



**DORA ANN LANGE CANHOS**

**SISTEMAS DE INFORMAÇÃO EM BIODIVERSIDADE E A  
FORMULAÇÃO DE POLÍTICAS PÚBLICAS NA ERA DIGITAL**

**Campinas – 2013**





**NÚMERO: 282/2013**

**UNIVERSIDADE ESTADUAL DE CAMPINAS  
INSTITUTO DE GEOCIÊNCIAS**

**DORA ANN LANGE CANHOS**

**SISTEMAS DE INFORMAÇÃO EM BIODIVERSIDADE E A FORMULAÇÃO DE  
POLÍTICAS PÚBLICAS NA ERA DIGITAL**

**TESE DE DOUTORADO APRESENTADA AO  
INSTITUTO DE GEOCIÊNCIAS DA UNICAMP  
PARA OBTENÇÃO DO TÍTULO DE DOUTORA  
EM POLÍTICA CIENTÍFICA E TECNOLÓGICA.**

**ORIENTADORA: PROFA. DRA. MARIA BEATRIZ MACHADO BONACELLI**

**ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA TESE  
DEFENDIDA PELA ALUNA DORA ANN LANGE CANHOS E ORIENTADA  
PELA PROF<sup>a</sup> DR<sup>a</sup> MARIA BEATRIZ MACHADO BONACELLI**

---

**CAMPINAS - 2013**

FICHA CATALOGRÁFICA ELABORADA POR  
CÁSSIA RAQUEL DA SILVA – CRB8/5752 – BIBLIOTECA “CONRADO PASCHOALE” DO  
INSTITUTO DE GEOCIÊNCIAS  
UNICAMP

Canhos, Dora Ann Lange, 1954-  
C162s      Sistemas de informação em biodiversidade e a formulação de  
políticas públicas na era digital / Dora Ann Lange Canhos. -  
Campinas, SP. : [s.n.], 2013.

Orientador: Maria Beatriz Machado Bonacelli.  
Tese (doutorado) - Universidade Estadual de Campinas, Instituto  
de Geociências.

1. Comunicação científica. 2. Diversidade biológica. 3.  
Tecnologia da informação e comunicação. I. Bonacelli, Maria Beatriz  
Machado, 1962- II. Universidade Estadual de Campinas, Instituto de  
Geociências. III. Título.

Informações para a Biblioteca Digital

**Título em inglês:** Biodiversity information systems and public policies in the digital age.

**Palavras-chaves em inglês:**

Scientific communication

Biological diversity

Information and communication technology

**Área de concentração:** PC&T – Política Científica e Tecnológica

**Titulação:** Doutora em Política Científica e Tecnológica.

**Banca examinadora:**

Maria Beatriz Machado Bonacelli (Orientador)

Ana Lúcia Delgado Assad

André Tosi Furtado

Ariadne Chloe Mary Furnival

Mercedes Maria da Cunha Bustamante

**Data da defesa:** 26-02-2013

**Programa de Pós-graduação em:** Política Científica e Tecnológica



UNICAMP

UNIVERSIDADE ESTADUAL DE CAMPINAS  
INSTITUTO DE GEOCIÊNCIAS  
PÓS GRADUAÇÃO EM  
POLÍTICA CIENTÍFICA E TECNOLÓGICA

**AUTORA:** Dora Ann Lange Canhos

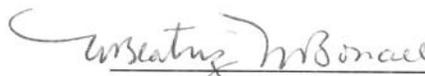
"Sistemas de Informação em Biodiversidade e a Formulação de Políticas Públicas na era Digital".

**ORIENTADORA:** Profa. Dra. Maria Beatriz Machado Bonacelli

Aprovada em: 26 / 02 / 2013

**EXAMINADORES:**

Profa. Dra. Maria Beatriz Machado Bonacelli

  
\_\_\_\_\_  
Presidente

Dra. Ana Lúcia Delgado Assad

  
\_\_\_\_\_  
Examinadora

Prof. Dr. André Tosi Furtado

  
\_\_\_\_\_  
Examinador

Profa. Dra. Mercedes Maria da Cunha Bustamante

  
\_\_\_\_\_  
Examinadora

Profa. Dra. Ariadne Chloe Mary Furnival

  
\_\_\_\_\_  
Examinadora

Campinas, 26 de fevereiro de 2013.



Dedico este trabalho aos meus filhos, Paula, André e Luísa.



## AGRADECIMENTOS

É sempre difícil agradecer às pessoas que direta ou indiretamente influíram na realização de um trabalho.

Primeiro, gostaria de agradecer à Profa. Dra. Maria Beatriz Machado Bonacelli, que teve a difícil missão de orientar uma pessoa com uma boa base prática, mas distante por muitos anos da academia. Certamente não foi uma tarefa fácil. Agradeço também aos membros da banca, pelas críticas e sugestões que seguramente enriqueceram o trabalho.

Agradeço a todos os professores e alunos do DPCT com quem tive a oportunidade de trocar ideias e experiências e aprender muito. E às secretárias Val e Gorete, sempre muito prestativas e eficientes.

A minha experiência profissional foi muito importante para compreender o tema desta tese, e assim devo agradecer a toda equipe do CRIA, passado e presente, Alexandre, Luiz, Thiago, Renato, Marinez, Paula, Érica, Ingrid, Daniel, Cristina, Rosely, Silvia, Fátima, Mariane, Letícia e tantos outros que por lá passaram. Mas devo um agradecimento muito especial ao Sidnei, que, com extrema paciência, ao longo dos últimos 30 anos vem me explicando esse mundo fascinante da informática e transformando ideias em produtos fantásticos.

Ainda no CRIA, devo agradecer aos membros do Conselho Deliberativo, Rubens Naves, Márcio de Miranda Santos, José Fernando Perez, Pedro Paulo Branco, Roberto Waack, Giselda Durigan, Vera Lúcia Imperatriz-Fonseca, fundamentais na manutenção do sonho e dos ideais da instituição. E à minha querida amiga Rosana, também conselheira do CRIA, um agradecimento muito especial, pela sua amizade, sua postura ética e seus valores, tão próximos aos meus.

Devo também agradecer aos curadores, taxonomistas, técnicos e diretores das coleções biológicas participantes da rede *speciesLink*. Este trabalho só foi possível graças ao compartilhamento livre e aberto dos dados de seus acervos e da interação ativa via rede e em reuniões presenciais. Em especial, gostaria de agradecer aos membros do Comitê Gestor do INCT Herbário Virtual da Flora e dos Fungos – Leonor, Ariane, Maria Regina, Mariângela, João Renato e Ana Odete – que, ao longo dos últimos quatro anos, têm ajudado a viabilizar este sonho de estruturar uma infraestrutura pública de dados sobre biodiversidade neste país.

Por fim, agradeço à minha família. Aos meus pais pelo exemplo, valores e oportunidades oferecidas. Aos meus irmãos, sobrinhos, tios, primos, genros, nora e amigos pela diversidade de

experiências, boas e não tão boas, mas que certamente me tornaram uma pessoa melhor. Ao meu marido, Vanderlei, que não só me estimulou, mas muitas vezes me fez trilhar caminhos que eu não imaginava ser capaz. E aos meus filhos, Paula, André e Luísa, que despertaram em mim um eterno aprendiz.

*If I have seen further it is only by standing on the shoulders of giants.*

Isaac Newton (1676)

*What you see depends upon where you stand and how you choose to focus your attention*

(Shapin & Thackray, 1974)





UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE GEOCIÊNCIAS

DEPARTAMENTO DE POLÍTICA CIENTÍFICA E TECNOLÓGICA

**SISTEMAS DE INFORMAÇÃO EM BIODIVERSIDADE E A FORMULAÇÃO DE  
POLÍTICAS PÚBLICAS NA ERA DIGITAL**

**RESUMO**

**Tese de Doutorado**

**Dora Ann Lange Canhos**

A tese apresenta como tema a influência das tecnologias da informação e comunicação (TICs) na circulação do conhecimento científico e o seu efeito na elaboração de políticas públicas em biodiversidade. O avanço das TICs está afetando a forma com que o conhecimento é produzido e como os resultados são difundidos. As TICs tornaram possível o envolvimento de mais atores na gênese do conhecimento, atores de diferentes disciplinas, especialidades, instituições, localidades, países, culturas e realidades sociais. O paradigma tradicional de compartilhamento de dados e resultados científicos por meio apenas de publicações em livros e revistas especializadas não é mais suficiente para atender à demanda contemporânea que necessita não só da disponibilidade dos dados, mas dos processos, análises e métodos e, quando possível, em tempo real, proporcionando uma maior transparência e reprodutibilidade dos resultados. Também não basta a disponibilidade dos dados *on-line*; estes precisam estar organizados de forma padronizada em formatos úteis e utilizáveis, acessíveis, tanto por interfaces humanas, como também via serviços *web*. Além da tecnologia e da necessidade científica de compartilhar dados, métodos e análises em diferentes escalas e disciplinas, existe uma necessidade de acesso e uso dos dados e aplicativos para processos de tomada de decisão, também em diferentes escalas, do local ao global. Isso é particularmente verdadeiro quando o tema é meio ambiente e desenvolvimento sustentável. A tese argumenta que a política para dados sobre biodiversidade deve promover o seu acesso livre e aberto. Destaca a importância de políticas de longo prazo voltadas ao desenvolvimento e manutenção contínua de infraestruturas de dados para armazenar, organizar, preservar, recuperar e disseminar *on-line*, de forma livre e aberta, em formato útil e utilizável, dados e informações sobre biodiversidade, geradas com recursos públicos. Indica a necessidade das agências e do poder público se capacitarem para se apropriarem desses dados e informações disponibilizadas nessas e-infraestruturas. Na tese são analisadas três infraestruturas de dados sobre biodiversidade: a rede global GBIF (Global Biodiversity Information Facility); a rede mexicana Conabio (Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, Mexico) e a rede brasileira *speciesLink*. Para cada rede é apresentada uma breve análise dos pontos fortes e fracos, buscando elementos para avaliar a adequação das políticas de fomento para o desenvolvimento e manutenção desses sistemas. São também apresentados três exemplos de uso da infraestrutura brasileira, a rede *speciesLink*: (i) para o monitoramento e acompanhamento de programas e projetos e para avaliação e valorização dos serviços científicos prestados por curadores e taxonomistas; (ii) para a identificação de lacunas de dados e conhecimento sobre a ocorrência de espécies da flora e fungos do Brasil; e (iii) para a definição de estratégias de apoio à informatização dos acervos de coleções biológicas do país. A tese procura mostrar a oportunidade, viabilidade e importância de usar e-infraestruturas em biodiversidade, não só para ampliar o acesso e a usabilidade dos dados para o desenvolvimento científico, mas para a elaboração e avaliação de políticas públicas, inclusive para melhorar a qualidade, confiabilidade e completude dos dados e informações.

**Palavras Chaves:** comunicação científica, convenção da diversidade biológica, tecnologia da informação e da comunicação





UNIVERSITY OF CAMPINAS  
INSTITUTE OF GEOSCIENCE

**Biodiversity Information Systems and Public Policies in the Digital Age**

**ABSTRACT**

**PhD Thesis**

**Dora Ann Lange Canhos**

The theme of this thesis is the influence of information and communication technologies (ICTs) in the circulation of scientific knowledge and its effect on the formulation of public policies on biodiversity. Advances in ICTs are affecting the way in which knowledge is produced and how results are disseminated. ICTs have enabled the involvement of more actors in the genesis of knowledge, actors from different disciplines, specialties, institutions, localities, countries, and from different social and cultural backgrounds. Thus, the traditional paradigm of sharing scientific data and results through books and journals is not sufficient to attend contemporary demands that require not only the availability of data, but also processes, analysis, and methods and, when possible, in real time, providing increased transparency and reproducibility of results. But just making data available on-line is not sufficient. It must be organized following accepted standards, in useful and usable formats, and accessible, both through human interfaces and as web services. In addition to the technology and the need to share scientific data, methods and analysis in different scales and disciplines, there is the need to use a data infrastructure with tools for decision making, also in different scales, from local to global. This is particularly true when the theme is environment and sustainable development. The thesis argues that policy for biodiversity data must promote its free and open access. It also stresses the importance of long term policies for ongoing maintenance and development of electronic biodiversity data infrastructures, responsible for storing, organizing, preserving, recuperating, and for on-line free and open dissemination of information and data, generated with public funding. The thesis also indicates the need for public agents and authorities to make efficient and effective use of the data and information provided by such e-infrastructures. The thesis analyses three biodiversity data infrastructures: the global network GBIF (Global Biodiversity Information Facility); the Mexican network Conabio (Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, Mexico); and the Brazilian network *speciesLink*. A brief analysis is presented for each network, evaluating strengths and weaknesses, seeking elements to assess the adequacy of policies to promote the development and maintenance of these systems. Three examples of possible uses of the Brazilian network *speciesLink* are also presented: (i) for monitoring programs and projects and for the evaluation and acknowledgement of scientific services provided by curators and taxonomists; (ii) to identify data and knowledge gaps on the occurrence of species of flora and fungi of Brazil; and, (iii) to define strategies to support the computerization of biological collections in the country. The thesis aims at showing the opportunity, feasibility and importance of using e-infrastructures in biodiversity, not only to increase access and usability of data for scientific development, but for the development and evaluation of public policies, including those to improve the quality, reliability, and completeness of data and information.

**Key words:** scientific communication, convention on biological diversity, information and communication technology



# Sumário

<b>ÍNDICE DE FIGURAS.....</b>	<b>XIX</b>
<b>ÍNDICE DE TABELAS .....</b>	<b>XXI</b>
<b>SIGLAS E ABREVIATURAS USADAS.....</b>	<b>XXIII</b>
<b>INTRODUÇÃO .....</b>	<b>1</b>
<b>CAPÍTULO 1. A EVOLUÇÃO DA COMUNICAÇÃO CIENTÍFICA .....</b>	<b>9</b>
1.1. CIVILIZAÇÕES ANTIGAS .....	9
1.2. CULTURA CLÁSSICA (600 A.C. A 500 D.C.).....	10
1.3. IDADE MÉDIA (500 A 1450) .....	12
1.4. REVOLUÇÃO CIENTÍFICA (1450 A 1700) .....	13
1.4.1. <i>Livros, Patentes e Correspondências</i> .....	14
1.4.2. <i>Academias e Sociedades Científicas</i> .....	15
1.4.3. <i>Revistas Científicas</i> .....	16
1.4.4. <i>Obras de Destaque</i> .....	17
1.5. SÉCULO XVIII .....	19
1.6. SÉCULO XIX.....	22
1.7. SÉCULO XX .....	28
1.7.1. <i>O Movimento Ambientalista</i> .....	29
1.7.1. <i>A Internet e Sistemas de Informação em Biodiversidade</i> .....	30
1.8. SÉCULO XXI.....	34
1.8.1. <i>Movimento do Acesso Livre e Aberto a Dados</i> .....	36
1.8.2. <i>Iniciativas Inovadoras no Uso das TICs</i> .....	40
<b>CAPÍTULO 2. COMUNICAÇÃO CIENTÍFICA E MEIO AMBIENTE .....</b>	<b>45</b>
2.1. DADOS, INFORMAÇÃO E CONHECIMENTO .....	46
2.2. CIÊNCIA ABERTA ( <i>OPEN SCIENCE</i> ).....	49
2.3. DADOS E INFORMAÇÕES SOBRE BIODIVERSIDADE.....	52
<b>CAPÍTULO 3. E-INFRAESTRUTURAS SOBRE BIODIVERSIDADE .....</b>	<b>55</b>
3.1. GLOBAL BIODIVERSITY INFORMATION FACILITY – GBIF .....	56
3.1.1. <i>Pontos Fortes do GBIF</i> .....	60
3.1.2. <i>Pontos Fracos do GBIF</i> .....	61
3.2. COMISIÓN NACIONAL PARA EL CONOCIMIENTO Y USO DE LA BIODIVERSIDAD – CONABIO (MÉXICO).....	62
3.2.1. <i>Pontos Fortes da Conabio</i> .....	66

3.2.2. <i>Pontos Fracos da Conabio</i> .....	67
3.3. A REDE <i>SPECIESLINK</i> .....	67
3.3.1. <i>Pontos Fortes da rede speciesLink</i> .....	73
3.3.2. <i>Pontos Fracos da rede speciesLink</i> .....	74
3.4. POLÍTICAS E ESTRATÉGIAS PARA INFRAESTRUTURAS ELETRÔNICAS DE DADOS SOBRE BIODIVERSIDADE NO BRASIL .....	78
<b>CAPÍTULO 4. E-INFRAESTRUTURAS E POLÍTICAS PÚBLICAS .....</b>	<b>81</b>
4.1. MONITORAMENTO E ACOMPANHAMENTO DE PROGRAMAS, PROJETOS E SERVIÇOS CIENTÍFICOS UTILIZANDO AS E-INFRAESTRUTURAS.....	82
4.1.1. <i>Monitoramento da E-Infraestrutura</i> .....	84
4.1.2. <i>Valorização dos Serviços Científicos de Curadores e Taxonomistas</i> .....	89
4.2. USO DA E-INFRAESTRUTURA NA DEFINIÇÃO DE ESTRATÉGIAS PARA PREENCHER LACUNAS DE DADOS E CONHECIMENTO EM BIODIVERSIDADE .....	95
4.3. OS PROVEDORES DE DADOS: DIAGNÓSTICO E DESAFIOS PARA O FORTALECIMENTO DE E-INFRAESTRUTURAS.....	105
4.3.1. <i>Metodologia para Diagnóstico das Coleções Biológicas Provedoras de Dados</i> ....	106
4.3.2. <i>Resultados</i> .....	109
4.3.3. <i>Proposta de Aplicação</i> .....	117
4.4. DISCUSSÃO .....	120
<b>CONCLUSÕES .....</b>	<b>123</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>129</b>
<b>ANEXO 1. PESQUISA DE AVALIAÇÃO E PLANEJAMENTO .....</b>	<b>139</b>
<b>ANEXO 2. E-MAILS ENVIADOS .....</b>	<b>147</b>
<b>ANEXO 3. CARTA DA ORIENTADORA .....</b>	<b>149</b>
<b>ANEXO 4. TABELAS ORIGINAIS COM AS RESPOSTAS DAS COLEÇÕES BIOLÓGICAS .....</b>	<b>151</b>

## ÍNDICE DE FIGURAS

Figura 1.1. Exemplos das xilogravuras da obra <i>Historiae animalium</i> de Conrad Gesner. Fonte: <i>Historiae animalium</i> , 1551 .....	17
Figura 1.2. Os níveis de complexidade da biosfera (NRC, 2009).....	36
Figura 2.1. Pirâmide do Conhecimento (ACKOFF, 1989) .....	46
Figura 3.1. Contribuição de dados dos países participantes à rede GBIF. Fonte: GBIF (2011) ...	59
Figura 3.2. Diagrama da arquitetura da rede <i>speciesLink</i> Fonte: CRIA (2011).....	70
Figura 3.3. Evolução do número de registros <i>on-line</i> da rede <i>speciesLink</i> Fonte: Indicadores <i>speciesLink</i> (outubro, 2012) .....	72
Figura 4.1. Diagrama <i>on-line</i> com a evolução do número de registros e coleções integrantes do INCT Herbário Virtual da Flora e dos Fungos. Fonte: rede <i>speciesLink</i> (2012) .....	84
Figura 4.2. Índice de espécimes por km <sup>2</sup> . Fonte: Rede <i>speciesLink</i> (2012).....	85
Figura 4.3. Aplicativo que mostra o número de coleções e subcoleções por estado brasileiro participante da rede <i>speciesLink</i> . Fonte: Rede <i>speciesLink</i> (2012) .....	86
Figura 4.4. Indicador de informatização dos acervos dos provedores nacionais do INCT Herbário Virtual da Flora e dos Fungos. Fonte: Rede <i>speciesLink</i> (2012).....	89
Figura 4.5. Gráficos <i>on-line</i> sobre o histórico de movimentação de dados e do número de registros por ano de coleta de uma mesma coleção biológica. Fonte: rede <i>speciesLink</i> (2012).....	91
Figura 4.6. Ficha de material não identificado e o acesso à imagem. Fonte: Rede <i>speciesLink</i> (2012) .....	93
Figura 4.7. Ferramenta para comparar imagens de diferentes amostras. Fonte: Rede <i>speciesLink</i> (2012) .....	93
Figura 4.8. Gráficos produzidos dinamicamente indicando (1) os herbários que usaram os serviços do taxonomista e (2) o número de determinações realizadas pelo especialista por ano. Fonte: rede <i>speciesLink</i> (2012) .....	94
Figura 4.9. Visualização dos pontos de ocorrência de espécies de plantas do INCT Herbário Virtual da Flora e dos Fungos no Brasil e no Estado da Bahia. Fonte: Rede <i>speciesLink</i> (2012).....	96
Figura 4.10. Gráficos com o número de espécies de Angiospermas agrupadas nas quatro categorias de quantidade de registros <i>on-line</i> gerados em maio (gráfico A) e em setembro (gráfico B) de 2012. Fonte: Lacunas (2012).....	101

Figura 4.11. Vista parcial do relatório do status dos dados disponíveis <i>on-line</i> para as espécies da família Alstroemeriaceae. Fonte: Lacunas (2012) .....	102
Figura 4.12. Relatório do aplicativo <i>Lacunas</i> para o Anexo I das espécies da flora brasileira ameaçadas de extinção. Fonte: Lacunas (2012) .....	103
Figura 4.13. Trechos do relatório sobre o status dos dados disponíveis para a espécie <i>Attalea funifera</i> Mart. Fonte: Lacunas (setembro de 2012). .....	104
Figura 4.14. Gráfico do percentual dos dados <i>on-line</i> versus a avaliação do critério <i>Informatização</i> por herbário com o tamanho da bola proporcional ao tamanho do acervo. ....	119

## ÍNDICE DE TABELAS

Tabela 3.1. Acesso aos dados da rede <i>speciesLink</i> .....	73
Tabela 3.2. Resumo dos principais pontos levantados sobre as três infraestruturas de dados analisadas.....	76
Tabela 4.1. Contribuição vs dependência de registros por estado. ....	88
Tabela 4.2. Resultado da avaliação dos dados para espécies de Angiospermas .....	99
Tabela 4.3. Tamanho dos acervos das coleções que preencheram o formulário .....	110
Tabela 4.4. Frequência das respostas sobre a disponibilidade e qualificação de recursos humanos. ....	111
Tabela 4.5. Frequência das respostas sobre infraestrutura física .....	112
Tabela 4.6. Frequência das respostas sobre informatização.....	113
Tabela 4.7. Frequência das respostas sobre atividades dos herbários.....	114
Tabela 4.8. Frequência das respostas sobre governança. ....	115
Tabela 4.9. Priorização das ações das coleções .....	116



