiversity Research Center, University of Kansas	desktop GARP, Specify, Species Analyst Architecture		
New York Botanical Garden Herbarium	KE Meu		
National Museum of Natural		formal training and fellowship programs	

Institution	software	training	other
History, Smithsonian Institution			
Herbarium, Royal Botanic Gardens, Kew		access, imaging software, and procedures; specific knowledge in collection curation, care and repatriation; nomenclature; taxonomic scientific writing, bibliography search, reporting, etc.	
The Natural History Museum, London			databasing techniques, data capture techniques
CABI Bioscience		introduction to general functionality of databases and web servers, web site development	hardware - several working PC´s; basic networking equipment
Botanische Staatssammlung Munich	we develop in the context of the German GBIF node for mycology software modules of the Diversity Workbench which will be freely available		
Botanic Garden and Botanical Museum Berlin Dahlem	cold fusion server; trueSpectra Image Server, Macromedia		ABCD data exchange standard, BioCASE protocol
National Botanic Garden of Belgium	DiGIR, BioCASE provider		
African Biodiversity Information Centre		training where necessary in required programs (mostly MS Office)	
South East Asian Botanical Collections Information Network - National Herbarium Netherlands-Leiden University Branch	Brahms (Oxford) including special programming for individual herbaria	courses in data entry and querying	
Repatriated Biological Collections of India Information Division, National Chemical Laboratory			
FaunaBase: Western Australian Museum			
National Commission on Biodiversity, CONABIO	Mallos and gregalis, registered		
Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia			
Centro de Referência em Informação Ambiental - CRIA	Collaborative development: DiGIR, modeling, on-line mapping tools	training in the use of collection management software (Brahms, Biota, etc.); modeling; demonstrative talks about DiGIR	installation software, hardware, network at collections; data import for collections

2.7. Funding, Time Scale, and Costs

Table 15 shows the answers as to funding. It is clear that Project Grants and Fellowships (72%) and Institutional Funding (67%) are the most significant items. It would be important to know what percentage comes from the institute and from projects. It is clear that the importance of data-sharing and the necessity of having funds for digitizing scientific data should be stressed to include this within the strategy of research funding agencies.

Table 15. Funding

Institution	Project Grants & Fellowships	Internal Institutional Funding	Contribution from Country Receiving Data	Elaborate
Missouri Botanical Garden	1	1		
Natural History Museum & Biodiversity Research Center, University of Kansas	1			almost all development funded by US Agency research grants
New York Botanical Garden Herbarium	1	1		funding for hardware, software, website, imaging labor and equipment, data editing and maintenance has all been borne by NYBG; funding from National Science Foundation for data entry

Institution	Project Grants & Fellowships	Internal Institutional Funding	Contribution from Country Receiving Data	Elaborate
National Museum of Natural History, Smithsonian Institution	1	1	1	
Herbarium, Royal Botanic Gardens, Kew	1	1	1	Darwin Initiative Grant was used as a pilot, money raised specifically by Kew Foundation was used for the 2nd phase, fellowships are used to bring specialists over and for workshops, researchers' time and manager's time and IT personnel's time, as well as hardware and ac commodation is obtained from internal funding, repatriation officers obtain paid leave from the Brazilian institute sending them to Kew, staff time from Brazilian collaborators.
The Natural History Museum, London	1	1	1	All forms of funding have their place, but increasingly the NHM is dependent on external funding to be able to mobilise data. The availability of such funding will also assist in determining priorities
CABI Bioscience				no formal project funded yet - work to date in association with other projects
Botanische Staatssammlung Munich	1			we have a several cooperating projects funded by the Bavarian State Ministry, German State Ministry, etc.
Botanic Garden and Botanical Museum Berlin Dahlem		1		
National Botanic Garden of Belgium	1			Funding was obtained from ENBI-WP13 "Making non-European biodiversity data in European repositories globally available" to produce a prototype in the frame of a feasibility study
African Biodiversity Information Centre	1	1		
South East Asian Botanical Collections Information Network - National Herbarium Netherlands-Leiden University Branch	1	1		50% of costs covered by European Commission, rest by Institutes
Repatriated Biological Collections of India Information Division, National Chemical Laboratory		1		as part of the biodiversity informatics activities we felt it necessary to acquire data on specimens of Indian origin which could be linked with other species genetic, ecological datasets being developed by the NCL Centre for biodiversity informatics
FaunaBase: Western Australian Museum		1		
Royal Botanic Gardens Sydney, Repatriation of data to Papua New Guinea	1	1	1	Costs of current project: GBIF grant funding 70% of cost, NSW 20%, LAE 10%, in kind costs prior to this project are considerable, estimated to be approximately 11 times cost of current project
National Commission on Biodiversity, CONABIO		1		
Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia				
Reference Center on Environmental Information - CRIA	1	1		Largest support from São Paulo State's research grant agency (Fapesp), and some from the Ministry of Science and Technology.
Total	12	13	4	
% of total	67	72	22	