## SIGLAS E ABREVIATURAS USADAS

ABCD	Access to Biological Collections Data
ANSP	Academic Network at São Paulo
BIN21	Biodiversity Information Network – Agenda 21
BINbr	Biodiversity Information Network – Brazil
BioCASE	Biological Collection Access Service for Europe
Biota-Fapesp	Programa de Pesquisas em Caracterização, Conservação e Uso Sustentável da Biodiversidade do Estado de São Paulo
Biotasp	Embrião do programa Biota-Fapesp, conhecido como Biota São Paulo
CDB	Convenção da Diversidade Biológica
Cenapad	Centro Nacional de Processamento de Alto Desempenho
CGEN	Conselho de Gestão do Patrimônio Genético
CGI	Common Gateway Interface
CHM	Clearing House Mechanism da Convenção da Diversidade Biológica
CITES	Convention on International Trade in Endangered Species of Wild Fauna and Flora
CLF	Institut des Herbiers Universitaires, Clermont-Ferrand
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
CONABIO	Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, Mexico
CRIA	Centro de Referência em Informação Ambiental
C&T	Ciência e Tecnologia
CT&I	Ciência, Tecnologia e Inovação
DiGIR	Distributed Generic Information Retrieval
DwC	Modelo de dados DarwinCore
ECO 92	Conferência das Nações Unidas para o Meio Ambiente e o Desenvolvimento realizado no Rio de Janeiro em 1992
ERIN	Environmental Resources Information Network, Australia
Fapesp	Fundação de Amparo à Pesquisa do Estado de São Paulo
Fiocruz	Fundação Oswaldo Cruz
ftp	File Transfer Protocol
GBIF	Global Biodiversity Information Facility
GEF	Global Environment Facility
GRDI	Global Research Data Infrastructures
GSPC	Global Strategy for Plant Conservation
GTI	Global Taxonomy Initiative
HTTP	Hypertext Transfer Protocol
IBGE	Instituto Brasileiro de Geografia e Estatística
IBAMA	Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis
IDC – RNP	Internet Data Center da Rede Nacional de Ensino e Pesquisa
INBio	Instituto Nacional de Biodiversidad, Costa Rica
INCT-HVFF	Instituto Nacional de Ciência e Tecnologia Herbário Virtual da Flora e dos Fungos
INPE	Instituto Nacional de Pesquisas Espaciais
INPI	Instituto Nacional da Propriedade Industrial
IUCN	International Union for Conservation of Nature

JBRJ	Jardim Botânico do Rio de Janeiro
MCTI	Ministério da Ciência, Tecnologia e Inovação
MinC	Ministério da Cultura
MIT	Massachusetts Institute of Technology
MMA	Ministério do Meio Ambiente
MNHN	Muséum National d'Histoire Naturelle de Paris
MoU	Memorandum of Understanding
NSF	National Science Foundation
OCDE	Organização para a Cooperação e Desenvolvimento Econômico
OECD	<i>Organisation for Economic Co-operation and Development</i>
ONU	Organização das Nações Unidas
OS	Organização Social
OSCIP	Organização da Sociedade Civil de Interesse Público
PPBio	Programa de Pesquisa em Biodiversidade
P&D	Pesquisa e Desenvolvimento
RNP	Rede Nacional de Ensino e Pesquisa
SBSTTA	Subsidiary Body on Scientific, Technical and Technological Advice da Convenção da Diversidade Biológica
SiBBR	Sistema de Informação sobre a Biodiversidade Brasileira
SICol	Sistema de Informação de Coleções de Interesse Biotecnológico
SinBiota	Sistema de Informação Ambiental do Programa Biota-Fapesp
SISBIOTA Brasil	Sistema Nacional de Pesquisa em Biodiversidade
<i>speciesLink</i>	Rede de dados sobre espécies e espécimes
TAPIR	<i>TDWG Access Protocol for Information Retrieval</i>
TCP/IP	<i>Transmission Control Protocol / Internet Protocol</i>
TDWG	<i>Biodiversity Information Standards antes conhecido como Taxonomic Database Working Group</i>
TICs	Tecnologias de Informação e Comunicação
USPTO	<i>US Patent Office</i>
UNEP	United Nations Environment Programme
UNICAMP	Universidade Estadual de Campinas
USP	Universidade de São Paulo
WWW	<i>World Wide Web</i>

## INTRODUÇÃO

A tese tem como tema a influência das tecnologias da informação e comunicação na circulação do conhecimento científico e o seu efeito na elaboração de estratégias e políticas públicas visando a ampliação do conhecimento, a conservação e o uso sustentável da biodiversidade.

O conhecimento da diversidade de espécies no planeta é central para ações de conservação e de desenvolvimento sustentável. O relatório Brundtland (WCED, 1987) introduziu o conceito de desenvolvimento sustentável, definindo-o como aquele que satisfaz as necessidades do presente sem comprometer a capacidade das gerações futuras suprirem suas próprias necessidades. Colocou a sustentabilidade ambiental como uma questão central na tomada de decisão política, junto com o crescimento econômico e a equidade social. Alertou para a necessidade de ações e políticas coordenadas, indicando que o desenvolvimento sustentável ultrapassa as fronteiras dos países e as disciplinas científicas compartmentadas.

Também critica as políticas nacionais, igualmente compartmentadas, com ministérios de mandatos estreitos, preocupados apenas com o desenvolvimento do seu setor. Cita como exemplo os ministérios da indústria que focalizam apenas em metas de produção, sendo que a poluição causada pelas indústrias é da competência dos ministérios do meio ambiente. Assad (2000) chamou a atenção para esse problema no Brasil - *A ausência de uma coordenação central que assuma o papel e a liderança na definição de uma Política Nacional de Biodiversidade, fornecendo as diretrizes, definindo e coordenando as ações, sendo responsável pela articulação entre os diversos atores, é um dos problemas do Brasil.*

O relatório Brundtland estabeleceu como prioritário o problema do desaparecimento de espécies e de ameaça aos ecossistemas, recomendando o estabelecimento de uma convenção internacional sobre espécies com arranjos internacionais de financiamento. Em 1992, foi realizada no Rio de Janeiro, a Conferência das Nações Unidas sobre o Meio Ambiente e Desenvolvimento, ocasião em que foi assinada a Convenção sobre Diversidade Biológica (CDB), ratificada pelo Congresso Brasileiro em 1993. A CDB tem como objetivos a conservação da diversidade biológica, a utilização sustentável de seus componentes e a repartição justa e equitativa dos benefícios derivados da utilização dos recursos genéticos. Define biodiversidade ou diversidade biológica como sendo a variabilidade de organismos vivos de todas as origens, compreendendo, dentre outros, os ecossistemas terrestres, marinhos e demais ecossistemas

aquáticos e os complexos ecológicos de que fazem parte; a biodiversidade compreenderia ainda a diversidade dentro de espécies, entre espécies e de ecossistemas<sup>1</sup>.

A CDB prevê em seu artigo 7º. que cada país deve identificar e monitorar componentes da diversidade biológica importantes para sua conservação e utilização sustentável e manter e organizar os dados derivados dessas atividades. Destaca-se aqui a importância do trabalho dos taxonomistas<sup>2</sup> na identificação e classificação de espécies e da estruturação de sistemas de informação.

Em 1998, a Conferência das Partes da CDB estabeleceu a Iniciativa Mundial sobre Taxonomia (*Global Taxonomy Initiative - GTI*)<sup>3</sup> recomendando uma série de objetivos e atividades acordadas pelos governos, como facilitar o acesso ao material para uso em sistemática e disponibilizar todas as informações associadas aos espécimes. Foi o Brasil quem apresentou o tema “taxonomia” no âmbito da CDB por entender que havia um “impedimento taxonômico” ou lacunas de conhecimento que prejudicavam a nossa habilidade em conservar, usar e compartilhar os benefícios advindos da nossa biodiversidade. Portanto, foi a proposta brasileira que resultou no estabelecimento da Iniciativa Mundial sobre Taxonomia.

A preocupação com o impedimento taxonômico é pertinente uma vez que o Brasil é considerado um país megadiverso<sup>4</sup>, abrigando cerca de 15% a 20% de toda a diversidade biológica mundial, o que lhe confere uma extraordinária competitividade diante de demandas ambientais e biotecnológicas. Tal biodiversidade, se bem manejada, pode gerar grandes benefícios econômicos e sociais para o país. No entanto, nossa base de dados e de informações é incipiente e desagregada, com enormes lacunas de conhecimento, tanto geográficas como taxonômicas. Essa situação torna-se mais crítica diante da rapidez com que estão sendo destruídos os ecossistemas e do longo tempo que se leva para formar um especialista em taxonomia. Portanto, toda iniciativa no sentido de formar competência e de gerar informação

---

1 Esta tese, quando se refere a dados e informações sobre biodiversidade, refere-se à diversidade de espécies.

2 Taxonomia é a disciplina da biologia dedicada à descrição, atribuição de um nome e catalogação dos organismos e suas relações.

3 Para mais informações sobre o *Global Taxonomy Initiative*, acessar <http://www.cbd.int/gti>

(Nota: Ao longo da tese são apresentados *links* para informações disponíveis na Internet, não só para servir de referência àqueles que quiserem acessar a informação citada, mas para demonstrar a crescente acessibilidade à informação *on-line*.)

4 O conceito de megadiversidade foi criado para priorizar os esforços de conservação pelo mundo. País de megadiversidade é o termo usado para designar os países mais ricos em biodiversidade do mundo. A Conservation International em 1998 identificou 17 países megadiversos, entre os quais está o Brasil.

sobre esse tema, tornando-a disponível para o maior número possível de pessoas, apresenta-se de grande relevância.

Estimativas indicam que o número de espécies descritas no mundo é da ordem de 1,9 milhão, enquanto o número estimado de espécies só de plantas e animais é maior que sete milhões (CHAPMAN, 2009). Significa que, após cerca de três séculos de história de coletas e estudos sobre a biodiversidade do planeta, se as estimativas do nível de conhecimento em relação ao número de espécies existentes se confirmarem, nosso conhecimento atual é de apenas cerca de 25% das espécies de animais e plantas, sem levar em conta os microrganismos.

Uma importante fonte de dados são as coleções biológicas, criadas com o objetivo de armazenar amostras representativas da fauna, flora e microbiota para apoiar o ensino e a realização de estudos taxonômicos. Também existem diferentes estimativas sobre o número de amostras de espécies depositadas nas coleções biológicas do mundo. Ariño (2010) propôs um modelo probabilístico chegando a valores de cerca de 1,2 a 2,1 bilhões de amostras e algumas estimativas indicam que os acervos brasileiros contribuem com cerca de 1% do total (EGLER e SANTOS, 2006), um valor pouco representativo, principalmente quando se leva em conta a dimensão da nossa biodiversidade.

Esses números indicam a necessidade de realizar mais pesquisas voltadas ao conhecimento da biodiversidade do planeta, muito provavelmente usando outras técnicas de observação, coleta, identificação de material e registro dos dados, além de esforços internacionais voltados à construção de uma infraestrutura compartilhada de dados, elementos que colocam desafios importantes para o país.

Com a introdução da computação e de redes de comunicação, o trabalho do taxonomista passa a ter outra dimensão, sendo possível relacionar esse conhecimento a outras disciplinas e áreas de interesse. O nome da espécie permite a indexação do conhecimento e a sumarização da informação disponível sobre cada organismo nomeado (PIRANI, 2005). Permite a integração dos dados históricos e recentes e o seu uso em inúmeras aplicações novas, como estudos biogeográficos, definição de áreas prioritárias para conservação, projetos de zoneamento ecológico-econômico, estudos de mudanças climáticas, recomposição de áreas, entre inúmeras outras (CHAPMAN, 2005). Assim, dada a importância da disponibilização *on-line* de dados taxonômicos de qualidade, uma das preocupações deste trabalho de tese é a valorização das

atividades de taxonomia e de curadoria de coleções biológicas<sup>5</sup> como um importante serviço para a pesquisa científica, reconhecido pela academia e pelo sistema de CT&I do país, passando a integrar de forma mais sistemática critérios de avaliação em diversos âmbitos.

Por sua vez, LUCAS (2003), no contexto de publicações científicas, indica a importância do desenvolvimento de bases de dados pertinentes à realidade local sendo potencialmente espelho da atividade de pesquisa do país ou da região. Lastres e Albagli (1999), discutindo a maior difusão de informação e conhecimento graças às novas tecnologias de informação e comunicação, indicam que *somente aqueles que estão envolvidos na criação do conhecimento dispõem de possibilidades reais de absorver e fazer uso deste maior acesso.*

A estruturação de sistemas próprios de informação, fruto do desenvolvimento de pesquisas locais, significa uma maior probabilidade de apropriação desse conhecimento para a geração de políticas públicas. A estruturação de sistemas de informação é, portanto, um componente estratégico na organização e disseminação de dados para o avanço científico, para a tomada de decisão informada (*informed decision making*) e, desse modo, para a elaboração de políticas públicas voltadas à biodiversidade e meio ambiente.

Esta tese parte do pressuposto que estamos vivenciando um momento de inflexão (*tipping point*) da comunicação científica, no qual a disseminação de dados *on-line* ganha importância com consequências no desenvolvimento de estratégias e políticas públicas para ciência, tecnologia e inovação. Também pressupõe que a evolução para um cenário de maior equidade e justiça social e de sustentabilidade ambiental passa necessariamente por um processo de ampla circulação de dados, de informações e do conhecimento científico. Portanto, a tese defende que a política para dados sobre biodiversidade deve promover o seu acesso livre e aberto.

O acesso livre e aberto a dados sobre biodiversidade pode trazer grandes benefícios para a ciência e tecnologia, viabilizando o seu uso e reuso mais amplo, por diferentes públicos, de diferentes áreas de conhecimento, tornando os resultados mais claros, transparentes e reproduzíveis. Pode inclusive realimentar as políticas e estratégias públicas nos diferentes níveis, desde o local ou institucional até o nacional ou mesmo o global, uma vez que o desenvolvimento sustentável depende cada vez mais de ações cooperativas e interativas em todos os níveis.

---

<sup>5</sup> Coleções biológicas são centros de documentação da biodiversidade composto pelo acervo (amostras de espécimes), dados associados (dados taxonômicos - nome da espécie; geográficos - local da coleta; data da coleta, imagens etc.) e bibliografia associada. O curador é responsável pela organização, preservação e documentação do acervo.

É evidente que setores econômicos como a indústria farmacêutica e de cosméticos, que usam a biodiversidade como fonte de elaboração de seus produtos, trabalham com dados sigilosos de acesso restrito. Dados sobre biodiversidade considerados sensíveis<sup>6</sup> também merecem tratamentos diferenciados. Ou seja, uma política de acesso livre e aberto a dados sobre biodiversidade deve considerar algumas situações como exceção e promover estratégias diferenciadas.

O objetivo da tese é compreender e demonstrar a importância estratégica das infraestruturas eletrônicas de dados de acesso livre e aberto sobre biodiversidade. Procura analisar como vem ocorrendo a disseminação de dados sobre biodiversidade como elemento da comunicação científica e propõe o uso de infraestruturas eletrônicas de dados, por agentes públicos, para valorizar o trabalho dos profissionais envolvidos e subsidiar a definição de estratégias (públicas ou privadas) para a pesquisa em biodiversidade.

São muitos os termos usados para dar nome às iniciativas que dão suporte a pesquisas colaborativas através de redes eletrônicas como *e-Science*, *cyberinfrastructure*, *e-infrastructure*, entre muitos outros. Trata-se de uma área em franco desenvolvimento, cujos termos ainda são usados de maneira bastante livre. O termo *e-Research* engloba todos os termos anteriores e pode ser definido como o uso de ferramentas avançadas de computação e de redes de alto desempenho para colaboração e compartilhamento de recursos em pesquisas acadêmicas e científicas (SCHROEDER, 2007).

As infraestruturas eletrônicas de informação e comunicação estão se tornando sistemas semelhantes às grandes infraestruturas de pesquisa. Para dar uma ideia mais fiel dos desenvolvimentos mais recentes dos sistemas de informação sobre biodiversidade, ao longo da tese será utilizado o termo infraestrutura eletrônica de dados ou simplesmente e-infraestrutura, que inclui o sistema de dados e informações propriamente dito, as ferramentas associadas e a rede de provedores de dados e usuários.

A tese está estruturada em quatro capítulos. O Capítulo 1, **A Evolução da Comunicação Científica**, apresenta uma breve análise desse percurso ao longo da História, procurando destacar as formas de organização da comunidade, os meios de comunicação existentes e a evolução tecnológica, sempre que possível ilustrando com alguma produção científica de destaque para o tema biodiversidade. Esse capítulo procura mostrar a evolução na organização da comunidade

---

<sup>6</sup> A definição de dados sensíveis é apresentada no Capítulo 2

científica e de sua comunicação – oral, escrita, impressa, *on-line* – na qual cada elemento novo, decorrente dos avanços científicos e tecnológicos, vem se somar àqueles existentes. Procura mostrar as inovações na comunicação científica que apresentaram elementos de ruptura ou grandes mudanças para compreender o impacto das tecnologias de informação e comunicação (TICs) em período mais recente.

O Capítulo 2, **Comunicação Científica e Meio Ambiente**, tem como objetivo apresentar a fundamentação teórica sobre a dinâmica da ciência e tecnologia na atualidade. São apresentados os conceitos de dados, informação, conhecimento e ciência aberta, assim como a classificação descrita no Manual Frascati para as atividades de ciência e tecnologia (STA - *Scientific and Technological Activities*), que incluem os serviços científicos e tecnológicos (STS – *Scientific and Technological Services*) que englobam atividades de bibliotecas e museus, por exemplo, dentre várias outras (OECD, 2002). Por analogia as atividades das coleções biológicas e as infraestruturas eletrônicas de dados e informações sobre biodiversidade de acesso público são serviços científicos e como tal precisam ser reconhecidos. A compreensão desse conceito é importante para valorizar o papel estratégico das coleções, das e-infraestruturas e dos profissionais que nelas atuam. É também apresentado o desafio do desenvolvimento sustentável, que aliado aos avanços das TICs e da ciência aberta, justificam a criação das infraestruturas de dados *on-line* sobre biodiversidade (e-infraestruturas).

O Capítulo 3, **E-Infraestruturas sobre Biodiversidade**, apresenta uma análise de três e-infraestruturas sobre biodiversidade: GBIF, a rede global; CONABIO do México; e, *speciesLink*, a rede brasileira, apontando e discutindo suas principais características, pontos fortes e fracos, avaliando, enfim, elementos que possam sustentar uma política de fomento à e-infraestrutura sobre biodiversidade no país.

O Capítulo 4, **E-infraestruturas e políticas públicas**, analisa, a partir da e-infraestrutura brasileira, a rede *speciesLink*, a possibilidade de definir estratégias públicas para pesquisa em biodiversidade, objetivo da tese, a partir da apresentação de três (potenciais) usos da referida rede<sup>7</sup>. Para tal, gerou-se aplicativos, levantou-se informações e dados disponíveis *on-line* e mesmo dados primários, via aplicação de questionário junto a coleções biológicas da rede.

---

<sup>7</sup> Vale comentar que a autora desta tese coordenou o trabalho de análise dos dados e definição dos produtos desejados para o desenvolvimento dos aplicativos apresentados no Capítulo 4.

A primeira análise recai sobre o uso dos dados da rede para monitoramento e acompanhamento de programas e projetos e para serviços científicos. A segunda análise associa os dados da e-infraestrutura do Herbário Virtual da Flora e dos Fungos, um dos Institutos Nacionais de Ciência e Tecnologia do MCTI, com a Lista de Espécies da Flora do Brasil (2012) e a Lista Nacional das Espécies da Flora Brasileira Ameaçadas de Extinção, apresentando um diagnóstico *on-line* das possíveis lacunas de conhecimento da flora e dos fungos do Brasil. Propõe o uso de aplicativos em bases dinâmicas de dados para subsidiar a análise de especialistas na orientação de novas pesquisas e na caracterização e monitoramento do status de conservação de espécies. A última análise, que também usa a e-infraestrutura de dados sobre biodiversidade, baseou-se em levantamento de dados primários e secundários (*on-line*) junto às coleções biológicas provedoras de dados, visando obter um diagnóstico da situação e definição de ações prioritárias para aumentar a quantidade e qualidade dos dados compartilhados *on-line*.

Para fechar as discussões, as **Conclusões** destacam os principais pontos do trabalho, procurando demonstrar que as TICs realmente estão revolucionando a forma de fazer ciência e as infraestruturas eletrônicas de dados estão no centro desta mudança. Destaca a importância dos dados de coleções biológicas e o papel do taxonomista e do curador quanto à qualidade dos dados e a sua disponibilização na e-infraestrutura. Procura enfatizar a necessidade e a oportunidade de explorar as e-infraestruturas nos processos de tomada de decisão nas diferentes escalas, do local até global, e na elaboração de políticas públicas em biodiversidade. Além de uma política de acesso livre e aberto a dados sobre biodiversidade, é também destacada a importância do desenvolvimento das e-infraestruturas em ambientes abertos, evitando, sempre que possível, a utilização de software, ferramentas e padrões proprietários. Por fim, destaca a importância do desenvolvimento de e-infraestruturas de diferentes áreas de conhecimento de forma a permitir associar diferentes tipos de dados e informações necessários para análises multi e interdisciplinares.



## **CAPÍTULO 1. A EVOLUÇÃO DA COMUNICAÇÃO CIENTÍFICA**

Novas estruturas de informação e comunicação são moldadas pelas formas sociais e viabilizam padrões específicos de interação (LATHAM e SASSEN, 2005). Ciência e tecnologia são atividades sociais resultantes da geração, do uso e da aplicação de dados, informação e conhecimento, de ferramentas, métodos e técnicas<sup>8</sup>. A comunicação técnico-científica constitui parte essencial do processo de pesquisa e desenvolvimento, sendo que a maneira pela qual o cientista transmite informações depende do veículo utilizado, da natureza das informações e do público-alvo (MEADOWS, 1999).