The next items that were analyzed were time scale and costs (Table 16). Of these institutions, New York Botanical Garden and CONABIO are the "oldest" concerned with data-sharing and digitizing and the time scale in this case is of several years. This makes it very clear that digitization is a recent activity and data-sharing at this level of integration only became possible with the Internet and due to IT developments that followed.

As to costs, it is interesting to compare answers of those few that have a clear idea of time and costs per record. Answers vary between 40 seconds and 30 minutes per record and costs vary

between US$ 0.40 to US$ 12.70. Both time and cost per record would obviously require a more detailed analysis to verify what each provided answer entails.

Start up costs are also very difficult to compare as the answers are not uniform. Most institutional costs are not considered and answers as to hardware and software are also vague. We believe that the questionnaire would require more details probably followed by personal visits/interviews with selected institutions to evaluate start-up and maintenance costs.

Table 16. Funding, Time Scale, and Costs

Institution	Time Scale	start-up costs	time per record	cost per record	other comments
Missouri Botanical Garden	on-going	computers, printers, software			because of the careful agreements and collaborative research between the Garden and overseas institutions we are able to agree on the methods and formats of the data transfers resulting on an efficient and cost effective data transfer protocol
Natural History Museum & Biodiversity Research Center, University of Kansas	past 5 yrs	US$ 500,00 to start current research funding @ US$ 3 million			
New York Botanical Garden Herbarium	first data posted on line in 1996; NYBG Virtual herbarium web site established in current form in 1998	software, hardware, approximately US$ 500,000	6 min	US$ 1.92	
National Museum of Natural History, Smithsonian Institution	on-going				
Herbarium, Royal Botanic Gardens, Kew	attempting to complete 50% in 5 years, i.e. 10% of the area's flora deposited at Kew per year	Kew provided hardware, software as well as technicians to train officers in imaging	10 min to database and image each specimen	£7.50 per specimen with image	
The Natural History Museum, London	depending on project, but generally up to about 3 years	almost inevitably hardware part of start-up costs, as will employment of additional staff	variable by taxon and data required	variable by taxon and data required	The NHM is seeking to database all its specimens, but requires external partners and funding in order to be able to do this
CABI Bioscience	1-3 years		40 secs	US$ 0.40	
Botanische Staatssammlung Munich	GLOPP and InfoComp were funded between 2001 and 2003; other projects like the German GBIF node for mycology will be funded until 2005		10-15 min		
Botanic Garden and Botanical Museum Berlin Dahlem	on-going	US$ 30,000 excl. personnel	6 min for digital specimen with country and name record	US$ 3	
National Botanic Garden of Belgium	10 months	scanner, pc & software: 8500 euros	30 min	10 euros	
African Biodiversity Information Centre	since 1999	absorbed by institutional funds			efficiency of data transfer is not a major concern. Focus is more on training the receiving person in the data use and interpretation

Institution	Time Scale	start-up costs	time per record	cost per record	other comments
South East Asian Botanical Collections Information Network - National Herbarium Netherlands-Leiden University Branch	2.5 years	variable per country in principal 3-4 computers per institute and budget of data typists on a temporary basis	3 min data entry no validation	dependent on country	
Repatriated Biological Collections of India Information Division, National Chemical Laboratory	3 months, began in July 2003	software and hardware US$ 30,000	500 records per day as most is in electronic format	US$ 1	looking forward to develop collaborative programs for estimating size of Indian collections in museums abroad and also begin the process of their automation and digitization.
FaunaBase: Western Australian Museum					
National Commission on Biodiversity, CONABIO	variable	part of a several years institutional effort	1 year per database. Average 10K to 100 K records per database	US$ 1-10	
Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia					
Centro de Referência em Informação Ambiental - CRIA	on-going	US$ 400,000 which included basic hardware & software for each collection (12). Cost of digitizing and validating data not included. Average yearly cost of US$ 300,000			Fapesp also had a specific program for infrastructure for the collections

2.8. Results and Products

Answers requested under "Results and Products" included "taxa covered", "how many", "tools", "tool sharing", "data sharing on the Internet" and the "benefits of the project". Table 17 shows the answers given.

Table 17 Results and Products

Institution	Taxa covered	How many	Tools	Tool Sharing	Data available on the Internet	URL	Project Benefits
Missouri Botanical Garden	Vascular plants and bryophytes	Names: 906,762 Specimens: 1,750,039 References: 87,324 Types: 291,915	TROPICOS system contains html forms for data input and user selected output	yes	yes	www.tropicos.org	1. rapid flow of information for research and conservation 2. improved reporting of results to collaborating institutions 3. expanded data set for research 4. larger data set for web display
Natural History Museum & Biodiversity Research Center, University of Kansas	all vertebrates, insects, plants	SpeciesAnalyst serves more than 60 million records	desktop GARP for predictive modeling; Species analyst for data search and retrieval; Specify for data capture and management	yes	yes	speciesanalyst.net	worldwide sharing of biodiversity data for application to evolutionary and ecological studies, and studying spread of emerging diseases, invasive species, pests
New York Botanical Garden Herbarium	Vascular plants, bryophytes, lichens, fungi	Records: 700,000	Virtual Herbarium Express, a simple data entry system for herbarium specimen data written in Access and distributed free for use and customization.	yes	yes	www.nybg.org/bsci/hcol/hcol.htm	greater access to NYBG specimen data; improved data management less staff time required to answer individual queries and loan requests
National Museum of Natural History, Smithsonian Institution	most	Records: 4 million			yes		greater scientific research facility
Herbarium, Royal Botanic Gardens, Kew	45% of the flora of Northeast Brazil	Records: 17,000 Species: 2,114 Types: 1,000			yes	www.rbgkew.org.uk/data/repatbr/homepage.html	It increased access to our collections by helping the institution to develop a set of procedures and guidelines to deal with further projects of this kind It also has started to take off the pressure on specific members of staff to attend demands by sending piecemeal data to Brazilian collaborators
The Natural History Museum, London	Many animals and plants	no figures available	mils - system to search NHM databases and allow queries to all. Worldmap - system to plot data geographically and analyze for key	no single receiving country	yes	www.nhm.ac.uk	many, depending on individual project

Institution	Taxa covered	How many	Tools	Tool Sharing	Data available on the Internet	URL	Project Benefits
			conservation regions.				
CABI Bioscience	fungi + protozoan and chromistan fungi	300,000 records representing about 30,000 species			yes	not officially available yet	some real capacity building in the country providing the 'fingers on keyboards' manpower.
Botanische Staatssammlung Munich	diatoms, several taxonomic groups of fungi, compositae	about 22,000 records/specimens					
Botanic Garden and Botanical Museum Berlin Dahlem	plants and fungi	~18,000 specimens, ~6,100 taxa			yes	ww2.bgbm.org/Herbarium/Default2.cfm	reduction of mailing costs and risks for loans
National Botanic Garden of Belgium	8 pilot plant groups	~250 species; ~550 type specimens	Zooming process developed for plates in the frame of "Flora brasiliensis on line" project will be adapted to herbarium specimens	yes	yes	www.be.gbif.net	It is anticipated that this project will make information accessible world-wide for all biodiversity researchers, not only for those conducting basic groundwork studies, but also for those working in the specialized areas of environment and biotope protection. Moreover, it will be of great interest for a wider public interested in historical collections and archives. The proposed project, especially addressing the topic of accessibility, will particularly be able to support biodiversity research carried out in Brazil and bordering countries. Data basing of Martius' collection and digitization of Flora brasiliensis will greatly facilitate ongoing work of taxonomists in the region. The relevance for the country as such lies in the fact that the prototype delivers a structure useful for repatriating biodiversity data kept in foreign collections.
African Biodiversity Information Centre	vertebrates and invertebrates	8,5 million zoological specimens	mainly research collaboration	yes	yes	biodiversitycomores.com; www.fishbase.org	training & education, data-mining and improved data use
South East Asian Botanical	Dipterocarpaceae and several		tools to append data files via internet to the	yes	yes	not yet applicable	Benefits will be reduction in data entry per herbarium