Para compreender melhor o presente e propor ações para o futuro é importante conhecer o passado, porém este capítulo não se propõe a ser uma descrição detalhada de todos os eventos históricos relevantes para o tema da comunicação científica em biodiversidade. São apresentadas, em linhas gerais, a evolução da comunicação científica, as formas de organização da comunidade e os meios existentes, sempre que possível ilustrando com alguma produção científica de destaque para o tema biodiversidade.

Vickery (2000) apresenta uma retrospectiva histórica da comunicação científica dividida em sete períodos: as civilizações antigas, a cultura clássica, o período medieval, a revolução científica, os séculos XVIII, XIX e XX. É esta a divisão que será adotada nesta tese ao discutir a evolução da comunicação científica, acrescentando, quando possível, exemplos da história natural, contextualizando as informações relativas ao Brasil na área de ciências naturais a partir do século XVII. O objetivo aqui é mostrar as inovações na comunicação científica que apresentaram elementos de ruptura ou de grandes mudanças em relação ao padrão existente e, dessa forma, compreender o impacto das tecnologias de informação e comunicação (TICs) nos dias de hoje.

### **1.1. CIVILIZAÇÕES ANTIGAS**

As civilizações antigas tinham um grande conhecimento prático acumulado sobre o seu ambiente físico e biológico e já haviam desenvolvido muitas técnicas para o seu manuseio, inclusive técnicas agrícolas, a criação de animais, a pesca, a fiação e a tecelagem, a cerâmica e a carpintaria, a extração de metais e as construções em pedra e tijolo. Como consequência da

---

<sup>8</sup> Uma discussão sobre a definição e os conceitos de dados, informação e conhecimento é apresentada no Capítulo 2

evolução do conhecimento técnico e da crescente produtividade do homem, surgiram as civilizações urbanas por volta do ano 3.000 a.C.

Com a emergência das cidades veio a necessidade dos registros permanentes, desenvolvendo-se a escrita como meio de preservar e transmitir informações, ideias e sentimentos. Na Mesopotâmia, onde havia o uso de placas de argila para o ensino da escrita, matemática e música nos templos e palácios, foi possível encontrar arquivos e bibliotecas com coleções de informações sobre observações astronômicas, prescrições médicas, receitas químicas, além de leis, cartas oficiais e registros administrativos.

No Egito, apesar de ter a escrita se iniciado depois da Mesopotâmia, sendo talvez por ela influenciada, sobreviveram poucos registros porque se usava o papiro, muito menos resistente à ação do tempo. Mas existem vários registros nas paredes internas das pirâmides que, além de relatos da vida dos faraós, incluem textos administrativos e problemas de aritmética e geometria.

A civilização persa contribuiu muito para o avanço da comunicação, construindo a estrada real onde desenvolveu um sistema de mensageiros, com trocas de cavalo, o que aumentou a distância que poderia ser percorrida em um dia de 20 para 100 milhas.

Quanto aos fenícios, eles viviam na costa e usavam barcos a vela como meio de transporte, utilizando a para suas transações comerciais, motivo pelo qual foram considerados responsáveis pelo desenvolvimento da escrita alfabética.

É difícil caracterizar algo como atividade científica antes da Grécia, mas é importante registrar a mudança na comunicação, antes somente oral, de pessoa a pessoa, e depois, com a escrita, alcançando mais pessoas, ampliando também seu escopo geográfico e temporal.

## 1.2. CULTURA CLÁSSICA (600 A.C. A 500 D.C.)

A cultura clássica teve como destaque o desenvolvimento de cidades-estados (fenícios e gregos) e novos impérios (persa e romano). Não se pode afirmar quando teve início a pesquisa científica e, por conseguinte, a comunicação científica, mas os debates filosóficos dos gregos antigos na Academia (periferia de Atenas) nos séculos V e IV a.C. tiveram impacto na pesquisa científica moderna. As obras escritas, como as de Aristóteles, e os debates influenciaram primeiro a cultura árabe e depois a Europa Ocidental (MEADOWS, 1999).

O alfabeto grego, derivado do fenício, e sua simplificação, espalharam a alfabetização para um segmento muito mais amplo da sociedade. Antes restrito aos sacerdotes, governantes e escribas, a democratização do aprendizado com escritas mais simples teve consequências

culturais importantes, como a fundação de escolas onde se ensinava a leitura, escrita, aritmética e ginástica. A educação superior dependia de encontrar um professor, e, assim, com o crescimento da demanda, surgiram os sofistas ou professores profissionais.

Os sofistas eram professores itinerantes, mas Platão estabeleceu uma escola nos arredores de Atenas em um lugar que pertencia a Akademos. Platão em sua escola (Academia) ensinava filosofia, ética e política. A comunicação era promovida através de debates filosóficos e essas especulações filosóficas formaram a matriz do pensamento científico. Aristóteles foi um membro da Academia, tendo sido tutor do Alexandre, o Grande. Após a morte de Platão, ele montou um novo estabelecimento denominado Liceu onde seus pupilos estudavam filosofia, lógica e ciências e realizavam pesquisas. Alexandre o Grande, em suas viagens coletava animais, plantas, minerais e curiosidades e os enviava para Aristóteles como material de estudo. Suas campanhas, portanto, podem ser consideradas as primeiras expedições científicas da história (VICKERY, 2000).

Alexandre fundou Alexandria (no Egito, sendo que depois surgiram algumas dezenas de cidades homônimas), que viria a se tornar um dos maiores centros culturais, científicos e econômicos da Antiguidade. Em Alexandria foram estabelecidos uma biblioteca e um museu, onde trabalhavam mais de 100 estudiosos. Diferente de Atenas, onde a comunicação científica era predominantemente oral, a tradição em Alexandria era o ensino por meio dos livros, com uma grande produção de textos no próprio museu. Essas atividades promoveram a organização do mercado de livros, havendo na cidade uma fábrica de papiros.

Estima-se que Roma no ano 300 d.C. tinha 50 mil milhas de estradas pavimentadas, para o trânsito de pedestres, mensageiros, animais de cargas e carroças, o que aumentava a distância máxima de locomoção para 150 milhas/dia. Havia um grande comércio, inclusive de livros, que incluíam poemas, peças teatrais, história, filosofia, biografias, técnicas agrícolas, manuais, arquitetura e medicina. Muitos escravos gregos e homens livres tinham como ocupação copiar livros. Depois do século III o papiro foi sendo substituído pelo pergaminho.

Os romanos contribuíram na codificação e organização da comunicação do conhecimento. Foram produzidos muitos manuais, tratados e enciclopédias. Um marco de comunicação científica em história natural foi a publicação da obra *Naturalis Historia* (77-79 d.C.)<sup>9</sup> um manuscrito com 102 volumes de Gaius Plinius Secundus (*Pliny the Elder*) com os conhecimentos

---

<sup>9</sup> Ver produções sobre a obra de Pliny, inclusive traduções, no Biodiversity Heritage Library <http://www.biodiversitylibrary.org> acessado em 15/06/2011

que os romanos tinham sobre cosmologia, astronomia, geografia, zoologia, botânica, mineralogia, medicina, metalurgia e agricultura. O trabalho tornou-se modelo para futuras enciclopédias pela amplitude do assunto, por referenciar os autores originais e por apresentar um índice do conteúdo. É um trabalho que, abrangendo o mundo orgânico de plantas e animais e a matéria inorgânica, procura não só descrever a ocorrência das plantas e animais, mas também a sua exploração pelo homem, especialmente os romanos. No prefácio o autor afirma ter mencionado 20 mil fatos obtidos de 2 mil livros de 100 autores selecionados. A lista completa de autoridades traz mais de 400, incluindo 146 fontes romanas e 327 gregas.

### 1.3. IDADE MÉDIA (500 A 1450)

A queda do Império Romano caracteriza o fim da era clássica e o início da Idade Média na Europa. Foi um período de grandes transformações políticas e culturais, inclusive com a adoção do Cristianismo por Constantino (306-337) que também foi responsável pela construção da nova capital no oriente, Constantinopla. O império estava dividido em sua parte ocidental, centrada em Roma e que falava Latim e sua parte oriental centrada em Constantinopla onde se falava Grego. Durante o século V, os seus territórios na Europa ocidental, inclusive a Itália, e norte da África sofreram várias invasões e saques enquanto a parte oriental manteve-se praticamente intacta por vários séculos.

A história da comunicação científica nesse período é complexa porque a cultura clássica mediterrânea deu início a três linhas divergentes: o império bizantino, os árabes e os cristãos.

O império Bizantino, ou Império Romano do Oriente, centrado em Constantinopla sobreviveu ao saque de Roma, preservando a cultura grega, mas em um nível de vitalidade baixa.

Os árabes foram responsáveis pela preservação do conhecimento grego e por sua ampliação, principalmente nas áreas de álgebra e medicina. Entre os séculos VII e X, estudiosos de Gondeshapur e Bagdá traduziram grande parte da ciência grega disponibilizando-a para o mundo árabe. Em 826 foi fundada em Bagdá a “Casa do Conhecimento” para preparar as traduções, mas ao final do século X, os árabes começaram a produzir material original. Por todo o seu império construíram centros de aprendizagem e bibliotecas, além de observatórios. No século XIII, a ciência árabe entrou em declínio e o centro de avanço científico se deslocou para a Europa.

O império romano ocidental foi invadido pelos bárbaros, povos agrícolas que haviam abandonado a vida nômade. Apesar de mais primitivos, foram responsáveis por avanços

importantes em técnicas agrícolas. Na Europa, o feudalismo formou a base social para uma estrutura governamental, o crescimento de cidades e o desenvolvimento de novos ofícios.

Os mosteiros cristãos a partir do século VI, preservaram a tradição do aprendizado por meio de livros. Eram ilhas de aprendizagem e centros de transmissão da cultura clássica. Possuíam *scriptoriums*, locais ou espaços onde trabalhavam os monges copistas. No período entre 1050 a 1350 foram construídas as grandes catedrais, que, com suas bibliotecas, também se tornaram uma espécie de universidade para ensino religioso.

Para se fortalecerem diante dos senhores feudais, comerciantes e artesãos se organizaram em corporações (*guilds*). Essa forma de estruturação será a base para a organização dos pesquisadores e formação das sociedades científicas no século XVII.

Uma outra característica importante do período diz respeito às línguas vernáculas do Islã e da Europa que, aos poucos, também se tornam veículos literários.

A evolução técnica na comunicação incluía o códice de pergaminho, uma evolução importante do rolo, e o uso de papel e xilogravura. Tem-se ainda que, do ano de 1436 a 1439, Gutenberg inventa a impressão por tipos móveis, possibilitando a cópia muito mais rápida de livros e jornais. Ele inventou um processo de produção em massa do tipo móvel utilizando tinta a base de óleo com uma prensa de madeira. A combinação desses elementos criou um sistema prático, substituindo os manuscritos e a impressão em blocos de madeira. A invenção de Gutenberg iniciou a revolução da imprensa e teve papel fundamental no desenvolvimento da Renascença, Reforma e na Revolução Científica (SHAPIN, 1996).

#### 1.4. REVOLUÇÃO CIENTÍFICA (1450 A 1700)

Os períodos anteriores se caracterizaram pela preocupação em registrar e catalogar os conhecimentos existentes e em reproduzir e traduzir as obras clássicas. O período da Renascença rompe com a tradição escolástica medieval judaico-cristã, e tudo passa a ser pensado sob a ótica da condição humana, aliando a experiência e a razão. Merece destaque o início da experimentação científica, a base da ciência moderna, por Francis Bacon (1561-1626). Bacon *deu um novo ordenamento às ciências, propôs a distinção entre a fé e a razão para não cair nos preconceitos religiosos que distorcem a compreensão da realidade* (GADOTTI, 2005). O conhecimento científico, para Bacon, tem por finalidade servir o homem e dar-lhe poder sobre a natureza, sendo que a descoberta de fatos verdadeiros não depende do raciocínio silogístico aristotélico, mas sim da observação e da experimentação regulada pelo raciocínio indutivo

(BATISTA, 2010). Portanto, vem do século XVII a noção de que o monopólio da criação de conhecimentos novos não pertencia ao mundo antigo e que novas observações e ideias poderiam contribuir para elevar o nível do conhecimento existente, ou seja, que o conhecimento era cumulativo (MEADOWS, 1999).

O cientista renascentista buscava novos dados, novas experiências, novas teorias e novos conhecimentos e sentia a necessidade de compartilhar suas últimas descobertas, beneficiando-se das críticas e sugestões dos outros (VICKERY, 2000). Foi o período de filósofos notáveis como Copérnico, Kepler, Newton, Galileu e Descartes.

#### *1.4.1. Livros, Patentes e Correspondências*

A necessidade de compartilhar experiências aliada à tecnologia de impressão incentivou o aumento na produção de livros técnicos por engenheiros, farmacêuticos, navegadores etc., que dessa forma também passaram a divulgar suas práticas e técnicas. Aumentou o uso da língua vernacular e tornou a educação acessível às pessoas fora do mundo acadêmico. Houve importantes avanços técnicos em matemática, na compreensão do sistema solar, na física, anatomia, microscopia e história natural. A instrumentação científica também avançou muito. A navegação e o desenvolvimento de mapas, a mineração, a engenharia e a química também avançaram, inclusive com o início da fabricação de produtos químicos em escala industrial. O capitalismo começa a substituir o feudalismo, nações-estado se desenvolvem e os europeus começam a explorar o mundo, estabelecendo impérios coloniais.

Nesse período foram estabelecidas as patentes, um marco importante que afeta a comunicação científica até os dias de hoje. Existem registros de proteção semelhante a patentes na Grécia Antiga, na Inglaterra no século XIV com as *cartas patente*, e em Florença em 1421. Mas Veneza é considerada um marco, uma vez que, a partir de 1450, as patentes passaram a ser concedidas sistematicamente. Como os venezianos emigraram para outros países, eles passaram a exigir, nesses locais, proteção de suas invenções, contribuindo assim para a difusão do sistema patentário.

Um meio fundamental de comunicação científica nesse período foram as correspondências: muitos mercadores foram usados como mensageiros, contribuindo para criar verdadeiras redes de comunicação entre os cientistas. No século XVII tiveram início os serviços postais regulares.

#### 1.4.2. Academias e Sociedades Científicas

Para aumentar a interação e troca de informações é que surgiram as academias e sociedades científicas. As primeiras academias foram organizadas na Itália por humanistas que buscavam um espaço fora das universidades, que ainda defendiam as tradições aristotélicas. Nessas academias era ensinada a língua grega e estudados manuscritos gregos recém-descobertos.

A primeira academia de filósofos naturalistas de que se tem notícia é a *Accademia Secretorum Natureae* fundada em Nápoles em 1560, por Giovanni Baptista Della Porta, e tinha como meta estudar os *segredos da natureza*. Mas foi acusada de bruxaria e se dissolveu. A *Accademia dei Lincei*, fundada em 1603, tinha também por objetivo compreender as ciências naturais. Della Porta entrou nessa academia em 1610 e Galileu em 1611<sup>10</sup>. A academia queria estabelecer “mosteiros não monásticos” em todo mundo, ou seja, um local com museu, biblioteca, oficina tipográfica, laboratório e jardim botânico. Não conseguiu, mas se tornou um importante centro de pesquisa experimental. As experiências foram registradas nos anais da academia *Gesta Lynceorum*, manuscrito que pode ser considerado a primeira publicação de uma sociedade científica (VICKERY, 2000).

Na Inglaterra, a educação nas universidades era restrita a assuntos de interesse para a igreja. A maioria dos cientistas trabalhava fora de instituições formais. Em 1598, graças à doação do Sir Thomas Gresham, nasce a *Gresham College*, primeira casa da *Royal Society of London for the Improvement of Natural Knowledge*, ou simplesmente, *Royal Society*, fundada em 1660, ano da restauração da monarquia na Inglaterra, sendo aprovada pelo Rei Charles II em 1662<sup>11</sup>.

Seus fundadores foram influenciados por Francis Bacon, que em seu último livro descreveu as possíveis atividades de uma instituição de pesquisa. Um dos requisitos dessa instituição, de acordo com Bacon, seria dar prioridade à coleta e à análise de informações. Bacon sugeriu que alguns membros deveriam viajar ao exterior para coletar dados por meio de conversas e observações diretas, enquanto outros ficariam na sede lendo e fazendo resumos da literatura publicada no mundo (MEADOWS, 1999).

---

<sup>10</sup> Para mais informações sobre a Accademia dei Lincei, acessar <http://www.lincai.it/modules.php?name=Content&pa=showpage&pid=21> (acessado em 29/04/2012).

<sup>11</sup> Para mais informações sobre The Royal Society, acessar <http://royalsociety.org/about-us/history> (acessado em 18/07/2012).

Criada como local de pesquisa e debate, a *Royal Society* é possivelmente a mais antiga academia de ciências ainda ativa. Hoje ela representa a Academia de Ciências da Grã Bretanha e desempenha o papel de assessoria científica do governo britânico.

Em 1600 a Universidade de Paris tinha cerca de 30 mil estudantes, mas o currículo ainda era medieval, o ensino se fazia em Latim e muitos livros eram censurados. A Academia Francesa (literária) foi estabelecida em 1635 e a *Académie Royale des Sciences* em 1666.

Além do apoio do Estado e das doações privadas às atividades científicas, a formação das academias é também importante por ser uma das primeiras experiências coletivas de administração da ciência, não só financeira, mas estratégica, com as prioridades e linhas de ação sendo definidas por cientistas.

#### 1.4.3. Revistas Científicas

Em janeiro de 1665 foi publicada a revista *Journal des Sçavans*<sup>12</sup> na França e em março do mesmo ano a revista *Philosophical Transactions* da *Royal Society*<sup>13</sup>. O *Journal des Sçavans* deixou de ser publicado em 1792 durante a revolução francesa, retomando os trabalhos em 1816, mas como uma revista literária, sem material científico significativo. Por sua vez o *Philosophical Transactions of the Royal Society*, que, na realidade foi a primeira revista exclusivamente científica, continua ativa.

É importante destacar o real significado de uma revista para a comunicação científica. Podemos compreender a sua evolução lógica, com a ampliação geográfica do debate, antes somente oral, para a troca de correspondências em larga escala e com a organização da comunidade em Sociedades ou Academias. Como passo lógico, era necessário encontrar um modo mais eficiente de comunicação para atender a uma demanda cada vez mais crescente e interessada nos novos desenvolvimentos. Se definirmos como comunicação informal, uma comunicação mais restrita e de vida curta, como a comunicação oral e as correspondências, e a comunicação formal como algo disponível por longo tempo para um público amplo, como os livros, podemos entender que as revistas científicas surgiram no século XVII como um novo canal formal de comunicação científica (MEADOWS, 1999).

---

<sup>12</sup> A Biblioteca Nacional da França através do Gallica disponibiliza 244 anos (1665 a 1940) de publicações do *Le Journal des Sçavans on-line* <http://gallica.bnf.fr/ark:/12148/cb343488023/date.langPT> (acessado no dia 26 de julho de 2011).

<sup>13</sup> A revista *Philosophical Transactions* disponibiliza *on-line* todas as revistas desde 1665 até o presente (<http://royalsocietypublishing.org>).

#### 1.4.4. Obras de Destaque

Como exemplo de um trabalho importante para biodiversidade nesse período, tem-se a obra de Conrad Gesner (1516-1565) *Historiae animalum* que teve os quatro primeiros volumes publicados em Zurique de 1551 a 1558, enquanto o quinto e sexto volumes foram publicados após a sua morte (em 1587 e 1634, respectivamente). É considerada a base da zoologia moderna, sendo a primeira obra que procura descrever todos os animais conhecidos, inclusive animais da mitologia e do imaginário popular<sup>14</sup>. Os seis volumes contêm mais de 4.500 páginas e em seu prefácio, Gesner lista 52 colaboradores. O livro apresenta descrições científicas com trabalhos do próprio Gesner e colaboradores e uma revisão bibliográfica procurando incluir todos os animais descritos ao longo do tempo, com ou sem comprovação científica. Nas ilustrações apresentadas na Figura 1.1., tem-se a descrição completa de um peixe e uma ilustração de um rinoceronte indiano feita pelo artista Albrecht Dürer, o qual, embora nunca tivesse visto um rinoceronte, baseou-se em um desenho do animal, levado da Índia a Lisboa em 1515. A terceira ilustração é do monstro *Hydra* da mitologia grega, morto por Hércules; Gesner, apesar de considerar a hidra um ser imaginário, incluiu-o na obra por causa de um panfleto que havia sido recentemente publicado sobre uma cobra de sete cabeças encontrada na Turquia.



Figura 1.1. Exemplos das xilogravuras da obra *Historiae animalium* de Conrad Gesner.  
Fonte: *Historiae animalium*, 1551<sup>15</sup>

<sup>14</sup> Cambridge University Library – Rare Books – Featured book archive: Conrad Gesner *Historia animalium libri I-IV. Cum iconibus. Lib. I. De quadrupedibus uiuiparis*. Zurich: C. Froschauer, 1551. N\*. 1.19(A). [www.lib.cam.ac.uk/deptserv/rarebooks/gesner.html](http://www.lib.cam.ac.uk/deptserv/rarebooks/gesner.html) (acessado em 10 de abril de 2010).

<sup>15</sup> A obra de Conrad Gesner está disponível na National Library of Medicine no endereço <http://archive.nlm.nih.gov/proj/flash/gesner/gesner.html>

O trabalho de Gesner também representa um marco para a comunicação científica, uma vez que é resultante de um estudo colaborativo (52 colaboradores), que usou correspondências como principal meio de comunicação. Os primeiros quatro volumes deste trabalho foram publicados em sete anos, mas a obra completa levou 83 anos para ser concluída.

Uma outra importante iniciativa foi a criação dos primeiros jardins botânicos na Itália nos séculos XVI e XVII para o estudo de plantas medicinais<sup>16</sup>. Com o período das grandes navegações, jardins como Royal Botanic Gardens, Kew e Real Jardín Botánico de Madrid foram estruturados para cultivar novas espécies enviadas dos trópicos. Eram centros de pesquisa que trabalhavam no melhoramento das espécies de interesse econômico e dos métodos de cultivo. Esse trabalho ganhou grande importância nos séculos seguintes com o aumento do comércio internacional (BROCKWAY, 1979).