Institution	Taxa covered	How many	Tools	Tool Sharing	Data available on the Internet	URL	Project Benefits
Collections Information Network - National Herbarium Netherlands-Leiden University Branch	genera		central database, tool to change and query these data, in future tools for generating reports, distribution maps, biodiversity analysis tools				because of numerous widespread duplicates, upgrade of trustworthy identifications, more complete data per species, and tools for analysis of data.
Repatriated Biological Collections of India Information Division, National Chemical Laboratory	all taxa of plants, animals, virus, microbes, fungi and fossils		web-based data retrieval tool	yes	yes	www.ncbi.org.in/rbci/	Scientific and academic community in India and neighboring region have access to information on specimens of Indian origin at a click of a mouse
FaunaBase: Western Australian Museum	vertebrates (invertebrates will be included soon)	373,000 vertebrates records	"snapshot" model		yes	www.museum.wa.gov.au/faunabase	on-line access to museum collections information (public) and records (researchers)
National Commission on Biodiversity, CONABIO	plants, animals	2,349,074 specimens 65,374 names	Mallos, gregalis, Biotica, and many ad hoc tools for quality control	yes	yes	www.conabio.gob.mx	Many. Invasive species, vectors of diseases, conservation, regional planning,
Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia	more the 3,000 species of vascular plants	~70,000 records	Distribution maps, gap analysis, taxonomic complementarity among localities	yes	no		Project demonstrates viability of information sharing process among different institutions, using very informal channels of communication and cooperation. Information exchanged has been used in many ways, mainly to research planning and definition of conservation priorities
Centro de Referência em Informação Ambiental - CRIA	plants, animals, microbial strains	184761 specimen records included in the speciesLink network and 46,015 species included in SinBiota	DiGIR software modules; modeling tools; mapping tools; data entry/update; etc.	yes	yes	sinbiota.cria.org.br/atlas splink.cria.org.br/ sicol.cria.org.br/ www.cria.org.br/sp	There are many benefits: - integration of the scientific community; - interaction scientific community and software developers; - national and international cooperation; - dissemination of biodiversity data to a larger public; incentive to digitalization; - possibility of migrating from "reaction" to "prediction"

Reviewing the answers it is possible to start understanding why "data-sharing with country of origin" is becoming tangible. The first important point, in our opinion is that **digitizing and data basing collections are fundamental for the day-to-day operation of a collection**. Within the benefits listed we can see those that are directly important for the information provider, such as:

- Improved reporting;
- Expanded data set for research;
- Larger data set for web display;
- Greater access to specimen data;
- Improved data management and less time to answer individual queries and loan requests; and,
- Reduction of mailing costs and risks for loans.

The second important point is that international collaboration, which always existed in the taxonomic field, is largely enhanced by on-line dissemination of data/information as can be seen by the answers as benefits:

- World wide sharing of biodiversity data for application to evolutionary and ecological studies and studying spread of emerging diseases, invasive species, pests;
- Greater scientific research facility;
- Training and education, data-mining and improved data use;
- Integration of the scientific community; and,
- National and international cooperation.

Another point is inter-disciplinary applications, dissemination to a wider public and the use of scientific data for other purposes (not only taxonomic) as can be seen in the answers:

- Many. Invasive species vectors of diseases, conservation, regional planning, we give about one answer per day…; and,
- The possibility of migrating from "reaction" to "prediction"

2.9. Problems and hurdles addressed

Problems and hurdles were categorized as financial constraints, bureaucratic red tape, technological problems, human resources constraints, training, tools, languages and others (Table 18).

Table 18. Problems and hurdles addressed

Institution	financial constraints	bureaucratic red tape	technological problems	human resources constraints	training	tools	language	Others
Missouri Botanical Garden	1		1	1	1	1		
Natural History Museum & Biodiversity Research Center, University of Kansas		1		1				Non-collaborative attitude among some natural history institutions
New York Botanical Garden Herbarium	1		1					
National Museum of Natural History, Smithsonian Institution	1		1	1				
Herbarium, Royal Botanic Gardens, Kew	1		1	1			1	
The Natural History Museum, London	1	1	1	1	1	1	1	
CABI Bioscience	1	1		1				
Botanische Staatssammlung München	1		1	1				
Botanic Garden and Botanical Museum Berlin Dahlem	1	1		1				
National Botanic Garden of Belgium								
African Biodiversity Information Centre			1	1	1	1	1	
South East Asian Botanical Collections Information Network - National Herbarium Netherlands-Leiden University Branch			1					
Repatriated Biological Collections of India Information Division, National Chemical Laboratory	1			1			1	We have been experiencing delay or no response from some of the collection managers / curators to a call for exchange / sharing initiative or even to develop collaborative programs
FaunaBase: Western Australian Museum	1			1				
Royal Botanic Gardens Sydney, Repatriation of data to Papua New Guinea	1		1		1			
National Commission on Biodiversity, CONABIO	1		1					willingness to share data
Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia			1					
Centro de Referência em Informação Ambiental - CRIA	1		1	1				
Total	13	4	12	12	4	3	4	
% of total	72	22	67	67	22	17	22	

Financial Constraints was the most mentioned (13 institutions – 72%) followed closely by technological problems and human resources constraints (12 institutions – 67%). It is important to note that "bureaucratic red tape" which could be expected to be more significant with such a new technology as is the Internet, is relatively small, having only been mentioned by 4 institutions. Here again as "others" we have "non-collaborative attitude", "delay or no response", and "willingness to share data" being mentioned by initiatives that depend on third parties to get their data.

2.10. Recommendations and Advice

There are a number of recommendations from the institutions surveyed that are worth transcribing. This information in some cases has been summarized and in others is presented "as is". For precise information please refer to the questionnaires in annex.

Table 19. Recommendations and Advice

Institution	Recommendations and Advice
Missouri Botanical Garden	We are visitors, helpers, and supporters In the countries we undertake research, training, and conservation. The Garden, in promoting and participating with the local scientific establishment, has and will seek ways to make the information it holds freely and quickly available for scientific and conservation research.
Natural History Museum & Biodiversity Research Center, University of Kansas	Demonstrate benefits of distributed sharing/applying biodiversity data through demonstration projects contracted out to institutions leading this initiative; form country and region wide partnerships with institutions leading this initiative, and then others, which might now be recalcitrant, will follow.
New York Botanical Garden Herbarium	
National Museum of Natural History, Smithsonian Institution	All countries are users and all are providers. Scientists may focus on taxa in their country, but must have access to data and specimens from other countries to do appropriate taxonomic environmental and resource management work. The Science Committee of GBIF stressed that this is about the free and open access to data, not about repatriation of data. The report from this questionnaire should take that comment very seriously.
Herbarium, Royal Botanic Gardens, Kew	That the data gets checked periodically by an expert in a given plant family to ensure that data repatriated is meaningful for the users. The area of origin of the repatriated specimens should be well delimited in the beginning of the process and that the country of origin should be consulted and give input in the kind of information access is needed. To involve people from the country of origin to develop part of the data basing as they will have fewer language barriers and will receive capacity building during the project.
The Natural History Museum, London	Projects need to be adequately funded
CABI Bioscience	
Botanische Staatssammlung Munich	
Botanic Garden and Botanical Museum Berlin Dahlem	Digital imaging is highly recommended for suitable specimens
National Botanic Garden of Belgium	
African Biodiversity Information Centre	
South East Asian Botanical Collections, Information Network - National Herbarium Netherlands- Leiden University Branch	Full access worldwide should not be granted automatically, especially data on CITES species should be protected. In SEABCIN 3 layers of access will be present: only reading and part of specimen data (collector, year, identification, country) for general use; reading and querying plus downloading for all members of SEABCIN; per institute 2 managers who can also change the data. Outsiders can apply for more access via internet (e.g. collaborators in flora projects). Depending on the status of the applicant access will be free (researchers) or a fee has to be paid (industry).
Repatriated Biological Collections of India, Information Division, National Chemical Laboratory	1. Regional/Global scale proposal may be developed with funding from national governments, international funding agencies to digitize the specimens that especially have origin in developing / under-developed regions of the works. 2. exchange/sharing of data with the countries of origin must be treated as a part of the taxonomic capacity building activity in these regions. 3. specific proposals need to be developed for Asia/ Africa as the collections from these regions are located in Europe/North America. Further these regions would require help (financial, technical, human resource and capacity building) in digitizing the collections.
FaunaBase: Western Australian Museum	Option for a snapshot model as opposed to distributed if speed of online access is the top priority
Royal Botanic Gardens Sydney, Repatriation of data to Papua New Guinea	Not yet – Project initiated 1 October 2003
National Commission on Biodiversity, CONABIO	

Institution	Recommendations and Advice
Alexander von Humboldt Biological Resources Institute	
Reference Center on Environmental Information - CRIA	Repatriation should be seen as a collaborative effort. Emphasis should be given to "data and knowledge sharing".