Nesse período existem poucos relatos sobre a história natural do Brasil, com exceção do breve período de dominação da Companhia das Índias Ocidentais no Nordeste. No período do governo do Conde João Maurício de Nassau-Siegen entre 1637 e 1644, vieram ao Nordeste cientistas, cartógrafos, artistas e administradores. Nesse período foi instalado o primeiro observatório astronômico do hemisfério sul e foram criados o primeiro Jardim Botânico e o primeiro Jardim Zoológico das Américas, ambos no Recife.

Nassau trouxe ao Brasil Frans Post e Albert Eckhout, que retrataram paisagens nordestinas, tipos humanos, a fauna e a flora, constituindo um importante legado artístico do período colonial. Trouxe também o naturalista alemão George Marcgraf, que permaneceu no Brasil de 1638 até 1643, integrando a expedição científica e militar de Nassau, como ajudante do médico Willem Piso. Marcgraf percorreu a região do Nordeste brasileiro (Pernambuco, Paraíba, Alagoas, Rio Grande do Norte, Ceará, Sergipe e Bahia), coletando plantas e animais, e foi considerado por Martius como o pai da História Natural brasileira (DIETRICH, 1996). Em 1644, já doente, Marcgraf foi enviado a Angola, morrendo ao chegar em Luanda. Mas, ao deixar o Brasil, tinha confiado suas observações ao conde Maurício de Nassau juntamente com os espécimes coletados, tomando o cuidado de escrever suas observações em códigos para impedir que Piso se apropriasse da sua obra<sup>17</sup>. O trabalho de Marcgraf foi decifrado e editado por Johannes de Laet. Piso publicou a obra *Historia Naturalis Brasiliae* em 1648, na qual quatro volumes foram

---

<sup>16</sup> <http://www.bgci.org/resources/history>

<sup>17</sup> Jstor Plant Science. Marcgrave, Georg (1610-1644) disponível em <http://plants.jstor.org/person/bm000048985>

escritos por ele (*De medicina Brasiliense*) e oito com as observações de Marcgraf (*Historia rerum naturalium Brasiliae*). Essa obra foi praticamente a única referência à botânica do Brasil durante mais de 150 anos.

### 1.5. SÉCULO XVIII

O século XVIII caracteriza-se como o século da revolução industrial. As estradas, a navegação e os serviços postais tiveram avanços significativos impulsionando o aumento do mercado, tanto europeu quanto internacional, principalmente da indústria têxtil com o algodão sendo importado dos Estados Unidos e da Índia.

Um marco na ciência é o nascimento do *British Museum* em 1753 cujo objetivo era colecionar material para servir de referência e de apoio à pesquisa. Na França, em 1790, nasce o *Muséum National d'Histoire Naturelle* de Paris, como um centro de ensino e pesquisa.

As academias se multiplicam e a ciência começa a se especializar, estimulada pelo grande avanço na indústria. São formadas academias especializadas em produções industriais que estabeleceram um elo forte entre a ciência e a tecnologia. Outras academias lançavam projetos competitivos e davam prêmios para os melhores inventos, como a *Society for the Encouragement of Arts, Sciences, and Commerce* (VICKERY, 2000)<sup>18</sup>. Surgem novas oportunidades de carreiras, projetos colaborativos e a avaliação entre pares (*peer assessment*), parte da essência da ciência.

A maioria das academias publicava seus anais, divulgando resumos e popularizando os avanços do conhecimento. Problemas como o tempo longo entre a apresentação de um trabalho na academia e a sua publicação e a quantidade de trabalhos de diferentes áreas de interesse e línguas levou à especialização das publicações.

Com a proliferação das academias e revistas e a sua especialização, muda também a linguagem. A ciência, agora como profissão organizada, passa a também desenvolver a sua linguagem especializada, passa a ter necessidade de uma maior exatidão dos termos.

É um século marcante para a taxonomia de plantas e animais devido à publicação da obra *Systema Naturae* escrito por Carl Linneus, um botânico, zoólogo e médico sueco. A obra apresenta uma classificação hierárquica das espécies, dividindo a natureza em três reinos: *Animalia*, *Vegetalia* e *Mineralia*. O livro foi publicado em latim e a primeira edição é de 1735. Na sua 10ª. edição, em 1758, Linnaeus inicia a aplicação geral da nomenclatura binomial

---

<sup>18</sup> É interessante observar essa experiência no século XVIII e compará-la à experiência de empresas no século XXI que hoje lançam desafios na Internet premiando a melhor solução.

zoológica. Esta data é aceita como ponto de partida da nomenclatura zoológica e da lei da prioridade, ou seja, o nome disponível mais antigo aplicado a um táxon é o nome considerado válido.

O trabalho de Linnaeus teve um impacto enorme na ciência. Trata-se da fundação da nomenclatura biológica que nos dias de hoje é regulamentada por códigos de nomenclatura. Linnaeus orientava os trabalhos de vários estudantes e os enviava para coletar material no exterior, o que tornou o seu trabalho muito mais influente. Sua obra está disponível *on-line* no Biodiversity Heritage Library<sup>19</sup>. Graças a este trabalho, uma espécie possui um único nome científico válido, reconhecido no mundo todo. Trata-se de uma chave importante para a interoperabilidade de sistemas de informação. Somente no XVIII Congresso Internacional de Botânica, realizado em Melbourne, Austrália em julho de 2011, é que foi votada uma resolução permitindo que a partir de janeiro de 2012 a descrição de uma nova espécie pudesse ser feita também em Inglês e publicada em uma revista eletrônica<sup>20</sup>.

Outro naturalista importante da época foi o francês Georges-Louis Leclerc, Conde de Buffon, precursor de Lamarck e Darwin, publicando a obra *Histoire naturelle, générale et particulière* (1749–1788)<sup>21</sup>. Buffon foi considerado o pai da biogeografia. Ele separou a história natural dos relatos bíblicos, propondo uma teoria da reprodução contrária à teoria da preexistência.

Em Portugal, em 1772, o Marquês de Pombal, secretário do Rei D. João I, decretou a reforma da Universidade de Coimbra, promovendo uma revolução nos estudos universitários com forte ênfase em ciências e história natural. Domingos Vandelli, amigo de Linnaeus, foi nomeado para a cátedra de história natural, dando início à formação de uma geração de naturalistas (PRIORE e VENANCIO, 2010).

Em 1783 teve início a expedição denominada *Viagem Filosófica* (1783-1792) do naturalista luso-brasileiro Alexandre Rodrigues Ferreira (1756-1815), aluno de Vandelli. É importante lembrar, que nessa época, Portugal proibiu as expedições às terras brasileiras de James Cook (1768-1771) e de Alexandre Von Humboldt e Aimé Bonpland (1799-1804). Ferreira veio a mando da Rainha D. Maria I de Portugal para conhecer melhor o Centro-Norte da colônia.

---

<sup>19</sup> Veja [www.biodiversitylibrary.org/bibliography/542](http://www.biodiversitylibrary.org/bibliography/542)

<sup>20</sup> Antes dessa resolução, as descrições só podiam ser feitas em Latim e publicadas em revistas especializadas impressas.

<sup>21</sup> Várias obras de Buffon estão disponíveis *on-line* na Biodiversity Heritage Library <http://www.biodiversitylibrary.org>

Seu objetivo era descrever, recolher, aprontar e remeter para o Real Museu de Lisboa amostras de utensílios empregados pela população local, bem como de minerais, plantas e animais. Ficou também encarregado de tecer comentários filosóficos e políticos sobre o que visse nos lugares por onde passasse. Durante nove anos, ele percorreu aproximadamente 39 mil quilômetros por remo e vela no interior da Amazônia. Foi o primeiro a reconhecer a importância de registrar os nomes e usos indígenas dos recursos naturais, sendo por isso considerado o primeiro etnobotânico da história da Amazônia. Várias vezes fez comentários sobre a imensa biodiversidade da região e lamentou o uso não sustentável dos recursos naturais pelos portugueses. A excursão rendeu milhares de páginas de diários, monografias e relatórios, aquarelas e desenhos, amostras da cultura de muitos grupos indígenas, produtos vegetais e documentação científica de peixes, mamíferos, plantas e pássaros (DALY, 2001). Tudo foi enviado à Lisboa, mas após o retorno de Ferreira a Portugal, durante 15 anos as coleções e manuscritos permaneceram desorganizados, sem publicações. Com a invasão de Portugal pelo exército de Napoleão em 1808, muitos manuscritos e documentos da viagem de Alexandre Ferreira foram levados. A maior parte da coleção foi devolvida sete anos depois, tempo suficiente para que naturalistas em Paris lançassem publicações sobre dezenas de espécies novas de plantas e animais da Amazônia, sem nunca terem ido lá. Ferreira faleceu no mesmo ano que seu material foi devolvido.

A maior parte da coleção etnobotânica de Alexandre Rodrigues Ferreira e das amostras de animais que ficaram em Lisboa foram perdidas em um incêndio no Museu Bocage na década de 1970. Seu acervo, incluindo aproximadamente 2.000 aquarelas e desenhos de indígenas, cidades, povoados, plantas, paisagens, animais, outros povos, minas e artigos de índios, está depositado no Herbário de Lisboa (LISU), na Biblioteca Nacional e no Museu Nacional do Rio de Janeiro.

Destaque deve ser dado aos desenhos produzidos na sua expedição. Alguns desenhos têm uma identificação direta com a obra de Buffon e acredita-se que alguns desenhos da Viagem Filosófica tenham sido usados na obra *Flora Brasiliensis* publicada entre 1840 e 1906 pelos editores Carl Friedrich Philipp von Martius, August Wilhelm Eichler e Ignatz Urban (MONTEIRO, 2010).

Outro brasileiro importante do século XVIII é o Frei José Mariano da Conceição Veloso (1742-1811), considerado pai da botânica do Brasil e autor da obra *Flora Fluminensis* (publicada após a sua morte), resultado de investigações científicas realizadas na província do Rio de Janeiro

durante oito anos, a pedido do Vice-Rei Luiz de Vasconcellos e Souza. Frei Veloso terminou o extenso levantamento da flora da região em 1790, descrevendo 400 novas espécies, sendo autor de 66 gêneros (LUNA, 2009). A partir do século XIX, quando as ciências naturais no Brasil começaram a conquistar o seu espaço por meio da criação dos museus, Veloso passou a ser considerado um dos principais cientistas brasileiros do século XVIII.

## 1.6. SÉCULO XIX

O Século XIX tem como grande marco o desenvolvimento da moderna indústria, principalmente na Alemanha, Estados Unidos e Grã Bretanha, com o aprofundamento da substituição da manufatura pela maquinofatura, sendo que as máquinas passam a ser movidas à energia (carvão e depois eletricidade, no final do século). A unidade industrial passa a ser majoritariamente a fábrica, que começa a oferecer um novo espaço de trabalho para cientistas. Nasce laboratórios de pesquisa e desenvolvimento no âmbito do processo da II<sup>a</sup> Revolução Industrial (ou Técnico-Científica para Braverman, 1974).

Marcante é o desenvolvimento do transporte e comunicações. A locomotiva a vapor foi responsável pela grande expansão das linhas ferroviárias, primeiro na Inglaterra e depois rapidamente expandindo por toda a Europa e outros países. Através de ferrovias e navios a vapor, a indústria da Europa alcançou novas terras e novos mercados, expandindo para os Estados Unidos, Canadá, Argentina, Uruguai, Brasil, África do Sul e Austrália.

Inúmeros inventos transformaram a comunicação científica no século XIX, como o telégrafo de William Cooke e Charles Wheatstone em 1837, a primeira máquina de datilografia moderna inventada por Christopher Latham Sholes em 1868 e fabricada pela Remington em 1873, o telefone elétrico (Alexander Graham Bell) e o mimeógrafo (Thomas Edison), ambos em 1876, e o filme Kodak para máquinas fotográficas patenteado por George Eastman em 1888.

A facilidade no transporte e comunicação das pessoas promoveu uma maior transferência de tecnologia e de conhecimento. Estima-se que, entre 1846 e 1932, 51 milhões de pessoas emigraram da Grã Bretanha, Itália e Alemanha para os Estados Unidos, Argentina, Canadá, Brasil, Austrália, África do Sul e Caribe. Esses emigrantes traziam conhecimento sobre sistemas de manufatura e habilidades técnicas. A transferência da tecnologia através da informação pessoal foi um importante meio de comunicação do conhecimento técnico (VICKERY, 2000).

As patentes passam a ganhar maior importância. A lei de patentes nos Estados Unidos foi aprovada em 1790, na França em 1791 e na Alemanha em 1842 e a primeira convenção

internacional foi celebrada em 1883. Na Inglaterra, cuja primeira patente data de 1552, o governo promoveu em 1850 ações para tornar pública a informação das patentes, imprimindo e distribuindo toda patente importante a partir de 1617 e estabelecendo uma biblioteca no escritório de patentes. As revistas começaram também a divulgar notícias sobre esses registros.

O século XIX marca também o crescimento do financiamento às instituições científicas pelo Estado. As universidades alemãs eram majoritariamente mantidas pelo Estado e eram orientadas para a formação de profissionais para a indústria com laboratórios de química e física. Em 1820 foram estabelecidas escolas técnicas. Em 1822 foi fundada a associação alemã de cientistas *Deutscher Naturforscher Versammlung*, que introduziu a prática de promover congressos científicos nacionais. O modelo alemão foi seguido por outros países que também criaram instituições de ensino voltados à formação de engenheiros.

Como importante instrumento de difusão, foram estruturadas as *Exposições Universais* que, em seu início, eram focadas no comércio e na exibição de invenções e avanços tecnológicos. Eram o local onde o estado-da-arte da ciência e tecnologia era reunida. A primeira grande exposição, aberta pela Rainha Victoria, foi *The Great Exhibition of the Works of Industry of all Nations* ou, simplesmente, *The Great Exhibition*, realizada em Londres de maio a outubro de 1851. Outras exposições se seguiram em Paris (1889), Chicago (1893), e em outras cidades até a data de hoje, sendo a última o Expo 2010 na cidade de Shanghai, China.

Outro importante espaço de interação e comunicação foi criado em 1899, a Associação Internacional de Academias (International Association of Academies – IAA)<sup>22</sup>, reunindo academias de ciências da Europa e dos Estados Unidos<sup>23</sup>.

Nas ciências naturais, de impacto internacional tem-se a obra de Charles Darwin (1809-1882) e sua teoria sobre a seleção natural como mecanismo básico da evolução de espécies. Darwin viajou no navio de pesquisas HMS Beagle do capitão FitzRoy no período de dezembro de 1831 a outubro de 1836, que teve por objetivo completar o levantamento da Patagônia e Terra do Fogo, percorrer o litoral do Chile, Peru e algumas ilhas do Pacífico, para fixar as longitudes com exatidão (TAYLOR, 2009).

---

<sup>22</sup> Para mais informações sobre a Associação Internacional de Academias, acessar <http://www.britac.ac.uk/pubs/src/found.cfm>

<sup>23</sup> A Sociedade Brasileira de Ciências somente foi criada em 1916 e em 1922 passou a se chamar Academia Brasileira de Ciências (ABC).

A narrativa da viagem publicada por FitzRoy em 1839 incluiu o relato de Darwin “*Darwin’s Journal and Remarks*” que se tornou muito popular. Darwin apresentou o seu trabalho no Geological Society of London, no Zoological Society e foi eleito *Fellow of the Royal Society* em 1839.

O duplo anúncio da teoria de seleção natural de Darwin e Alfred Russel Wallace foi realizado em um encontro do *Linnean Society* em Londres em julho de 1858, sem a presença dos autores (BROWNE, 2006). O livro *On the Origin of Species* foi publicado em novembro de 1859 em linguagem simples e se tornou muito popular.

Com o aumento das expedições científicas e o maior acesso a novas espécies, cresce a importância da botânica e dos jardins botânicos especialmente antes do advento da indústria química e da produção de substitutos sintéticos à matéria-prima. Em 1841 o *Royal Botanic Gardens Kew*, antes praticamente um hobby da realeza, torna-se uma instituição do Estado. No início focado em plantas medicinais, novas espécies de uso econômico são selecionadas e seu método de cultivo aprimorado. Muitas espécies foram distribuídas em colônias com mão de obra escrava tornando-se elas matéria-prima para produtos industrializados na Inglaterra (BROCKWAY, 1979).

No Brasil, merece destaque a vinda da família real portuguesa em 1808, o que promoveu a abertura dos portos, facilitando o intercâmbio intelectual com outros países e a vinda de numerosos naturalistas. Com a transferência da corte para o Brasil tem início a criação da infraestrutura necessária para o desenvolvimento da pesquisa no país. No dia 18 de fevereiro de 1808, em Salvador, foi inaugurada por D. João VI a primeira escola de ensino superior, a Faculdade de Medicina da Bahia. Em 13 de junho do mesmo ano, no Rio de Janeiro foi criado o Jardim de Aclimação com o objetivo de aclimatar as especiarias vindas das Índias Orientais. No dia 11 de outubro do mesmo ano, foi instalado o Jardim, denominado Real Horto, hoje Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, que foi aberto à visitação pública após 1822. O acervo da Real Biblioteca, com cerca de 60 mil peças foi transferido de Lisboa nos anos de 1810 e 1811, para a Biblioteca Nacional no Rio de Janeiro, fundada em 1810. No dia 6 de junho de 1818 foi criado o Museu Real por D. João VI, hoje denominado Museu Nacional, que é o maior museu de história natural e antropológica da América Latina<sup>24</sup>. A biblioteca do Museu Nacional

---

<sup>24</sup> Para mais informações sobre o Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, acessar <http://www.jbrj.gov.br> e sobre o Museu Nacional <http://www.museunacional.ufrj.br/MuseuNacional/Principal/omuseu.htm>

foi fundada em 1863 e a primeira revista científica brasileira de certa projeção - *os Arquivos do Museu Imperial* - teve início em 1876 (NOGUEIRA, 2000), 211 anos após a publicação da primeira revista científica na Europa.

Em 1866, foi fundada em Belém a Sociedade Philomática, por Domingos Soares Ferreira Penna, responsável pelo núcleo do futuro Museu Paraense (SCAFF, 1981). Em 25 de março de 1871, o Museu Paraense (hoje Museu Goeldi) foi instalado oficialmente pelo Governo do Estado. No final do século XIX, foram criados os primeiros institutos de pesquisa voltados à agricultura e saúde, como o Instituto Agrônômico de Campinas, fundado em 1887 pelo Imperador D. Pedro II, o Instituto Butantan, fundado em 1898 com o objetivo de produzir soro para combater a peste bubônica devido à epidemia no porto de Santos, e o Instituto Soroterápico Federal, hoje Fundação Oswaldo Cruz, fundado em maio de 1900 também com o objetivo de fabricar soros e vacinas contra a peste bubônica.

O século XIX foi importante para a ampliação do conhecimento da fauna e flora brasileira. Considerada uma das maiores obras botânicas de todos os tempos, a publicação da *Flora brasiliensis* teve início com a missão austríaca enviada ao Brasil em 1817, que além da princesa Leopoldina que veio para se casar com D. Pedro I, trouxe os cientistas Martius, Spix, Pohl, Schott, Ludwig Riedel, Mikan e Raddi com a finalidade de levar amostras botânicas e zoológicas para o Museu de História Natural da Áustria. O roteiro de viagem de Martius e Spix incluiu os estados de São Paulo, Rio de Janeiro, Minas Gerais, Bahia, Pernambuco, Piauí, Maranhão, Pará e Amazonas.

Voltando a Munique em 1822, Spix e Martius iniciaram os trabalhos de publicação dos resultados. Spix faleceu em 1826 e ficou a cargo de Martius a tarefa de publicar os resultados da viagem. Martius publicou as obras *Nova Genera et Species Plantarum* (3 volumes, 1823 – 1832), *Icones Selectae Plantarum Cryptogamicarum Brasiliensium* (1 volume 1827), *Historia Naturalis Palmarum* (3 volumes 1823 – 1853), entre outras. Sua obra mais ambiciosa foi a *Flora brasiliensis*, projeto iniciado em 1839 que contou com a colaboração de 65 autores, 38 alemães, 7 austríacos, 5 ingleses, 5 suíços, 4 franceses, 2 belgas, 2 dinamarqueses, 1 holandês e 1 húngaro. Não figura entre os colaboradores nenhum brasileiro, mostrando que não houve preocupação em formar escola no Brasil, nem foram consideradas as descrições feitas por brasileiros (SÁ, 2001). Foram publicados 140 fascículos com a descrição de 22.767 espécies. São 20.733 páginas e 3.811 pranchas que ilustram 6.246 espécies. O último volume foi publicado em 1906, 38 anos após a

morte de Martius. A obra foi digitalizada em 2006, como resultado de um projeto financiado pela Fapesp, Natura e Fundação Vitae, desenvolvido pelo CRIA (Centro de Referência em Informação Ambiental), Jardim Botânico de Missouri e Unicamp e está disponível *on-line*<sup>25</sup>.

Um trabalho recente (MONTEIRO, 2010) compara alguns desenhos botânicos da *Viagem Filosófica* com os da *Flora Brasiliensis* mostrando claramente que vários elementos foram copiados, o que comprova a consulta aos desenhos da Viagem Filosófica pelos artistas no período de preparação da obra *Flora Brasiliensis*. Copiar desenhos era usual para multiplicar o conhecimento entre cientistas, havendo especialistas neste ofício. “*O redesenho era o recurso utilizado na etapa de produção gráfica para conferir valor artístico à publicação e tornar os desenhos originais propensos à gravura. É notadamente a partir do século XIX e no século XX que surgem os critérios de valorização de desenhos originais*” (BELUZZO, 1994).

Um outro naturalista de destaque que chegou ao Brasil em 1816 foi o francês Auguste de Saint-Hilaire. Durante seis anos ele percorreu os Estados do Espírito Santo, Rio de Janeiro, Minas Gerais, Goiás, São Paulo, Santa Catarina, Rio Grande do Sul, além do rio da Plata e a Província de Missiones na Argentina e parte leste do Paraguai, retornando à Europa em setembro de 1822. Além do acervo botânico preservado no Museu Nacional de História Natural (MNHN) em Paris estão também disponíveis os seus cadernos de campo que, além de descrever a flora das diferentes regiões por que passou, relata a vida e os costumes brasileiros. Todo o seu acervo está sendo digitalizado e disponibilizado *on-line* graças à parceria entre o Instituto de Botânica de São Paulo, o *Muséum National d’Histoire Naturelle* (MNHN) de Paris, o *Institut des Herbiers Universitaires*, CLF, Clermont-Ferrand, França e o CRIA, com o apoio da Fapesp<sup>26</sup>.