Comments as to Recommendations and Advice will be made in item 4.2.

2.11. List of Key Publications, Reports and Websites

In the summary (Table 20) only cited websites are included. For lists of publications and reports please refer to the original answers to the questionnaires.

Table 20. Reference Web sites

Institution	Websites
Missouri Botanical Garden	www.mobot.org; www.mobot.org/MOBOT/Research/; www.tropicos.org
Natural History Museum & Biodiversity Research Center, University of Kansas	lifemapper.org
New York Botanical Garden Herbarium	www.nybg.org/bsci/hcol/hcol.html
National Museum of Natural History, Smithsonian Institution	www.nmnh.si.edu/; www.mnh.si.edu/rc/db/databases.html
Herbarium, Royal Botanic Gardens, Kew	www.rbgkew.org.uk/data/repatbr/homepage.html
The Natural History Museum, London	www.nhm.ac.uk/science
CABI Bioscience	URL not officially available
Botanische Staatssammlung Munich	www.botanischestaatssammlung.de/projects/coll_online.html; www.glopp.net
Botanic Garden and Botanical Museum Berlin Dahlem	ww2.bgbm.org/Herbarium/Default2.cfm
National Botanic Garden of Belgium	
African Biodiversity Information Centre	www.africamuseum.be; www.fishbase.org; www.biodiversitycomores.com; www.birdingcomores.com
South East Asian Botanical Collections Information Network - National Herbarium Netherlands-Leiden University Branch	in development
Repatriated Biological Collections of India Information Division, National Chemical Laboratory	www.ncbi.org.in/rbci
FaunaBase: Western Australian Museum	www.museum.wa.gov.au/faunabase
Royal Botanic Gardens Sydney, Repatriation of data to Papua New Guinea	
National Commission on Biodiversity, CONABIO	www.conabio.gob.mx ; www.conabio.gob.mx/institucion/conabio_espanol/doctos/biodiversitas.html
Alexander von Humboldt Biological Resources Research Institute, Bogotá, Colombia	
Reference Center on Environmental Information – CRIA	www.cria.org.br; splink.cria.org.br/; sinbiota.cria.org.br/atlas; www.biotaneotropica.org.br; www.bioline.org.br/

3. General Analysis

3.1. Brief Summary of the Answers

A total of 27 institutions were contacted of which 18 answered the questionnaire. An attempt was made to cover all continents and to include both biological collections and institutions responsible for structuring online information systems. Of the 18 institutions, 4 are from the USA, 8 from Europe, 1 from India, 2 from Australia, and 3 from Latin America. Included are important herbaria and natural history museums that curate specimens from practically the whole world; initiatives concerned with digitizing data for their own institutional purposes; and explicit bilateral agreements directly dealing with data sharing with countries of origin. From a technological view there are institutions concerned in bringing data into the country and maintaining it in a centralized database and others developing distributed systems that will harvest data at remote sites. Although a thorough study on data sharing with countries of origin, including costs and identification of opportunities would demand much more time and resources, we believe that due to the fact that the projects analyzed are heterogeneous, and with the team's relatively long term experience with biodiversity on-line information systems, it was possible to produce an analysis which we believe will help GBIF in determining a strategy for data-sharing. The following points were analyzed by this report:

- data sharing;
- criteria for the selection of taxonomic groups and collections, priorities and user groups
- contractual arrangements and intellectual property rights
- receiving institution / country contribution
- process
- data format and transfer
- exchange formats and standards
- funding, time scale, and costs
- results and products
- problems and hurdles addressed
- recommendations and advice

All answers have been tabulated and are presented in the report and the complete questionnaires are annexed. Below is a brief summary of our analysis.

Data Sharing: most projects, regardless of whether the main purpose is to share data with the country of origin or not, are making information freely available on the internet.

Criteria for the selection of taxonomic groups and collections, priorities and user groups: Herbaria and natural history museums set as their priority for digitizing data and making it available, recent taxonomic revisions, taxonomic groups that have been recently curated and the presence of research groups to work the material. Institutions concerned with developing information systems include the willingness to share data and information. Taxonomic and regional scope are the most important criteria for the selection of a collection. As to priorities addressed, taxonomic and conservation were the most selected criteria. Policy-making was also included which may show the increasing concern to make quality scientific information a basis for the elaboration of policies. As to user groups, as was expected, all included the scientific community as a target user. Policy makers, the "general public" and educators were also frequently mentioned but less then 30% of the institutions included private companies as target users.

Contractual arrangements: An interesting fact is that not one institution included the option "formal contract", making it clear that this is still very informal and most probably carried out as a collaborative effort through informal agreements.

Intellectual Property Rights: IPR issues didn't come across as the 'number 1' issue of concern for most institutions. This may be true for species data in general or maybe it reflects the fact that very few institutions included in this survey have private companies as target users. This may be because most biological collections do not have private companies as target users, or, this is something still new and a growing field, or the survey that was carried out was not representative.

For those institutions not dealing with patents, we believe that important issues are proper attribution or credit, custodianship and ownership (i.e. each contributing museum retains ownership of its records) and acknowledgement. With the evolution of technology it is becoming more possible to let custodians have full control over their data. This way, sensitive or non-validated data may be easily withheld without having to withhold full datasets.

Receiving Institution / Country Contribution: Taxonomy, collaboration among specialists is a necessity, and the proximity of nations due to the advance of communication technology is a fact. We believe, therefore, that the tendency is (aside from political and biosafety reasons) toward the enhancement of collaborative efforts.

Process: As to what type of data is being exchanged virtually all institutions included "label data" and "location information" and most included "Geocoding". It is important to note that these represent the fundamental data necessary for biogeographic studies and predictive distribution modeling. It is also significant that "image" appears in 61% of the institutions. As to data exchange, there is no predominant format but interactive web search is answered by over 70% of the

institutions. With exchange formats and standards, it seems that XML has become a standard. It is also interesting to see DiGIR, which unfortunately was not included as an option within the questionnaire, being mentioned by 4 institutions, and this way becoming the second most used standard. The last aspect analyzed within the process framework was transfer of technology specifically requesting information on software and expertise. Around 40% of the institutions indicated that they transfer software and offer training.

Funding: Project Grants and Fellowships and Institutional Funding are the most important funding mechanisms for data-sharing. It is clear that the importance of data-sharing and the necessity of having funds for digitizing scientific data should be stressed and to include this within the strategy of research funding agencies.

Time Scale and Costs: Time scale showed that data-sharing and digitizing is a recent activity, the oldest initiatives only being several years old. This was expected as data-sharing at this level of integration only became possible with the Internet and the IT developments that followed. As to costs, only few have a clear idea of time and costs per record. Answers vary between 40 seconds and 30 minutes per record and costs vary between US$ 0.40 to US$ 12.70. It appeared that different institution included different parts of the process within their estimates. Both time and cost per record would obviously require a more detailed analysis to verify what each answer entails.

Results and Products. Reviewing the answers it is possible to start understanding why data-sharing with country of origin is becoming a fact. The first important point, -, is that digitizing and data basing collections are fundamental for the day-to-day operation of a collection. Regardless whether linked to a data-sharing program or not collections **must** database their collections and document their activities. The second important point is that international collaboration, which has always existed in taxonomy, is largely enhanced by on-line dissemination of data and information. A third point which is becoming more relevant with the evolution of information and communication technology is the possibility of developing inter-disciplinary applications, disseminating information to a wider public and promoting the use of scientific data for other purposes (not only taxonomic), for example, conservation and policy making.

Problems and hurdles addressed. Problems and hurdles were categorized as financial constraints, bureaucratic red tape, technological problems, human resources constraints, training, tools, languages and others. Financial Constraints, as expected, was the most mentioned, followed closely by technological problems and human resources constraints. It is important to note that "bureaucratic red tape" which could be expected to be more significant with such a new technology as the Internet, is relatively small, having only been mentioned by 4 institutions. Here again as "others" we have "non-collaborative attitude", "delay or no response", and "willingness to share data" being mentioned by initiatives that depend on third parties to get their data. Here again we would like to stress the importance of having digitization and development of tools as strategic for the development of our knowledge base and therefore needs to be considered by funding agencies worldwide.