A vinda dos naturalistas estrangeiros também causou críticas da comunidade científica do Brasil que reivindicava maior apoio governamental aos estudiosos nacionais. Esse movimento deu origem a um fórum para discutir e divulgar atividades e contribuições científicas, a Sociedade Velloziana, que se reuniu pela primeira vez em julho de 1850, 190 anos depois da fundação da *Royal Society* na Inglaterra.

Outro francês de destaque foi o botânico Auguste François Marie Glaziou<sup>27</sup>, que chegou ao Brasil, a convite de D. Pedro II, em 1858 para coordenar a Diretoria de Parques e Jardins da Casa Imperial no Rio de Janeiro, permanecendo no Brasil até 1897. Além dos seus trabalhos em

---

<sup>25</sup> A obra *Flora brasiliensis* está disponível no endereço <http://florabrasiliensis.cria.org.br>

<sup>26</sup> O Herbário Virtual A. de Saint-Hilaire está disponível no endereço <http://hvsh.cria.org.br/>

<sup>27</sup> Para mais informações, acessar <http://www.casaruibarbosa.gov.br/glaziou/biografia.htm>

jardins públicos e na Floresta da Tijuca, integrou, como botânico, a Comissão de Estudos do Planalto Central do Brasil. Além de descrever várias novas espécies do Brasil, devemos a Glaziou o uso de plantas nativas em praças e ruas do país. O material coletado por Glaziou está depositado principalmente no Museu Nacional de História Natural de Paris e no Museu Nacional do Rio de Janeiro.

Um brasileiro que merece destaque no período é João Barbosa Rodrigues, que esteve em missão científica na Amazônia entre 1872 a 1875 e dirigiu o Jardim Botânico de Manaus inaugurado em 1883 sob o patrocínio da Princesa Isabel e que foi extinto após a proclamação da república. Em 1890 tornou-se diretor do Jardim Botânico do Rio de Janeiro onde foi responsável pela criação da carpoteca e biblioteca e deu os primeiros passos para a criação do herbário<sup>28</sup>, hoje um dos mais importantes do país.

Em 1884 chega ao Rio de Janeiro, como subdiretor da seção zoológica do Museu Nacional, o suíço Emílio Goeldi (Emil August Göldi, 1859-1917). Diferente dos naturalistas que vieram ao Brasil para coletar material, documentar suas impressões e retornar ao país de origem, Goeldi constituiu família no Brasil, residindo cerca de 23 anos no país, sem perder os laços com a comunidade científica da Europa. Além dos trabalhos taxonômicos, realizou várias pesquisas aplicadas, estudando problemas locais como, por exemplo, uma praga do café causada por um nematóide (*Meloidogyne exigua*), uma pesquisa de repercussão internacional.

No Museu Paraense, onde foi diretor de 1894 a 1907, procurou desenvolver linhas de pesquisas especializadas nas questões amazônicas, estruturando coleções científicas e construindo o jardim zoológico e o horto botânico. Além dos trabalhos com inventários biológicos, Goeldi traduziu muitas obras, criou uma biblioteca local especializada e uma revista científica em português, e instituiu um programa de viagens, mantendo uma forte interação com pesquisadores e instituições europeias. Além das atividades de pesquisa com a criação de laboratórios e coleções científicas, Goeldi transformou o museu em local de visitação pública.

Outro fato muito relevante foi a sua atuação na disputa franco-brasileira pelo território do Amapá conhecida como Contestado franco-brasileiro, quando Goeldi foi assessor do Barão do Rio Branco. Pelo Tratado de Arbitramento (1897), França e Brasil deveriam habilitar o árbitro a proferir a sentença, ou seja, deveriam enviar informações para substanciar a sua decisão. Goeldi teve um papel fundamental ao reunir documentos e informações para comprovar a posse

---

<sup>28</sup> Ver [http://www.jbrj.gov.br/historic/barb\\_rod.htm](http://www.jbrj.gov.br/historic/barb_rod.htm)

portuguesa e depois brasileira da região. Ele foi à região e coletou espécimes e objetos, entrevistou e fotografou os moradores, retratando também as vilas e paisagens e produzindo relatórios e mapas.

Além dos relatórios apresentados ao governador do Pará e ao Ministro das Relações Exteriores, Goeldi e sua equipe publicaram doze trabalhos científicos em periódicos da Suíça, Alemanha e Áustria. Ele foi à Europa onde apresentou o seu trabalho, procurando também mostrar o domínio do território pelo Brasil, sem falar no Contestado. A divulgação do seu trabalho entre seus pares deu aos seus relatórios a credibilidade necessária, e acabaram sendo utilizados como fonte para a descrição física e geográfica da região pelos árbitros suíços que, ao final, deram um parecer favorável ao Brasil (SANJAD, 2009).

O conhecimento científico na área de ciências naturais que já vinha sendo aplicado na medicina e saúde, também passa a ter importância na política e na tomada de decisão. Outro fato importante é que, a partir do século XIX, existe uma tendência exponencial de especialização da ciência (POMBO, 2005).

## 1.7. SÉCULO XX

O século XX é um período de grandes transformações na sociedade em geral e na comunicação científica em particular principalmente após a IIa. Guerra Mundial que trouxe profundas mudanças na concepção do papel da ciência como estratégia de Estado. Um marco foi o relatório *Science the endless frontier* coordenado por Vannevar Bush, diretor da Secretaria de Pesquisa e Desenvolvimento Científico do governo federal americano. O relatório salientou a importância do fluxo de novo conhecimento científico como instrumento de fomento à pesquisa industrial e da formação de recursos humanos, considerado como o elemento que promove ou limita o avanço científico (BUSH, 1945). Também mereceram destaque os incentivos e o fortalecimento do sistema patentário. O relatório recomendou a criação de uma nova agência permanente de fomento, a “*National Research Foundation*”, para apoiar a pesquisa básica na academia e a educação científica (graduação e pós-graduação) com fundos estáveis, para permitir o desenvolvimento de programas de longo prazo. Em maio de 1950 foi criada a *National Science Foundation*.

O Brasil acompanhou esse movimento pós-Guerra quando a ciência passou a ser considerada estratégica para o Estado, criando em 1951 o Conselho Nacional de Pesquisa

(CNPq<sup>29</sup>) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes). A FAPESP, cuja criação estava prevista na Constituição Estadual Paulista de 1947, foi fundada em 1962.

Em 1950 foi criada a Sociedade Botânica do Brasil (SBB), em 1956 a Sociedade Brasileira de Microbiologia (SBM) e em 1978 a Sociedade Brasileira de Zoologia. Em 1973 foi criada a Secretaria Especial do Meio Ambiente transformada em ministério em 1993.

Essas ações, especialmente das sociedades científicas, têm sido fundamentais na organização da comunidade científica para o desenvolvimento do sistema brasileiro de informação *on-line* sobre biodiversidade, descrito no Capítulo 3.

### 1.7.1. O Movimento Ambientalista

É também após a II Guerra Mundial que a questão ambiental começa a ganhar relevância. Um trabalho que causou grande impacto foi o relatório apresentado ao Clube de Roma em 1972 – *The Limits of Growth* (MEADOWS, MEADOWS, *et al.*, 1972) que concluiu que se os níveis de crescimento da população, indústria, alimentos e esgotamento dos recursos naturais fossem mantidos, o limite de crescimento do planeta seria alcançado em 100 anos e propôs ações para que fosse estabelecido um estado de equilíbrio.

Na década de 1970 foram realizadas várias conferências importantes e foi criada a UNEP - United Nations Environment Programme (em português PNUMA - Programa das Nações Unidas para o Meio Ambiente).

Em 1987 foi apresentado outro relatório importante, *Our Common Future* ou Relatório Brundtland (WCED, 1987), onde foi cunhado o conceito de *desenvolvimento sustentável*. Em 1988 tiveram início os trabalhos do IPCC (The Intergovernmental Panel on Climate Change), órgão intergovernamental estabelecido pela ONU, UNEP e pela Organização Meteorológica Mundial para avaliar informações científicas, técnicas e sócio-econômicas, com o objetivo de compreender os riscos de mudanças climáticas induzidas pelo homem.

Em 1992 foi realizada a Conferência das Nações Unidas para o Meio Ambiente e o Desenvolvimento (também conhecida como ECO-92 e Rio-92) que teve como produtos a Conferência da Diversidade Biológica (CDB), a Conferência do Clima e a Agenda 21. A CDB tem como base a conservação da biodiversidade, o uso sustentável de seus componentes e a

---

<sup>29</sup> Hoje denominado Conselho Nacional de Desenvolvimento Científico e Tecnológico

distribuição justa e equitativa dos benefícios derivados do uso dos recursos genéticos e é um dos principais instrumentos internacionais relacionados ao meio ambiente.

O intercâmbio de informações, a cooperação técnica e científica, a educação e a conscientização pública são apenas alguns dos compromissos assumidos pelo governo brasileiro ao ratificar a CBD em 1998<sup>30</sup>. A informação científica sobre biodiversidade passa a ser considerada estratégica, não só para o desenvolvimento científico, mas também para processos de tomada de decisão e para a formulação de políticas em escala global.

### 1.7.1. A Internet e Sistemas de Informação em Biodiversidade

O intercâmbio de informações e a cooperação científica foram beneficiados pela evolução das TICs que viabilizaram não só uma maior acessibilidade à informação sobre biodiversidade, mas também o desenvolvimento de redes colaborativas.

Apesar de ser a Internet fruto de um programa militar norte-americano, seu desenvolvimento se deu nas universidades. Os protocolos e as ferramentas, incluindo correio eletrônico, TCP/IP, telnet e ftp<sup>31</sup>, à medida que foram desenvolvidos foram disponibilizados na rede para os pesquisadores do projeto. À medida que outros grupos de pesquisadores foram se integrando à rede, esses aplicativos passaram a ser disponibilizados a todos os interessados. Esse fato, além de estimular e promover uma maior utilização dessas ferramentas, iniciou um hábito cultural da Internet de que *software* disponível na rede em domínio público, aparecesse como *freeware* ou na forma de *shareware*<sup>32</sup> (CANHOS e CANHOS, 2001).

A arquitetura desenvolvida para a Internet catalisou o desenvolvimento de uma cultura colaborativa. A nova tecnologia adotada partia do conceito de comutação de pacotes. O sistema evitaria o modelo centralizado e os “nós” seriam interligados por caminhos redundantes utilizando diversos meios físicos (cabo, rádio, satélite) e teriam autonomia para gerar, transmitir e receber mensagens (LEINER, CERF, *et al.*, 1997). Ao invés da uniformização do *hardware* e do sistema operacional, optou-se por desenvolver um protocolo comum de transmissão de dados

---

<sup>30</sup> Decreto 2.519, de 16/3/1998

<sup>31</sup> TCP/IP significa *transmission control protocol/Internet protocol* e é um conjunto de protocolos de comunicação usado pela Internet; Telnet é um protocolo cliente-servidor usado para permitir a comunicação entre computadores ligados em rede; e ftp (*file transfer protocol*) é um protocolo de transferência de arquivos entre computadores na Internet.

<sup>32</sup> Freeware é *software* gratuito, podendo ser *livre* (código fonte disponível) ou não, podendo ainda ter licenças com restrições proibindo o uso comercial ou a redistribuição e modificação não autorizada. O *software* denominado *shareware* normalmente tem uma taxa (muitas vezes simbólica) de uso ou uso gratuito de funcionalidades ou tempo limitados.

para as diferentes plataformas, o TCP/IP. Essa visão, de reconhecer que existem plataformas diferentes e que é necessário desenvolver protocolos comuns de comunicação, até hoje norteia o desenvolvimento não só da Internet, mas dos principais sistemas de informação distribuída *on-line*.

Os países começaram a investir na espinha dorsal (*backbone*) da rede criando as redes acadêmicas. Graças à arquitetura desenvolvida, as instituições começaram a se conectar à Internet pagando apenas o custo da ligação do seu próprio computador ao nó mais próximo, usando o *hardware* e ambiente computacional de que dispunham, o que reduziu, de forma significativa, o custo necessário para sua conexão.

O primeiro acesso à Internet no Brasil foi realizado em fevereiro de 1991, através da Fapesp e sua rede ANSP (*Academic Network at São Paulo*). O primeiro *backbone* da RNP (Rede Nacional de Ensino e Pesquisa) interligando 11 capitais foi lançado no início de 1992.

Vários aplicativos e protocolos foram desenvolvidos para facilitar o acesso a documentos na Internet. Em 1991 foi lançado por pesquisadores da Universidade de Minnesota o *Gopher*, uma interface amigável e muito útil para a organização de diretórios de documentos. Mas a verdadeira mudança veio em 1993 com o lançamento do *Mosaic*, navegador (*browser*) do *World Wide Web* (WWW), como *software* livre.

O fato de ser gratuito, ter interface gráfica e ser capaz de acessar sistemas que usassem os demais protocolos existentes, fez com que rapidamente se transformasse na principal via de acesso a sistemas de informação (VIEIRA, 2003). Antes do lançamento dessa interface, o usuário tinha que saber como usar cada sistema disponível *on-line*. Com o *Mosaic*, bastava saber usar o *Mosaic*, que na maioria das vezes apenas apresentava links clicáveis para acessar qualquer sistema ou arquivo. Caberia ao provedor dos dados e informações preparar o seu servidor de forma que ele pudesse ser acessado pelo navegador, ou seja, não era mais exigido do usuário conhecimento especializado sobre o sistema a ser acessado, bastava saber usar um único *software*.

Uma das primeiras iniciativas que reconheceram a Internet como um novo paradigma para divulgação e disseminação de dados sobre biodiversidade foi o ERIN, *The Australian Environmental Resources Information Network*, que em 1992, iniciou um programa de digitação dos acervos de herbários da Austrália e disponibilizou os dados *on-line* integrados a uma base cartográfica. O trabalho desenvolvido pelo ERIN serviu de exemplo para a CONABIO (Comisión

Nacional para el Conocimiento y Uso de la Biodiversidad)<sup>33</sup>, estabelecida em 1992 no México e do INBio (Instituto Nacional de Biodiversidad)<sup>34</sup>, uma associação privada sem fins lucrativos da Costa Rica, que se estabeleceu em 1989.

No início de 1992, antes da Eco 92, foi organizada uma reunião no Brasil com especialistas da Austrália, Estados Unidos, Canadá, e Europa para discutir uma estratégia de coordenação de ações sobre sistemas de informação em biodiversidade. Essa iniciativa, denominada *BIN21 – Biodiversity Information Network, Agenda 21*, propunha a criação de um diretório que desse visibilidade às diferentes iniciativas existentes nessa área.

Após a Eco-92, a BIN21 passou a ter como objetivo dar suporte à Convenção sobre a Diversidade Biológica, facilitando o acesso eficiente à informação relacionada a todos os aspectos da biodiversidade. Tinha como foco de sua ação o estabelecimento de *links* a fontes de dados, a promoção do conceito da biodiversidade e o estímulo ao desenvolvimento de bancos de dados e redes, esperando, dessa forma, estimular o desenvolvimento científico e a conservação<sup>35</sup>. Foi estabelecida como uma rede de interesse especial (*Special Interest Network*), na época um novo paradigma para cooperação que estava surgindo com as redes internacionais de computadores (GREEN e CROFT, 1994).

Em 1995 foi realizado um workshop no Brasil para discutir o documento UNEP/CBD/COP/2/6 *Establishment of the Clearing-House Mechanism (CHM) to Promote and Facilitate Technical and Scientific Cooperation* que tinha como proposta o estabelecimento de um centro de informação que seria responsável único pela captação e divulgação da informação de interesse para a CDB de forma centralizada. O relatório da reunião<sup>36</sup> recomendou o desenvolvimento de uma rede distribuída e foi base para a proposta brasileira apresentada na reunião de especialistas *Building the Clearing-House Mechanism* realizada em Bonn, Alemanha de 25 a 29 de Junho de 1997, e na terceira reunião do SBSTTA (*Subsidiary Body for Scientific, Technical and Technological Advice*) em setembro de 1997 (DIAS, GROVER, *et al.*, 1997). O CHM foi estabelecido como uma rede distribuída com pontos focais nacionais e um nó central no Secretariado da CDB. O ponto focal brasileiro para o CHM é o Ministério do Meio Ambiente.

---

<sup>33</sup> <http://www.conabio.gob.mx>

<sup>34</sup> <http://www.inbio.ac.cr>

<sup>35</sup> BIN21 Introduction, disponível no endereço <http://www.anbg.gov.au/bin21/introduction.html> (acessado no dia 01/08/2011)

<sup>36</sup> Report of the 1995 BIN21 Meeting disponível no endereço <http://www.cbd.int/doc/ref/chm-cbd-en.pdf> (acessado no dia 01/08/2011)

Em 1997, já pensando na estruturação mundial de dados sobre biodiversidade como nós distribuídos, teve início o projeto para a implantação da Rede Brasileira de Informação em Biodiversidade – BINbr (Biodiversity Information Network – Brazil) com financiamento do GEF/Banco Mundial através do Ministério do Meio Ambiente, concluído no ano 2000 (CANHOS, CANHOS, *et al.*, 2000).

No mesmo período da discussão e implementação da BINbr foi estabelecido um grupo de coordenação para discutir um programa sobre biodiversidade para o Estado de São Paulo, o Biotasp – Biota São Paulo. Durante os três anos de discussão para a formulação de uma proposta de programa (1996 a 1998), o grupo de coordenação usou ferramentas *on-line* para estruturar e dar publicidade às discussões. O trabalho foi apresentado em um workshop<sup>37</sup> onde foram formados grupos de trabalho para a definição das ações e elaboração final do programa a ser apresentado à Fapesp. Nascia, em março de 1999, o Programa de Pesquisas em Conservação Sustentável da Biodiversidade, denominado "Programa Biota-Fapesp, O Instituto Virtual da Biodiversidade" (JOLY e SPEGLICH, 2003).

O fato inovador foi a concordância da comunidade científica em compartilhar dados não sensíveis, não somente com seus pares, mas disponibilizando-os na Internet de forma livre e aberta a todo usuário interessado, sem distinção. Um dos primeiros projetos aprovados foi o de desenvolver um sistema de informação capaz de receber e disseminar os dados dos inventários realizados com recursos do programa, o SinBiota - Sistema de Informação Ambiental do Programa Biota-Fapesp. O SinBiota foi desenvolvido com uma interface de acesso restrito para que os pesquisadores do programa pudessem depositar seus dados. A interface pública permite a busca e recuperação dos dados sobre a ocorrência geográfica de espécies em São Paulo, integrada com a base cartográfica do estado<sup>38</sup>. O sistema desenvolvido pelo CRIA – Centro de Referência em Informação Ambiental em colaboração com a Unicamp e Instituto Florestal, hoje está hospedado no Centro Nacional de Processamento de Alto Desempenho em São Paulo (Cenapad-SP), localizado na Unicamp.

---

<sup>37</sup> Workshop: Bases para a Conservação da Biodiversidade do Estado de São Paulo, 1997 disponível no endereço <http://www.biota.org.br/info/historico/workshop> (acessado no dia 02/08/2011)

<sup>38</sup> Sistema de Informação Ambiental do Programa Biota-Fapesp (SinBiota) disponível no endereço <http://sinbiota.biota.org.br/atlas> (acessado no dia 02/08/2011)

No âmbito internacional também merece destaque o Fórum de Megaciência da OCDE<sup>39</sup> (Organização para a Cooperação e Desenvolvimento Econômico), originalmente constituído por um grupo de informática em biologia que posteriormente foi desdobrado em dois subgrupos: informática em biodiversidade e neuroinformática.

O subgrupo *biodiversidade* reuniu-se entre 1996 e 1998 e apresentou um relatório em janeiro de 1999<sup>40</sup> indicando ser necessário integrar os recursos em informática e as pessoas que deles se utilizam, em um sistema interoperável, o que tornaria a informática em biologia um esforço de megaciência. Para biodiversidade, a recomendação primária foi a criação de um mecanismo internacional que teria como missão tornar dados e informações sobre biodiversidade mundialmente acessíveis. Os governos dos países da OCDE deveriam estabelecer e dar suporte a um sistema distribuído de *módulos* (bancos de dados, *software*, ferramentas de rede, sistemas de busca, algoritmos etc.) interconectados e interoperáveis que juntos formariam a infraestrutura global de dados sobre biodiversidade, o *Global Biodiversity Information Facility* (GBIF). Conceitualmente, esse foi um marco histórico, pois colocou a questão da infraestrutura distribuída de dados de acesso público no foco da política global da OCDE.

Com o desenvolvimento da Internet no final do século pode-se perceber claramente um movimento de articulação e organização da comunidade científica e dos governos no sentido estruturar sistemas *on-line* como ferramenta para a comunicação científica. Esse movimento se consolida e apresenta avanços significativos ainda no início do século XXI.

Castells (1999) considera a revolução das tecnologias de informação e comunicação como um evento histórico da mesma importância da Revolução Industrial do século XVIII, induzindo a um padrão de descontinuidade nas bases materiais da economia, sociedade e cultura.

## 1.8. SÉCULO XXI

A evolução das tecnologias de informação e comunicação (TICs) tem conduzido o processo de mudança da constituição e disponibilização de dados e informações, estando os sistemas de dados digitais no centro dessa mudança de paradigma. Eles viabilizam análises com níveis de precisão e sofisticação sem precedentes, possibilitando uma nova compreensão por

---

<sup>39</sup> O *Megascience Forum* da OCDE hoje é denominado *Global Science Forum*

<sup>40</sup> Veja o documento: *Final Report of the OECD Megascience Forum Working Group on Biological Informatics*, Janeiro, 1999 ([www.gbif.org/fileadmin/Temp\\_for\\_New\\_Web\\_Site/BIrepfin.pdf](http://www.gbif.org/fileadmin/Temp_for_New_Web_Site/BIrepfin.pdf)).

meios inovadores de integração da informação (NSB, 2005). Novas disciplinas científicas nem existiriam sem essas tecnologias.