3.2. Recommendations and Advice

The study that was requested was specifically contracted as a "Study on Data Repatriation", but at an early stage there was a change, not only rewording the term "repatriation" to "data-sharing", but an overall change of concept. "Repatriation" during the first discussions at the Convention on Biological Diversity (CBD) meant (to many) actually transferring scientific holdings to countries where the specimens had been collected. With the evolution of information and communication technology the term "data" was added to the term "repatriation" to mean a specific action concerning transfer or access to data to the country where the specimens had been collected. This continues to be a one-way action, and has, in our opinion, a paternalistic attitude. A much broader concept that was adopted in this work was "data-sharing with countries of origin". The idea behind this is that all have something to contribute. If we wish to change the pattern of destruction and loss we have today one must necessarily share, not only data, but experience, know-how, time, expertise, resources, information and knowledge. This issue was discussed with the GBIF secretariat and the term "data repatriation" was changed to "data sharing with countries of origin".

We recommend that all countries be rightfully considered users and providers of data and information. Not only is it important to promote digitization of existing collections, and it is clear that the largest holdings are held in collections in Europe and the US, but it is fundamental to promote digitization of smaller and perhaps more specialized collections and promote capacity building in countries with rich biological diversity, but also in the fields of curatorial practices (to guarantee quality information) and in information and communication technology.

An important point is that developing countries in the recent past had to consider a large amount of resources to help make larger "world" collections data available. CONABIO's effort in repatriating data in the 90s is a good example. It seems to be logical to consider that developing countries can invest more time and money on working on the organization of their own collections, making this data available and benefiting from the digitization that is already an on-going activity in larger collections. Bilateral agreements continue playing a fundamental role in establishing priorities and enhancing collaboration.

It was worrisome to note that there still are some groups (see South East Asian Botanical Collections Information Network) that apparently believe information should be restricted to its members and a group of collaborators. Users outside this group will have either restricted access or will pay a fee. It is evident that there may be situations were sensitive information exists, but as a rule, we recommend that GBIF promote the concept that information must be made freely and quickly available for educational purposes and for scientific research.

It is important to note the evolution of information systems and we recommend that GBIF explore the development of tools and standards that are open to different architectures and that guarantee that the ultimate control over data remains with the data providers/custodians. There is also a need for standards and tools to aid in the improvement of the quality and accuracy of the data, and the development of these needs supporting. Digitization is a trend that can be catalyzed by adequate policies and technologies to accelerate the process, and to ensure the accuracy and quality of the resultant product.

People must be involved and GBIF must acknowledge, respect and consider different levels of technological development. We recommend that GBIF promote regular meetings with developers of different countries to exchange experience and expertise.

It is clear from the questionnaire that the majority of institutions are making their information available via Interactive Web forms. It is apparent that there are many other institutions, some small – some large that are also making their collection information available in this manner - either individually, or through broader collaborative efforts such as speciesLink, Australian Virtual Herbarium, MaPSTeDI, Species Analyst, etc.

At the moment, we cannot make an assessment as to how many institutions are making their information available on the Internet, and how many collections are involved. It will be important, once the GBIF data discovery system becomes available, that as many as possible of these collections be “discovered” and register themselves.

GBIF could consider the development of a short on-line questionnaire that would be open to any collection to respond to, and that would be advertised through a number of fora such as Taxacom, BioCASE users group, etc. The questionnaire would need to be simple, and possibly automatically collated and analyzed. Most answers could be supplied through a tick-box response. Questions such as

- Name of Collecting Institution
- Web Address for data access
- Number of collections (specimens)
- Available directly by institution, or through secondary organization, or both.
- Groups (plants, animals – possible broken down into 10 subgroups)
- Access conditions (freely available, available with charge, password access, etc.)

- Delivery mechanism (DiGIR, XML, csv, dbf, other)
- Information available (name, location, geocode, collector, image …)
- Contact e-mail

In this way, once the GBIF Discovery system is available, an automated email could be sent to each of the respondents with details on the Discovery system and on registration, etc. It would also quickly give an idea of the number of possible registrants before hand, along with the number of digitized collections likely to be available for sharing/exchange, etc.

Lastly, we recommend that GBIF support demonstrative projects that clearly show the benefits of distributed systems, data sharing and collaborative efforts (applications). We believe that those that are recalcitrant will follow (internet is new – therefore there are cultural barriers to overcome) and through new and renewed partnerships, new ideas and concepts will appear.