Nos séculos passados, estudos ambientais eram desenvolvidos por naturalistas, com uma visão mais holística e interdisciplinar, observando não somente a fauna e a flora, mas a geografia, a geologia, a física etc. A partir do século XIX houve uma grande especialização e essa abordagem reducionista ajudou a revelar os processos básicos – moleculares, celulares, fisiológicos, ecológicos – que regem a vida.

Mas, a partir da metade do século XX a ciência deixou de ser pensada como linear e cresce a demanda para a resolução de problemas. A ciência passa a ser influenciada pela política e pela economia. A especialização tem que ser complementada ou substituída por uma compreensão interdisciplinar, e novas disciplinas são constituídas, na fronteira de disciplinas, como a bioquímica, e inter-disciplinas, como a engenharia genética (POMBO, 2005).

Os estudos especializados estão avançando, mas paralelamente, biólogos estão adquirindo capacidade para ir além das interações entre componentes de um determinado nível da organização biológica. É o momento da biologia entrar no mundo não linear (NRC, 2009). Trata-se de um dos desafios de uma infraestrutura de dados, a reintegração das subdisciplinas da biologia, junto com uma integração maior com outras áreas do conhecimento como se vê na Figura 1.2.

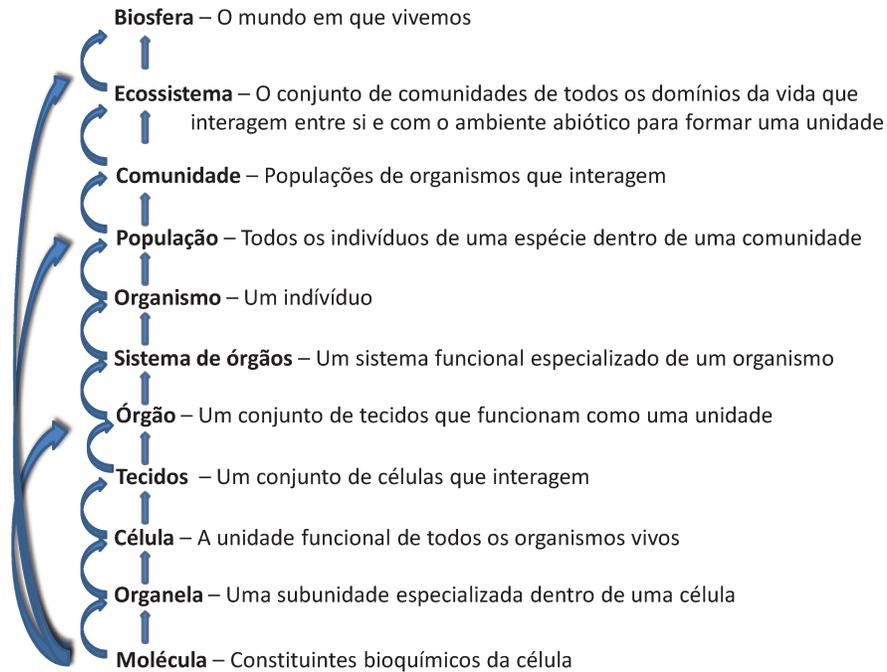


Figura 1.2. Os níveis de complexidade da biosfera (NRC, 2009)

A integração dos dados das subdisciplinas tecnicamente é possível graças aos desenvolvimentos das TICs. Uma característica do início do século XXI é a crescente quantidade e complexidade de dados, representando ao mesmo tempo uma oportunidade e um desafio (WOOD, ANDERSSON, *et al.*, 2010). Antes do enorme avanço tecnológico na computação e instrumentação, os dados eram gerenciados, lidos e interpretados por um especialista ou por um número pequeno de especialistas e as descobertas também eram feitas individualmente ou por pequenos grupos. Dava-se publicidade ao resultado da interpretação e análise dos dados na forma de livros e revistas científicas. Hoje os dados são gerenciados por equipes, e muitos deles não podem sequer ser lidos por indivíduos, mas por computadores, sendo necessários *softwares* para a sua interpretação. As descobertas cada vez mais resultam das análises de programas, que depois são verificadas e validadas por pessoas (RUMBLE, 2005).

### 1.8.1. Movimento do Acesso Livre e Aberto a Dados

Nesse início de século cresce o movimento em prol do acesso livre e aberto a dados. Em 2004, o Comitê de Política para Ciência e Tecnologia da OCDE discutiu o tema *Ciência, Tecnologia e Inovação para o Século XXI* (OECD, 2004). Os ministros de países membros da

OCDE adotaram a *Declaração sobre o Acesso a Dados de Pesquisa realizadas com Financiamento Público* (OECD, 2007) na qual, entre outros pontos, reconhecem que:

- o intercâmbio de dados, informações e conhecimento contribui de forma decisiva para o avanço da pesquisa científica e da inovação;
- o acesso aberto e o uso irrestrito de dados promove o progresso científico e facilita o treinamento de pesquisadores;
- o acesso aberto maximiza o valor derivado dos investimentos públicos em esforços de coleta de dados;
- os benefícios que a ciência, a economia e a sociedade de maneira geral poderiam obter das oportunidades que o uso de dados digitais têm a oferecer, reconhecendo que restrições indevidas ao acesso e uso de dados de pesquisa financiada com recursos públicos, podem diminuir a qualidade e eficiência da pesquisa científica e da inovação;
- a disponibilidade de dados de pesquisas realizadas com financiamento público, para países em desenvolvimento, promoverá a sua participação no sistema global de ciência, dessa forma contribuindo para o seu desenvolvimento social e econômico
- é necessário respeitar as leis nacionais, protegendo a privacidade dos cidadãos e as leis de propriedade intelectual; e;
- medidas precisam ser tomadas com respeito a alguns aspectos de acessibilidade a dados científicos obtidos com recursos públicos, uma vez que as disparidades nos regulamentos nacionais podem prejudicar o uso de dados obtidos com financiamento público em escala nacional e internacional.

Ao reconhecer esses pontos, declaram o compromisso de trabalhar no sentido de estabelecer regimes de acesso a dados digitais de pesquisa financiada com recursos públicos, de acordo com os objetivos e princípios de acesso livre e aberto, transparência, conformidade legal, profissionalismo, proteção à propriedade intelectual, interoperabilidade, qualidade e segurança, eficiência e responsabilidade.

Além da OCDE, existem várias iniciativas que têm por objetivo tornar dados e informações disponíveis de forma livre e aberta, sem qualquer restrição. A *open archives*

*initiative*<sup>41</sup> (ou iniciativa de arquivos abertos) nasceu como um movimento da comunidade científica, visando ao acesso gratuito a publicações. O Instituto da Sociedade Aberta (*Open Society Institute*) em 2001 organizou uma reunião em Budapeste (Budapest Open Access Initiative<sup>42</sup>) com o objetivo de acelerar o progresso dos esforços internacionais em tornar artigos de pesquisa de todos os campos acadêmicos, livremente disponíveis na Internet. Por acesso aberto a essa literatura, entende-se a disponibilidade livre na Internet pública, permitindo a qualquer usuário ler, copiar, distribuir, imprimir, pesquisar, criar *links* para o documento e utilizá-lo para quaisquer fins lícitos, sem barreiras financeiras, legais ou técnicas. A única restrição diz respeito à manutenção da integridade da obra e ao direito autoral, devendo este ser reconhecido e citado. Graças ao apoio da Mellon Foundation e da NSF, foi criada a *Open Archives Initiative* para o desenvolvimento de padrões e protocolos para facilitar a disseminação eficiente dos artigos.

Outra iniciativa que merece destaque é o *Creative Commons*<sup>43</sup> que disponibiliza licenças para obras intelectuais, oferecendo opções mais flexíveis de proteção. Tem como objetivo aumentar a disponibilidade de obras por meio de licenças que permitem a cópia e o compartilhamento menos restritivo. Autores podem, dessa forma, abrir mão de seus direitos, parcialmente ou integralmente, escolhendo a licença que lhes for mais adequada. Dada a sua popularização, o projeto foi ampliado para setenta países<sup>44</sup>, sendo que no Brasil é representado pelo Centro de Tecnologia e Sociedade da Faculdade de Direito da Fundação Getúlio Vargas, no Rio de Janeiro.

Em 2003, o MinC adotou a licença *Creative Commons* visando uma maior circulação das obras criadas sob o patrocínio do governo federal, mas o novo ministro que assumiu em 2011 retirou o selo da licença do site, gerando grande polêmica. Por sua vez, a Secretaria Municipal da Educação de São Paulo colocou todo o conteúdo do seu site sob uma licença *Creative Commons* que permite compartilhar (copiar, distribuir, transmitir) e *remixar* (criar obras derivadas) todo o material disponível. Cabe ao usuário atribuir a autoria, sendo vedado o uso comercial. Essa

---

<sup>41</sup> Para maiores informações sobre o Open Archives Initiative, acessar <http://www.openarchives.org>

<sup>42</sup> Para maiores informações da Budapest Open Access Initiative, acessar <http://www.soros.org/openaccess/index.shtml>

<sup>43</sup> Para mais informações sobre a iniciativa Creative Commons, acessar <http://www.creativecommons.org.br>

<sup>44</sup> Veja CC Affiliate Network disponível no endereço [http://wiki.creativecommons.org/CC\\_Affiliate\\_Network](http://wiki.creativecommons.org/CC_Affiliate_Network) (acessado em 31 de outubro de 2011)

licença está mais próxima ao objetivo da Ciência Aberta, em que o autor normalmente busca apenas o reconhecimento público.

Contra as iniciativas de políticas abertas, existe também a tendência de políticas mais restritivas para assegurar o direito à propriedade. Para e-infraestruturas, além do direito autoral que pode dificultar o acesso a dados e artigos científicos e das patentes muito abrangentes, preocupa o uso de software, ferramentas e padrões proprietários. É importante que o desenvolvimento das e-infraestruturas ocorra em ambientes abertos para viabilizar a livre troca de informações, dados e ferramentas que possam contribuir para o avanço do conhecimento em áreas estratégicas, como é o caso da diversidade de espécies.

Dois fatos recentes criam um novo marco para o acesso livre e aberto a dados científicos. No Brasil, a Lei de Acesso à Informação<sup>45</sup> e a sua regulamentação<sup>46</sup> em maio de 2012 indica ser dever dos órgãos e entidades promover, independentemente de requerimento, a divulgação na Internet de informações de interesse coletivo ou geral por eles produzidas ou custodiadas. O acesso livre não se aplica no caso das informações referentes a projetos de pesquisa e desenvolvimento científicos ou tecnológicos, cujo sigilo seja imprescindível à segurança da sociedade e do Estado.

Nos Estados Unidos, no dia 22 de fevereiro de 2013, o Escritório de Política em Ciência e Tecnologia (Office of Science and Technology Policy - OSTP) enviou um memorando a todos os chefes de departamentos executivos e agências cujo assunto era “Aumentando o acesso aos resultados das pesquisas científicas financiadas com recursos federais”<sup>47</sup>. Comunica a alocação de recursos (US\$ 100 milhões) para que cada agência apresente ao OSTP, em um prazo de seis meses, o seu plano preliminar para maximizar o acesso público e aberto a dados e publicações resultantes de pesquisas realizadas com financiamento federal.

Outro importante documento que deverá balizar a política de dados abertos da Grã Bretanha é apresentado pela The Royal Society (2012). O relatório recomenda o acesso livre e aberto de dados e modelos desenvolvidos por cientistas, apontando barreiras e sugerindo políticas para a abertura de dados científicos que complementam políticas para dados governamentais (*open government data*). Indica que a Internet oferece novas oportunidades de colaboração e

---

<sup>45</sup> Lei Federal no. 12.527 de 2011 ([http://www.planalto.gov.br/ccivil\\_03/\\_Ato2011-2014/2011/Lei/L12527.htm](http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2011/Lei/L12527.htm))

<sup>46</sup> Decreto no. 7.724 e 16 de maio de 2012 ([http://www.planalto.gov.br/ccivil\\_03/\\_Ato2011-2014/2012/Decreto/D7724.htm](http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2012/Decreto/D7724.htm))

<sup>47</sup> Increasing Access to the Results of Federally Funded Scientific Research ([http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp\\_public\\_access\\_memo\\_2013.pdf](http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf))

cooperação entre cientistas profissionais e amadores que poderá abrir o caminho para uma segunda revolução científica aberta.

### *1.8.2. Iniciativas Inovadoras no Uso das TICs*

As tecnologias para produção, gestão, armazenamento, recuperação e disseminação de dados e informações estão evoluindo rapidamente, assim como as demandas e necessidades científicas. As redes digitais criam as condições para o livre fluxo de dados e informações, permitem a democratização e o compartilhamento do conhecimento, criam novas formas de relacionamento e estabelecem um novo padrão de produção de bens (OLIVO, 2004).

O termo *Web 2.0* está associado a um novo conceito de desenvolvimento da rede que inclui a interação, interoperabilidade, o compartilhamento e a colaboração. A característica fundamental do *Web* anterior à família 2.0 é o acesso ao conteúdo de diferentes sistemas na Internet através de uma única interface, um acesso do tipo *read only*. Graças ao desenvolvimento dos novos aplicativos para o *Web* a partir de 2002/3, foi dada a possibilidade aos usuários de interagir e colaborar através de uma mídia social, e fazer parte de uma comunidade virtual. De usuários ou consumidores passivos de conteúdo, passam a ser membros ativos de sites de redes sociais como blogs, twitters e facebook e a usar sistemas para compartilhar e veicular vídeos (youtube), fotos (flickr) e arquivos (dropbox), abrindo um enorme leque de possibilidades em relação à construção colaborativa de plataformas distribuídas de dados. Começa a ganhar importância a participação pública na pesquisa científica, a ciência cidadã (*citizen science*).

A seguir são apresentados quatro exemplos inovadores do uso dessas novas tecnologias, a *Lista de Espécies da Flora do Brasil*, o *Open Courseware* do *Massachusetts Institute of Technology*, o *Biodiversity Heritage Library* e a experiência *Peer To Patent*.

#### *1.8.2.1. A Lista de Espécies da Flora do Brasil*

Até 2010, o único estudo completo sobre a ocorrência de Angiospermas no Brasil foi realizado por Carl Friedrich Philipp von Martius, August Wilhelm Eichler e Ignatz Urban, com a participação de 65 especialistas de vários países, que, no período de 1840 a 1906, editaram a obra *Flora Brasiliensis*. Na obra foram descritas 22.550 espécies, 2.298 gêneros e 210 famílias.

Em 1992, a Sociedade Botânica do Brasil indicou como estratégia a realização de floras por unidade da federação, em um formato que poderia se iniciar pela listagem das espécies. Tiveram início vários projetos de floras estaduais, inclusive a do estado de São Paulo em janeiro

de 1994. Tratava-se de um projeto bastante ambicioso que partia de um levantamento dos materiais depositados em herbários paulistas, obtendo uma lista de binômios e número de ocorrências por município. Com base neste levantamento, foram convidados taxonomistas do país e do exterior para trabalhar nas listas preliminares e foram identificadas lacunas de conhecimento geográfico para orientar novas coletas. A Fapesp também investiu na infraestrutura dos herbários para o acondicionamento do material coletado nas expedições (MAMEDE, 2003).

O projeto começou poucos anos depois do início da Internet no Brasil. A equipe do projeto optou por única e exclusivamente comunicar os resultados através da publicação de livros em papel. De acordo com a apresentação do Volume 5 da Flora Fanerogâmica do Estado de São Paulo<sup>48</sup> em 2007, a equipe do projeto previa a publicação dos 10 volumes restantes da série até 2017. Se esta previsão estiver correta, para descrever a flora do Estado de São Paulo os trabalhos durarão cerca de 23 anos.

Um dos programas da CDB é a Estratégia Global para a Conservação de Plantas (GSPC). Seu objetivo mais amplo é reduzir a perda contínua da diversidade de plantas para garantir o bem estar humano e um futuro sustentável ao planeta. Uma das metas estabelecidas para o ano de 2010 foi a elaboração de uma "lista funcional amplamente acessível das espécies conhecidas de plantas de cada país, como um passo para a elaboração de uma lista completa da flora mundial".

O Jardim Botânico do Rio de Janeiro (JBRJ) foi designado pelo MMA para coordenar a elaboração da Lista de Espécies da Flora do Brasil. O JBRJ convidou o CRIA para desenvolver o sistema e estruturou um Comitê Organizador que teve um papel importante na articulação da comunidade científica.

O sistema partiu de cerca de 40 listas existentes, cujos dados teriam que ser validados. A coordenação do JBRJ e os coordenadores dos grupos taxonômicos conseguiram em pouco tempo estruturar uma rede de 412 taxonomistas e alunos sendo 358 do Brasil, 20 dos Estados Unidos, 20 da Europa e 14 de outros países latino-americanos. Graças a esta rede social voluntária, à coordenação do JBRJ e ao desenvolvimento técnico do CRIA, a lista *on-line* foi aberta em abril de 2009 e concluída em dezembro de 2009. Um grupo de editoração trabalhou até maio de 2010, padronizando os nomes e verificando possíveis inconsistências que voltavam para a correção dos coordenadores. O Catálogo impresso foi lançado no 61º Congresso Nacional de Botânica em

---

<sup>48</sup> Veja na página <http://www.ibot.sp.gov.br/publicacoes/livros/livro2.php> (acessada no dia 04 de setembro de 2011)

setembro de 2010 em Manaus e apresentado na Conferência das Partes da Convenção sobre Diversidade Biológica realizada em Nagoya, Japão em outubro de 2010.

O desenho usado tanto para a entrada de dados como as diferentes combinações de buscas transformaram a lista em uma importante ferramenta de análise. Graças ao fato dos dados estarem *on-line*, foi possível desenvolver aplicativos para analisar diferentes aspectos como número de espécies por bioma e endemismo (FORZZA, BAUMGRATZ, *et al.*, 2012).

A Lista de 2010 tinha como objetivo imediato obter rapidamente os nomes válidos de plantas que ocorrem no Brasil. Nas versões 2011 e 2012 foram acrescentados mais campos além das imagens dos *vouchers*. Ao longo do tempo, a lista, se for mantida, provavelmente se tornará uma verdadeira Flora do Brasil.

O uso da tecnologia para viabilizar a participação dinâmica, *on-line*, de uma rede de especialistas, cada um responsável por uma parte pequena da lista, permitiu a produção da Lista em um espaço curtíssimo de tempo. Também graças à tecnologia adotada, a lista é dinâmica, com atualização contínua e com a possibilidade de incluir novos campos, de forma a atender as demandas dos provedores e usuários. A tecnologia torna possível incluir a colaboração de especialistas do exterior, principalmente para grupos onde o Brasil não tem competência.

O seu formato digital permite ainda a geração de diferentes produtos, além do catálogo, de estatísticas e relatórios com diagnósticos. Permite também o seu uso como ferramenta na avaliação da qualidade dos dados de herbários ou listas botânicas e na identificação de lacunas de conhecimento taxonômico.

#### 1.8.2.2. MIT OpenCourseWare

Na linha de material didático para ensino superior, uma experiência importante que merece destaque é o MIT OpenCourseWare (OCW)<sup>49</sup> – *Unlocking Knowledge, Empowering Minds*. Trata-se de um sistema *web* com praticamente todo o conteúdo dos cursos do Massachusetts Institute of Technology (MIT) disponível *on-line*, acessível a qualquer pessoa interessada. O sistema vem sendo desenvolvido desde 2002 e em 2007 já havia a publicação de mais de 1.800 cursos de 33 disciplinas acadêmicas. Em novembro de 2012, todos os 2.100 cursos oferecidos estavam *on-line*. O conteúdo *on-line* pode ser utilizado, copiado, distribuído, traduzido e modificado por qualquer pessoa interessada. A proposta do OCW é que as pessoas usem esse

---

<sup>49</sup> Para mais informações acessar <http://ocw.mit.edu/about> (acessado no dia 04 de novembro de 2011).

material não só como referência, mas para criar material novo que também deverá ser compartilhado. Assim, não se trata de uma simples biblioteca digital *on-line*, mas pode se tornar uma plataforma para os usuários colaborarem, compartilharem suas experiências, e ajudarem a acrescentar conteúdo ao longo do tempo (TAPSCOTT e WILLIAMS, 2010). O site tem em média um milhão de visitas por mês.

#### 1.8.2.3. *Biodiversity Heritage Library*

De interesse para estudos sobre biodiversidade, merece destaque o consórcio de bibliotecas de museus botânicos e de história natural denominado *Biodiversity Heritage Library*<sup>50</sup>. As obras em domínio público e de grande valor histórico e científico estão sendo digitalizadas e disponibilizadas *on-line* de forma livre e aberta a qualquer pessoa interessada. O consórcio hoje disponibiliza mais de 57 mil títulos, 108 mil volumes e cerca de 40 milhões de páginas. Os textos podem ser buscados, recuperados e lidos *on-line* ou baixados como arquivos em pdf.

#### 1.8.2.4. *Peer To Patent*

Essa iniciativa refere-se ao processo de exame de patentes. No Brasil, o pedido de patente é publicado na revista semanal da Propriedade Industrial após o término dos 18 (dezoito) meses de sigilo contados a partir do depósito ou da prioridade mais antiga. O exame do pedido de patente deverá ser requerido pelo depositante ou por qualquer interessado no prazo de 36 meses da data do depósito. O exame não será iniciado antes de decorridos 60 dias da publicação do pedido (INPI, 2008).

A iniciativa denominada *Peer To Patent*<sup>51</sup> teve início em 2007, fruto de uma colaboração entre a *New York Law School* e o escritório de patentes dos EUA (USPTO). A ideia do projeto foi desenvolver uma plataforma para colaboração pública via Internet de maneira estruturada e produtiva para ajudar o escritório de patente na fase de exame (NOVECK, 2009). O objetivo é garantir que as patentes aprovadas realmente têm mérito, ou seja, atendem aos critérios de novidade, atividade inventiva e aplicabilidade industrial e também garantir que pedidos que não atendam a esses critérios não sejam aprovados. No sistema atual, o examinador da patente é o único responsável por buscar informações sobre o estado da arte e estudar o pedido de patente.

---

<sup>50</sup> As informações do consórcio estão disponíveis no endereço <http://www.biodiversitylibrary.org> (acessado no dia 07 de outubro de 2012)

<sup>51</sup> O site **PEERTOPATENT** pode ser acessado na página <http://www.peertopatent.org>

Para tanto necessitaria estar na fronteira do conhecimento para poder garantir que os três critérios básicos sejam atendidos. *Peer To Patent* propõe buscar através da Internet a colaboração voluntária daqueles que de fato estão na fronteira do conhecimento.

O *Peer To Patent* usou um *software* de rede social para facilitar a discussão entre especialistas que, de forma voluntária, contribuíram com referências sobre o estado da arte, participaram de discussões e classificaram o material enviado por outros colaboradores. O resultado dessa discussão e a relação das dez publicações mais relevantes (selecionadas pelos participantes) foram enviados para o examinador que, dessa forma, passou a ter em mãos o que havia de mais recente sobre o assunto da patente em estudo. O examinador não participou da discussão, apenas recebeu os resultados obtidos. 220 pedidos de patentes das áreas de *software* e administração de empresas foram submetidos a esse processo com bons resultados, tanto na qualidade como também no tempo de análise. Um novo projeto foi iniciado em outubro de 2010 e concluído em setembro de 2011 e incluiu patentes das áreas de telecomunicações, traduções, biotecnologia, bioinformática e biofarmacêuticos. Outros escritórios de patente que estão desenvolvendo programas-piloto seguindo esse modelo são os da Austrália, Japão, Coreia do Sul e Reino Unido.

Portanto, mesmo quando existem restrições ao acesso a dados e informações, como é o caso dos pedidos de patente, a Internet e seus vários aplicativos de colaboração e integração podem ser usados para melhorar a qualidade da análise.

Esses são apenas alguns exemplos do uso das TICs para a disseminação de dados, informação e conhecimento no século XXI. As e-infraestruturas sobre biodiversidade serão discutidas no Capítulo 3. As tecnologias da informação e comunicação são os fatores mais recentes de transformação da ciência. Da mesma forma que as revistas científicas surgiram no século XVII como um novo canal formal de comunicação científica, bancos de dados e sistemas de informação *on-line* no século XXI são infraestruturas essenciais para o avanço científico especialmente quando o assunto é meio ambiente e biodiversidade.

Capítulo 2, a seguir, apresenta a definição e a relação entre dados, informações e conhecimento. Apresenta também algumas características próprias de dados sobre biodiversidade e discute o conceito de *Open Science* e a sua demanda sobre os sistemas de informação *on-line*.

## **CAPÍTULO 2. COMUNICAÇÃO CIENTÍFICA E MEIO AMBIENTE**

A comunicação científica compreende o conjunto de todas as atividades que englobam a produção, disseminação e uso da informação, desde o início do processo de criação científica, no qual as idéias da pesquisa são geradas, até o momento da aceitação dos resultados como parte do corpo de conhecimento científico. Inclui a formação e o treinamento de recursos humanos, os encontros com trocas de experiências em reuniões científicas, a consultoria, a transferência de pessoas e intercâmbio de especialistas, a publicação de livros e artigos científicos e a disseminação de dados e informações (GARVEY e GRIFITH, 1979; LEITE e COSTA, 2006).

Busca-se nesse capítulo discutir o conceito de dados, informação e conhecimento, como suporte para as análises sobre as atividades de CT&I e o papel de curadores e provedores de dados e informações. Mais especificamente para dados sobre biodiversidade, são também definidos dados insubstituíveis ou estáveis, sigilosos ou abertos e dados sensíveis. São também apresentados o conceito de ciência aberta no mundo contemporâneo, as mudanças na produção de conhecimento e sua influência nos sistemas de comunicação científica.

Como último item, o capítulo analisa as especificidades dos dados e informações sobre biodiversidade e a sua importância frente ao desafio do desenvolvimento sustentável. Com o avanço das TICs, Lastres e Albagli (1999) visualizam a emergência de um novo regime ou ordem informacional e a necessidade de regulamentação de novas práticas e relações em torno de atividades como a ampliação e reformulação das garantias de direitos de propriedade intelectual e a necessidade de novas regras e normas que ordenem os processos de geração, acesso, fluxo, disseminação e uso de informações e conhecimento. Mas, ao destacar outras dimensões que vêm sendo marginalmente consideradas ou até mesmo excluídas, como a dimensão social, política, ambiental e ético-valorativa, oferecem uma visão alternativa que contempla, entre outros, a subordinação dos mecanismos de apropriação privada de informações, saberes e conhecimentos à universalização do acesso daqueles de interesse público e social.

Assim, a definição política sobre a estratégia para a disseminação dos dados, como predominantemente aberta ou fechada, depende não só da dimensão econômica, mas cada vez mais das dimensões ambiental e social.

Focando a ciência sobre biodiversidade<sup>52</sup>, neste capítulo procura-se mostrar que as políticas públicas devem promover, em geral, o intercâmbio livre e aberto dos dados. Que a regra deve ser o compartilhamento livre e aberto aos dados e informações sobre a ocorrência de espécies e que dados e informações sigilosas ou sensíveis são exceções à regra, merecendo tratamento específico.

## 2.1. DADOS, INFORMAÇÃO E CONHECIMENTO

A discussão nos campos da Ciência da Informação e da Gestão do Conhecimento sobre a definição e a relação entre dados, informação, conhecimento e, em menor grau, sabedoria (DIKW – *Data, Information, Knowledge & Wisdom*), data de algumas décadas, sendo muito divulgada a pirâmide do conhecimento que procura contextualizar esses elementos mostrando-os como uma estrutura hierárquica (ACKOFF, 1989).



Figura 2.1. Pirâmide do Conhecimento (ACKOFF, 1989)

Rowley (2007) realizou uma ampla análise da relação hierárquica e das definições apresentadas por diferentes autores, tanto da área da ciência da informação como da gestão do conhecimento, apresentando uma definição própria daquilo que é considerado consensual, como apresentado a seguir.

**Dados** são definidos pelo que eles não apresentam – *aos dados falta significado ou valor, não são organizados ou processados*<sup>53</sup>.

**Informação** é geralmente definida em relação a dados, sendo percebida como dados organizados ou estruturados. Essa organização dá aos dados relevância para um propósito

<sup>52</sup> Entendemos como ciência sobre biodiversidade a taxonomia, a ecologia e a biologia da conservação.

<sup>53</sup> “Interestingly, these definitions are largely in terms of what data lacks; data lacks meaning or value, is unorganized and unprocessed” (ROWLEY, 2007, 171).

específico ou contexto. Informação, portanto, é definida como *dados que têm significado, valor, utilidade e relevância*.

**Conhecimento** é tipicamente definido em relação à informação. Pode ser considerado como uma *mescla de informação, compreensão, capacitação, experiência, aptidão e valores*. Textos da área da Gestão do Conhecimento também discutem a diferença entre conhecimento tácito e explícito, sendo o *conhecimento tácito* aquele que está na mente das pessoas e o *conhecimento explícito* o conhecimento codificado e disponível em livros, bancos de dados e outros formatos registrados.

**Sabedoria** é um conceito ainda mais complexo de definir. Uma definição filosófica básica é que *sabedoria é fazer o melhor uso do conhecimento*.

Cada elemento é definido em relação ao outro, então existem elementos que claramente são dados (p.ex. datas, coordenadas geográficas) e outros que mesclam dados e informações, como os registros sobre a ocorrência de espécies, que incluem informações taxonômicas. A atribuição de um nome à espécie observada ou coletada requer um conhecimento especializado. Portanto, pode-se argumentar que o taxonomista responsável pela identificação da amostra, ao lhe atribuir um nome científico, está compartilhando o seu conhecimento (MCDADE, MADDISON, *et al.*, 2011), ou seja, o nome da espécie é uma informação.

Para bancos de dados esses dois termos são sinônimos (ITL EDUCATION SOLUTIONS LIMITED, 2006). A *National Science Foundation*, quando explicitou que a partir de janeiro de 2011 toda proposta de projeto deveria apresentar um plano de gestão dos dados (*Data Management Plan*), indicou que o que constitui dados seria determinado pela comunidade (*community of interest*) por meio do processo de avaliação pelos pares. Mas indicou que poderia incluir – mas não estar limitado a – dados, publicações, amostras, coleções físicas, *software* e modelos. Portanto a tese, quando se refere a dados sobre biodiversidade no contexto de bancos de dados, sistemas de informação ou infraestruturas eletrônicas de dados, utiliza esse conceito mais amplo.

Os canais de comunicação científica podem ser classificados como *informal* e *formal*. Antes dos avanços das TICs a distinção era relativamente clara. A comunicação oral e cartas eram classificadas como *canais informais de comunicação*, por serem de caráter pessoal, de acesso limitado e de pouco tempo de vida. Os livros e revistas científicas são classificados como canais formais de comunicação, cujas características são o amplo acesso e vida longa. Nos dias

de hoje, uma mensagem de correio eletrônico estará usando a mesma mídia que uma revista eletrônica e pode ter uma distribuição em massa. Pode ainda ser enviada a uma lista de discussão na qual é arquivada, sendo mantida durante anos em meio eletrônico, podendo ser pesquisada e recuperada por qualquer pessoa interessada. Assim, no caso da mensagem e do artigo, ambos podem ser enviados a um público de qualquer tamanho, ambos são divulgados pelos mesmos canais e podem ser acessados por leitores por meio das mesmas telas de computador. A distinção hoje não é mais clara (MEADOWS, 1999).

Pode-se ainda argumentar que o periódico passa por um processo de validação, de revisão por pares, enquanto uma mensagem ou uma contribuição *on-line* não. No entanto, com os atuais sistemas interativos, existem inúmeros exemplos de contribuições *on-line* discutidas e validadas por especialistas, o que torna essa distinção menos clara.

Dados podem ser classificados como *insubstituíveis* ou *estáveis*; *sigilosos* ou *abertos*; e ainda, *sensíveis*. Essas classificações são especialmente importantes para dados sobre a ocorrência de espécies.

Dados *insubstituíveis* (ou efêmeros) são aqueles que não podem ser novamente coletados, enquanto outros (os estáveis) podem ser regenerados. Para estratégias de preservação de dados, esse conceito é fundamental. Basicamente tem-se que dados efêmeros precisam ser preservados, enquanto que a preservação de dados estáveis provavelmente dependerá do custo comparativo entre a preservação ou uma nova medição (GRAY, SZALAY, *et al.*, 2002). Com a extinção de espécies e perdas de habitat, muitos dados sobre biodiversidade, antes considerados estáveis, hoje são efêmeros (GODDARD, WILSON, *et al.*, 2011).

Dado ou informação *sensível* é aquele(a) que, se publicado(a), resultaria em um efeito adverso sobre o táxon ou ao indivíduo vivo. Vários fatores precisam ser considerados na determinação da sensibilidade de um dado, incluindo o tipo e nível de ameaça, a vulnerabilidade do táxon, o tipo de informação, e se essa informação já está à disposição do público (CHAPMAN e GRAFTON, 2008). Os exemplos mais comuns de dados sensíveis são as coordenadas geográficas precisas de espécies raras ou ameaçadas de extinção com alto valor comercial. A divulgação desse dado pode aumentar o risco de ameaça à espécie. No entanto, se um dado for considerado sensível, é importante que esta condição seja sempre reavaliada.

Em relação ao *dado aberto* (*open data*), a definição usual, seguindo a linha do *software* aberto, é *aquela que pode ser livremente usado, reusado e redistribuído por qualquer pessoa* –

*sujeito apenas, e no máximo, às exigências de atribuição e compartilhamento usando a mesma licença (sharealike) (OKF, 2010).*

Esses atributos definem a estratégia e a adoção de políticas para preservar e disseminar dados e informações de maneira aberta ou restritiva. Não se pode afirmar que todo dado deve ser de acesso livre e aberto, nem que deve ser de acesso restrito. É necessário considerar seu contexto. Além do conceito de dados abertos, importante para esta tese é o conceito de ciência aberta (*open science*), apresentado a seguir.

## 2.2. CIÊNCIA ABERTA (*OPEN SCIENCE*)

Na Grécia clássica, a ciência desenvolveu-se no paradigma do debate público, estimulando a competição ao invés da colaboração científica. Na Idade Média, com forte influência religiosa, a prática era o sigilo em busca dos segredos da Natureza. Além das considerações morais e filosóficas contrárias à divulgação aberta dos segredos divinos, as tecnologias disponíveis para a comunicação científica nesse período eram muito primitivas e custosas, o que não favorecia uma difusão científica mais ampla.

A impressão por tipos móveis, introduzida por Gutenberg por volta de 1439, é um dos fatores que contribuiu para promover e aparelhar a divulgação científica. No primeiro momento, apenas substituiu os manuscritos na cópia de obras de referência, ou seja, não alterou os padrões da comunicação da época. Mas a distribuição muito mais ampla das cópias promoveu a alfabetização e incentivou a autoria e as recompensas derivadas do reconhecimento do autor.

No período da Renascença, tem início a experimentação científica, a base da ciência moderna, por Francis Bacon (1561-1626). A ideia de que a humanidade se beneficiaria de um sistema organizado e coletivo em busca do conhecimento público e as convenções e normas que davam suporte à prática da ciência aberta parecem ser um aspecto organizacional vital distinto da Revolução Científica (DAVID, 2008). Graças às vantagens da divulgação ampla e aberta, surgiram várias iniciativas que promoviam a colaboração científica como as sociedades científicas, revistas especializadas e as grandes exposições.

Na sociedade contemporânea, a forma com que o conhecimento é produzido está mudando e afeta não somente o tipo de conhecimento que é produzido, mas como é produzido, em que contexto, os seus sistemas de recompensa, o controle de qualidade e como ele é comunicado (GIBBONS, LIMOGES, *et al.*, 1994). Esses autores caracterizam dois modos de produção de conhecimento. O *modo 1* é caracterizado por uma produção de conhecimento

disciplinar, desenvolvida na ausência de um objetivo prático, em que os resultados são avaliados por pares e comunicados por meio dos canais institucionais. No *modo 2*, o conhecimento é produzido no contexto das aplicações, na resolução dos problemas e a difusão dos resultados é realizada durante o processo. Em ambos os modos, a comunicação é crucial, mas o modo 1 foca a divulgação dos resultados (principalmente em revistas científicas especializadas) enquanto que no modo 2 a comunicação é crucial durante todo o processo, muitas vezes ultrapassando barreiras institucionais. Além disso, por essa visão, caem por terra o modelo linear da evolução da ciência, tecnologia e inovação e as fronteiras entre geração, difusão e comunicação do conhecimento. Fazer e difundir ciência e conhecimento não diz mais respeito apenas a cientistas e sim a toda uma comunidade de especialistas que necessariamente faz parte desse processo (HOBSBAWM, 2008).

Junta-se a esse diferente entendimento da dinâmica da CT&I, o processo de massificação da pesquisa e educação nos países industrializados no pós-Guerra, o qual estabeleceu as bases para uma maior difusão do conhecimento técnico e científico e das aptidões científicas em toda a sociedade. Antes da massificação da educação, os cientistas eram os detentores do conhecimento e a sociedade em geral era considerada leiga, e assim a troca de informação basicamente se restringia à própria comunidade científica e se fazia presente nos processos de educação. Uma cidadania mais bem educada coloca novas demandas à ciência, que passa a prestar contas não só financeiras, mas também sociais. A ciência passa a ser objeto de discussão pública (p.ex. questões relativas ao meio ambiente, energia nuclear, ética), sua linguagem se torna vernacular e existe uma difusão muito maior de dados e informações científicas para a sociedade (NOWOTNY, SCOTT e GIBBONS, 2001). Os avanços das TICs tornaram possível o envolvimento de mais atores na gênese do conhecimento, atores de diferentes disciplinas, especialidades, instituições, localidades, países, culturas e realidades sociais. A produção do conhecimento está associada ao problema, projeto ou programa e não mais necessariamente a uma instituição.

Uma definição informal postada por Michael Nielsen na lista de discussão do projeto *OpenScience*<sup>54</sup> indica que *Ciência aberta é a ideia que o conhecimento científico de todos os tipos deve ser compartilhado de maneira aberta o mais cedo possível durante o processo de*

---

<sup>54</sup> Para informações adicionais sobre o OpenScience Project, acessar <http://www.openscience.org>

*descoberta*<sup>55</sup>. O projeto indica quatro objetivos fundamentais no compartilhamento de conhecimento: (i) transparência na metodologia experimental, na observação e na coleta de dados; (ii) disponibilidade pública e possibilidade de reutilização de dados científicos; (iii) acessibilidade pública e comunicação científica transparente; e (iv) uso de ferramentas baseadas em *Web* para facilitar a colaboração científica.

Uma das principais bases de sustentação da ciência aberta é o uso e o acesso à tecnologia da informação. Um segundo ponto é a aproximação da ciência e da tecnologia, pela qual elementos de apropriabilidade têm permeado a ciência e elementos de divulgação e abertura têm se ampliado no segmento da tecnologia e inovação. O terceiro ponto refere-se às estruturas de governança, bem mais complexas que as usualmente empregadas em atividades de pesquisa (SALLES-FILHO, BIN e FERRO, 2008).

Mas é importante novamente frisar que, apesar dos enormes avanços que hoje permitem o compartilhamento aberto de dados, informação, conhecimento e ferramentas, continuarão existindo dados proprietários, publicações de acesso restrito e redes de acesso limitado. Também não é uma questão de ser aberto ou fechado, podem existir diferentes níveis de acessibilidade. Quanto estará aberto ou quanto será de acesso restrito dependerá de uma série de fatores que incluem a política, a coerência organizacional dos atores envolvidos, soluções técnicas abertas e sistemas abertos como ideal de pesquisa (SCHROEDER, 2007).

Mas para a promoção da ciência aberta, outro ponto importante são os mecanismos de reconhecimento e valorização do trabalho dos profissionais responsáveis pela geração, qualidade, manutenção e disseminação de dados e informações (THE ROYAL SOCIETY SCIENCE POLICY CENTER, 2012). No entanto, a atividade de curadoria de dados e bancos de dados em rede, apesar de sua relevância, por muitas agências não é reconhecida como parte do domínio “atividade de pesquisa” (SCHOFIELD, EPPIG, et al., 2010). Esse fato representa um possível gargalo para o avanço das e-infraestruturas.

Um instrumento importante criado em 1963 para conceituar P&D e os diferentes recursos alocados (financeiros, materiais e humanos) no ensino superior, governo, indústria e organizações sem fins lucrativos, possibilitando sua mensuração, é o Manual Frascati da OCDE (OECD, 2002). A classificação proposta e revista ao longo dos anos, permite a produção de estatísticas

---

<sup>55</sup> “Open science is the idea that scientific knowledge of all kinds should be openly shared as early as is practical in the discovery process.”, Michael Nielson

comparáveis de dispêndios em P&D, atingindo um nível mínimo de compatibilidade dos dados no nível internacional. As estatísticas produzidas são utilizadas como indicador para discussões em políticas de ciência e tecnologia e representa o padrão global para a coleta de estatísticas em P&D. Um dos principais indicadores é o GERD (*Gross domestic Expenditure on Research and Development*) e a proporção GERD/PIB como alvo de política.

De acordo com o Manual Frascati, as atividades de Ciência e Tecnologia (STA - *Scientific and Technological Activities*) incluem Pesquisa e Desenvolvimento, Educação e Treinamento Científico e Tecnológico (STET – *Scientific and Technical Education and Training*) e Serviços Científicos e Tecnológicos (STS – *Scientific and Technological Services*). São considerados STS, entre muitas outras, as atividades de C&T de bibliotecas e museus (OECD, 2002).

As coleções biológicas e as infraestruturas eletrônicas de dados e informações sobre biodiversidade de acesso público, portanto, são classificadas como serviços técnico-científicos e, como tal, precisam ser reconhecidas. Ou seja, são atividades de C&T. A compreensão desse conceito é importante para valorizar o papel estratégico das coleções, das e-infraestruturas e dos profissionais que nelas atuam.

### 2.3. DADOS E INFORMAÇÕES SOBRE BIODIVERSIDADE

A Constituição de 1988 e as leis sobre a política nacional do meio ambiente indicam ser dever do Estado e direito do cidadão o acesso à informação. A CDB também prevê a obrigatoriedade, pelos países membros, do intercâmbio de informações pertinentes à conservação e à utilização sustentável da diversidade biológica. Assim, podemos afirmar que informações não sensíveis sobre biodiversidade, em princípio devem ser de acesso livre e aberto. É um dever do Estado e um direito do cidadão.

O acesso à informação ambiental como direito constitucional e o dever do Estado em provê-la, não só internamente mas em sistemas globais, demonstram a importância estratégica de dados<sup>56</sup> sobre biodiversidade.

No ano 2000, por iniciativa das Nações Unidas, foi assinada a Declaração do Milênio por 191 Estados Nacionais, considerado o mais importante compromisso internacional a favor do desenvolvimento e da eliminação da pobreza e da fome no mundo<sup>57</sup>. Foram estabelecidos oito

---

<sup>56</sup> Dados aqui são usados em seu conceito mais amplo que inclui dados, informações, imagens, sons etc..

<sup>57</sup> Para maiores informações, acessar <http://www.itamaraty.gov.br/temas/temas-multilaterais/direitos-humanos-e-temas-sociais/metas-do-milenio/objetivos-de-desenvolvimento-do-milenio>

Objetivos de Desenvolvimento do Milênio (*Millennium Development Goals* – MDGs) a serem alcançados até 2015. Para a Rio+20 foi estabelecido um painel de alto nível sobre sustentabilidade global que recomendou que, a partir de 2015, formuladores de políticas adotem metas universais de desenvolvimento sustentável (*Sustainable Development Goals*) para os próximos 15 anos, com marcos intermediários. Fundamental para o alcance dessas metas são dados precisos, adequados e disponíveis para gestores, tomadores de decisão e para o público (SACHS, 2012).

O desafio do desenvolvimento sustentável envolve cientistas de diferentes áreas de conhecimento, tecnólogos, formuladores de políticas e comunidades de todas as escala – de local a global. Um dos principais componentes é a capacidade de mobilizar e usar a C&T (CASH, CLARK, *et al.*, 2003). Assim, as infraestruturas de dados de acesso público e aberto desempenham um papel cada vez mais importante e complexo no desafio de vincular dados, informação e conhecimento a tomadas de decisão.

Além dos dados, é necessário apoiar o desenvolvimento de aplicativos capazes de realizar processos de complexidade cada vez maior, mas que, por atenderem a um público sempre mais diverso, necessitam ser intuitivos tanto quanto possível. Assim, a disseminação e o compartilhamento aberto de dados, informações e conhecimento científico, de uma maneira que podem ser apropriados pelos diferentes segmentos organizados da sociedade, além do investimento na formação de recursos humanos, dependem muito dos avanços das tecnologias de informação e comunicação, tema objeto da análise apresentada nos capítulos 3 e 4.

A partir das discussões sobre dados, informações, conhecimento e ciência aberta e da necessidade de integrar e disponibilizar dados e informações de diferentes áreas de conhecimento que devem ser apropriados por diferentes atores para enfrentar o desafio do desenvolvimento sustentável, procurou-se mostrar a importância de uma política aberta aos dados e informações resultantes da ciência sobre biodiversidade. Uma política aberta a dados sobre biodiversidade, em seu sentido mais amplo, passa a considerar dados sigilosos ou sensíveis como exceção, a regra é a sua disseminação livre e aberta a todos os interessados.

No próximo capítulo são apresentadas diferentes experiências de desenvolvimento e manutenção de e-infraestruturas de dados sobre biodiversidade de acesso livre e aberto. Essas e-infraestruturas têm suas ações focadas em seus públicos-alvo, mas podem ter a sua ação ampliada por serem de acesso livre e aberto.



### **CAPÍTULO 3. E-INFRAESTRUTURAS SOBRE BIODIVERSIDADE**

Importantes desafios para a próxima geração de infraestruturas globais de dados científicos incluem dar suporte efetivo e eficiente à ciência intensiva em dados e à ciência multi e interdisciplinar (GRDI2020, 2010). São ambientes gerenciados de dados científicos digitais e em rede que consistem em serviços e ferramentas que dão suporte: (i) ao ciclo completo de dados científicos: captura, coleção, curadoria, documentação, análise, visualização, preservação e publicação; (ii) ao movimento de dados científicos por meio das diferentes disciplinas; (iii) à criação de *links* abertos conectando bases de dados de disciplinas diversas; e, (iv) à interoperabilidade de dados científicos e da literatura.

A definição da política de dados da ciência sobre biodiversidade, em seu sentido mais amplo, como uma política de acesso livre e aberto deve nortear a estratégia para o desenvolvimento e manutenção das e-infraestruturas.

Nesse capítulo são analisadas três e-infraestruturas, todas de acesso livre e aberto, que estão disponibilizando dados sobre espécies e espécimes: o *Global Biodiversity Information Facility* (GBIF), a *Comisión Nacional para el Conocimiento y Uso de la Biodiversidad* (CONABIO) do México e a *rede speciesLink do Brasil*.

Essa seleção se justifica, pois das iniciativas internacionais, o GBIF é o sistema global e a CONABIO, com 20 anos de história, é um dos principais sistemas nacionais do mundo e foi desenvolvida por um país latino americano. A *rede speciesLink* é analisada por ser o sistema brasileiro que mais se aproxima da missão do GBIF, sendo internacionalmente reconhecida pelos grandes avanços e inovações no desenvolvimento de ferramentas e aplicativos.

Para as três redes são analisados diferentes elementos da e-infraestrutura, como a motivação que resultou no desenvolvimento dos sistemas, o provedor de dados, o usuário-alvo, a política de dados, a tecnologia adotada e os resultados obtidos. Um último elemento – governança – é também analisado, sendo talvez o elemento mais complexo. Salles *et al.* (2008) referem-se às novas estruturas de governança como um dos elementos que distinguem as abordagens abertas de criação e produção do conhecimento, como sendo capazes de congregarem coordenação, eficiência, eficácia e efetividade com liberdade e criação e questionamento. Governança para as e-infraestruturas também é um elemento central.

Para cada rede é apresentada uma breve análise dos pontos fortes e fracos, buscando elementos para avaliar a adequação das políticas de fomento para o desenvolvimento e manutenção de e-infraestruturas sobre biodiversidade.

### 3.1. GLOBAL BIODIVERSITY INFORMATION FACILITY – GBIF<sup>58</sup>

Conforme apresentado no Capítulo 1, o GBIF é resultante do Fórum de Megaciência (hoje denominado Fórum da Ciência Global) da OCDE que considerava que *o esforço necessário para integrar recursos de informática e seus usuários em uma unidade sinérgica e interoperável é que torna a informática biológica em um esforço de megaciência* (OECD MEGASCIENCE FORUM, 1999)<sup>59</sup>. O conceito de *megaciência* foi cunhado pela OCDE sendo aparentemente escolhido para enfatizar não somente a questão de escala ou apoio do Estado, como evidenciado nas formas tradicionais de “big science”, mas o caráter complexo de parcerias estratégicas transinstitucionais e transfronteira (ELZINGA, 2012).

A rede GBIF foi estabelecida em 2001 por governos para estimular o acesso livre e aberto a dados sobre biodiversidade através da Internet. Tem como visão *um mundo onde informações sobre biodiversidade estão disponíveis de forma livre e universal para ciência, sociedade e para um futuro sustentável*. E tem como missão *ser o recurso mais importante de informação global para a biodiversidade e gerar soluções inteligentes para questões ambientais e de bem-estar humano* (GBIF, 2011).

Trata-se de uma iniciativa multilateral estabelecida por acordos intergovernamentais, baseado em um memorando não vinculativo (*Memorandum of Understanding – MoU*). Enquanto os dois primeiros MoU tiveram cinco anos de duração (2001-2006 e 2007-2011), a partir de 2012 o MoU não tem prazo (*open-ended*). Participantes do GBIF são países (governos), economias, organizações intergovernamentais ou internacionais, ou organizações com uma missão internacional / global / regional, ou ainda, instituições designadas por essas organizações para assinar o MoU. Hoje, participam da rede 58 países e 46 organizações internacionais<sup>60</sup>. O Brasil aderiu ao GBIF em outubro de 2012.

**Governança:** O órgão máximo do GBIF é o Conselho de Administração (*Governing Board, GB*), hoje composto por 38 países membros votantes, 20 países associados, sem direito a

---

<sup>58</sup> O endereço do GBIF é <http://www.gbif.org> (acessado em 02/09/2012)

<sup>59</sup> *It is the need to link these informatics resources (and the people who use them) into a synergistic, interoperable whole that makes biological informatics a megascience endeavor.*

<sup>60</sup> Informação disponível no site <http://www.gbif.org> (acessado em 31/10/2012)

voto, e 46 instituições associadas. Subordinado ao GB está o Comitê Executivo, responsável por acompanhar as ações do Secretariado e dos demais participantes na implementação das decisões do GB. O Secretariado é a instância executiva do GBIF, juntamente com os nós dos países participantes. Participam também da governança os comitês permanentes que atualmente são quatro: científico, orçamento, regras e gestores dos nós GBIF. A estrutura conta também com grupos de trabalho e outras comissões estabelecidas pelo GB ou pelo Comitê Executivo. Quanto à sustentabilidade financeira, cada participante é responsável pelo financiamento de sua participação na rede, que pode incluir os custos para o estabelecimento do nó nacional, a estruturação de um programa de digitação, a participação em reuniões etc.. O Secretariado, a estrutura de comitês e o plano de trabalho são financiados com recursos pagos pelos países membros votantes. O custo dessa estrutura em 2010 foi de 3,7 milhões de euros. O relatório apresenta a distribuição de despesas com 42% para a área de informática, 31% para gestão, 23% para articulação e 4% para governança.

**Provedor de dados.** Para que uma instituição possa publicar seus dados na rede, é necessário ter o aval de um dos participantes da rede, um país ou uma organização internacional. Portanto, um provedor de dados somente integrará sua base informacional à rede GBIF através de um nó ou com o seu aval.

**Usuário-Alvo.** Fruto de um projeto do fórum de megaciência da OCDE, o principal usuário-alvo é a comunidade científica. Mas nos últimos anos tem havido um grande empenho em buscar novas parcerias e prover dados para as convenções (CDB, CITES etc.) e para tomadores de decisão.

**Política de Dados.** O GBIF tem como missão tornar dados sobre biodiversidade disponíveis na Internet de maneira livre e aberta<sup>61</sup>. Procura estruturar um ambiente positivo para o compartilhamento e uso dos dados para o benefício da ciência e da sociedade, e dessa maneira contribuir para um futuro sustentável. O MoU estabelecido com os provedores de dados<sup>62</sup> possui um parágrafo específico sobre propriedade intelectual onde indica claramente que as leis e acordos estabelecidos serão respeitados. O GBIF não detém qualquer direito sobre os dados compartilhados na rede, mas pode proteger as ferramentas que desenvolver. O MoU indica,

---

<sup>61</sup> Mais informações sobre a política de dados do GBIF estão disponíveis no endereço <http://www.gbif.org/communications/resources/gbif-brochures/data-policy>

<sup>62</sup> O novo Memorandum of Understanding firmado a partir de 2012 está disponível no endereço ([http://imgbif.gbif.org/CMS\\_NEW/get\\_file.php?FILE=b8160fdd6f5bcd938453a0e1cb9050](http://imgbif.gbif.org/CMS_NEW/get_file.php?FILE=b8160fdd6f5bcd938453a0e1cb9050))

todavia, que o GBIF deverá promover a transferência dos aplicativos desenvolvidos a instituições de pesquisa, principalmente às de países em desenvolvimento. Indica também que a autoria dos dados será explicitamente atribuída a cada provedor. Cabe a cada provedor determinar que dado será de acesso público e explicitar as restrições (quando houver) em relação ao uso dos dados. O GBIF dispõe de acordos tanto para o compartilhamento de dados<sup>63</sup> como também para o seu uso<sup>64</sup>. Ao realizar uma busca em seu sistema *on-line*, a única menção sobre propriedade intelectual é uma mensagem de rodapé em que se afirma que os provedores de dados detêm todos os direitos sobre os seus dados. Não há uma indicação clara sobre quais as condições de utilização dos dados determinadas por cada provedor.

**Tecnologia.** O GBIF foi estruturado como uma rede distribuída de provedores (nós estabelecidos nos países e instituições membros) e um servidor central mantido pelo Secretariado, responsável pela indexação dos dados. Além dos protocolos de comunicação inicialmente implementados - Biocase (coleções principalmente europeias que utilizam o modelo de dados ABCD) e DiGIR (coleções que utilizam o modelo de dados DarwinCore) - e a evolução desses protocolos - o TAPIR (que aceita os dois modelos de dados ABCD e DarwinCore) -, o Secretariado desenvolveu uma plataforma - o *Integrated Publishing Toolkit* (IPT) -, disponível para o envio de dados de ocorrência de espécies, *checklists* e metadados.

**Resultados.** A rede GBIF disponibiliza cerca de 390 milhões de registros *on-line*, integrando os dados de mais de 10 mil bancos de dados de 422 editores<sup>65</sup>. Numa avaliação mais criteriosa sobre os dados disponibilizados, usando como documento de análise o relatório anual de atividades de 2010 (GBIF, 2011) quando a rede disponibilizava 267 milhões de registros, tem-se que cerca de 72% eram dados de observação, enquanto apenas cerca de 28% eram dados de espécimes, ou seja, dados que têm uma amostra física (ou *voucher*) associada. Dados associados a um *voucher* são considerados de maior valor, uma vez que podem ser reavaliados.

A grande quantidade de dados de observação é predominantemente de aves. Daí também a maior proporção de dados de animais, que representam mais de 70% do total, enquanto plantas representam pouco mais de 26%. Um outro parâmetro importante é a origem do dado. Na rede

---

<sup>63</sup> GBIF Data Sharing Agreement disponível no endereço <http://www.gbif.org/participation/data-publishers/gbif-sharing-agreement/gbif-data-sharing-agreement>

<sup>64</sup> GBIF Data Use Agreement disponível no endereço <http://www.gbif.org/participation/data-publishers/gbif-sharing-agreement/gbif-data-use-agreement>

<sup>65</sup> O GBIF denomina as instituições que disponibilizam seus dados não de provedores, mas editores. Dados do site <http://www.gbif.org> acessado no dia 01/09/2012.

GBIF 89% dos dados são provenientes de instituições da América do Norte e Europa e 75% dos dados foram observados ou coletados no hemisfério norte. Tem-se, portanto, que a maior parte dos dados é de observação de aves no hemisfério norte.

Quarenta e quatro países e dezesseis iniciativas internacionais estão fornecendo dados à rede, sendo que esses países são responsáveis por cerca de 95% da totalidade dos dados disponibilizados na rede. A Figura 3.1 mostra a contribuição dos 15 países que fornecem mais dados a rede GBIF.

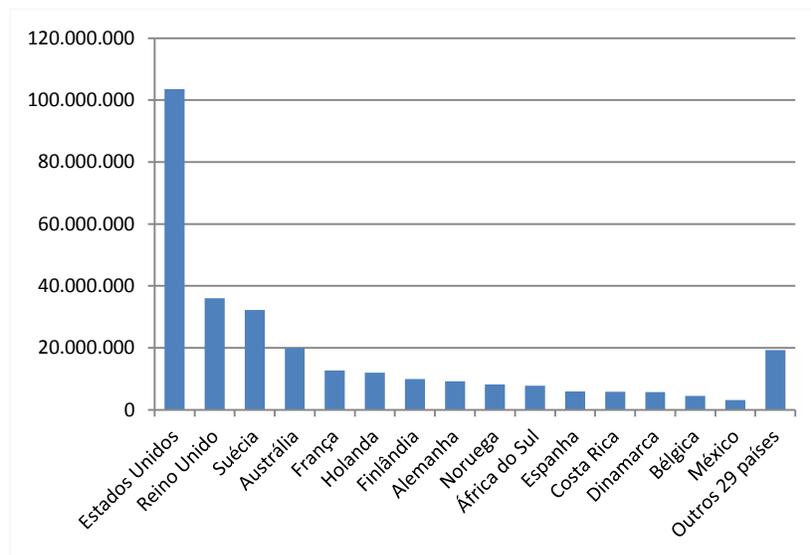


Figura 3.1. Contribuição de dados dos países participantes à rede GBIF.  
Fonte: GBIF (2011)

Os Estados Unidos, que atualmente fornecem mais de 100 milhões de registros por meio de 94 provedores, não enviam os dados através de um nó GBIF. Cada provedor tem seus dados indexados diretamente pela rede GBIF. Apenas 20 países membros têm um nó GBIF com atividade técnica e cerca de 70% dos registros desses países passam “fisicamente” por esses nós. A existência de nós nacionais bem estruturados representa uma excelente oportunidade para organizar e analisar os dados e fomentar o desenvolvimento local. Dos cinco países latino-americanos que estão fornecendo dados à rede (México, Costa Rica, Colômbia, Argentina, Peru), todos têm um nó técnico implementado e os nós são responsáveis pelo envio de cerca de 85% dos dados. No final de novembro de 2012, o Brasil assinou o memorando de entendimento e em breve deverá prover dados através do Sistema de Informação sobre a Biodiversidade Brasileira (SiBBr) e de sua parceria com a rede *speciesLink*, apresentada a seguir neste capítulo.

Com relação aos dados de coleções biológicas, ao final de 2010 a rede GBIF disponibilizava 58 milhões de registros de espécimes. Apesar dos inegáveis avanços, se considerarmos que o número estimado de espécimes existentes em coleções biológicas no mundo é de 1,2 a 2,1 bilhões de registros (ARÍÑO, 2010), tem-se que em 10 anos de atividade ainda não foi alcançado 5% do total estimado. É urgente um maior investimento na digitalização de dados de coleções biológicas e no engajamento de países na integração de seus sistemas de informação à rede GBIF, inclusive o Brasil.

Diferentemente da Convenção da Diversidade Biológica, na qual a grande maioria dos pontos focais é governamental, na rede GBIF tem-se também a participação de instituições de direito privado, sem fins lucrativos como representantes técnicos oficiais. Na América Latina, os nós técnicos mais atuantes são os do México – a Conabio, uma instituição governamental, interministerial –, da Costa Rica – o INBio, uma organização civil de direito privado sem fins lucrativos –, e da Colômbia – o Instituto Humboldt, uma organização civil de direito privado sem fins lucrativos.

Os dados de coletas e observações de ocorrência de espécies em países da América Latina representam apenas 5,45% do total, sendo que, desse total, apenas cerca de 35% são dados provenientes de instituições latino-americanas<sup>66</sup>. Portanto, a contribuição dos países latino-americanos à rede GBIF sobre dados da região é de cerca de 2% apenas. Essa contribuição é muito pequena, principalmente se levarmos em conta que 6 dos 17 países considerados megadiversos são países latino-americanos. Preocupa também por indicar que os países latino-americanos em geral não estão levantando, organizando e disponibilizando seus dados *on-line*, o que prejudica o desenvolvimento científico e a tomada de decisão local.

### 3.1.1. Pontos Fortes do GBIF

O fato do GBIF ser resultante de um estudo amplo realizado no âmbito da OCDE, fez com que a iniciativa já nascesse com o apoio das nações mais ricas que definiram um modelo colaborativo e financeiramente sustentável. Sua agenda científica e técnica atraiu parcerias em busca de soluções mais objetivas e pragmáticas, diferentemente da CDB cuja interlocução é diplomática e oficial. A natureza jurídica dos acordos bilaterais dos países com o GBIF através de um Memorando de Entendimento não vinculativo (*non-binding Memorandum of Understanding*)

---

<sup>66</sup> Acessar o módulo “*Data sharing with countries of origin*” disponível no endereço <http://data.gbif.org/countries/datasharing?view=percent&host=all&country=all> (acessado no dia 02/09/2012)

mostra uma relação colaborativa sem um compromisso jurídico formal, ambiente este mais próximo ao da cooperação científica.

O GBIF iniciou os seus trabalhos com a missão de deixar disponíveis na Internet de maneira livre e aberta os dados mundiais sobre biodiversidade. “Biodiversidade” no contexto do seu plano de trabalho significava dados sobre a ocorrência de espécies no espaço e no tempo e dados sobre taxonomia. O Secretariado tem como foco integrar os dados provenientes dos países e iniciativas internacionais parceiras, disponibilizando-os de maneira livre e aberta na Internet. Ter foco com metas claras é essencial para um sistema de informação sobre biodiversidade.

Sua arquitetura distribuída, com os dados sendo indexados pelo portal localizado no Secretariado, representa mais um ponto forte. A indexação local pode ser feita pelos nós dos países, responsáveis pela organização e mobilização dos dados de suas respectivas redes de provedores. Isso significa que cada país pode ter soluções locais para tratar as questões de conectividade, capacitação e infraestrutura, otimizando as condições no nó central. Uma rede distribuída em cada país, cujo nó central alimenta o portal no Secretariado, oferece a escalabilidade necessária à rede Global.

A rede de nós, por país, permite ainda um intercâmbio técnico por meio do comitê dos nós (*Nodes Committee*). A ideia é de uma colaboração de mão dupla, na qual os nós podem acessar os documentos e ferramentas disponíveis e compartilhar o seu conhecimento, contribuindo para tornar essa iniciativa verdadeiramente global. A estratégia delineada para os nós visa aumentar a relevância do GBIF para os países participantes, facilitar o trabalho dos nós e promover a colaboração.

### *3.1.2. Pontos Fracos do GBIF*

Para provar o conceito de estruturar uma rede global de dados sobre biodiversidade, tendo como objetivo principal disponibilizar dados de ocorrência de espécies, a estratégia inicialmente adotada foi a de envolver grandes coleções e iniciativas internacionais que já dispunham de dados digitalizados. Essa estratégia apresentou bons resultados e o GBIF pode provar o seu valor. Mas em 2007, quando a rede GBIF disponibilizava cerca de 140 milhões de registros, foi aprovada uma meta de um bilhão de registros *on-line* para ser alcançada em apenas um ano. Assim, para alcançar a meta, a preocupação deixou de ser a estruturação e integração dos nós dos países e o apoio à digitação de dados. Buscou-se apenas aquilo que estava facilmente disponível, sem se preocupar com o grupo taxonômico ou região geográfica dos dados. O resultado é que, além da

meta não ter sido alcançada, cerca de 70% dos registros hoje disponíveis são de observação de pássaros, principalmente do hemisfério norte. Esta mudança de foco foi muito criticada pelo grupo responsável pela avaliação do GBIF (BRIDGEWATER, KNAPP, *et al.*, 2010) pelo grande desequilíbrio geográfico e taxonômico dos dados. O plano estratégico para o período de 2012 a 2016 novamente dá ênfase aos dados dos acervos das coleções biológicas (NIELSEN, 2011).

Uma ameaça importante que precisa ser considerada é a pressão dos diferentes atores, principalmente dos países membros, em obter resultados numéricos mais expressivos em um curto espaço de tempo. Isso poderá significar a perda do foco, da qualidade e uma menor abrangência geográfica e taxonômica dos dados.

Vários países membros estão sendo representados pelos seus ministérios do meio ambiente e não estão estruturando os seus nós de informação, talvez por existir, por parte desses ministérios, uma maior preocupação em desenvolver mecanismos para usar os dados e não de participar da construção dessa infraestrutura de dados científicos para usos múltiplos. A grande ameaça ao dar maior prioridade ao uso dos dados é perder o foco central de antes criar uma plataforma global de dados sobre biodiversidade de acesso livre e aberto para depois (ou concomitantemente) desenvolver ferramentas e aplicativos para o seu uso.

Outro ponto fraco é o número pequeno de nós técnicos, o que, na realidade reflete o pequeno empenho dos países no desenvolvimento das suas estruturas locais. Dos 58 países participantes em onze anos apenas 19 estruturaram seus nós técnicos, sendo que o nó dos Estados Unidos, o *National Biological Information Infrastructure* (NBII)<sup>67</sup>, foi extinto em janeiro de 2012. Além da coordenação técnica e do apoio local aos provedores de dados, existe o intercâmbio de conhecimento e ferramentas entre os nós que poderia ser mais bem explorado.

### 3.2. COMISIÓN NACIONAL PARA EL CONOCIMIENTO Y USO DE LA BIODIVERSIDAD – CONABIO (MÉXICO)

A CONABIO foi criada em março de 1992 com o objetivo de coordenar as ações e estudos relacionados ao conhecimento e à preservação das espécies biológicas do México. Promove e fomenta atividades de pesquisa científica para a exploração, estudo, proteção e utilização dos recursos biológicos visando a conservação dos ecossistemas deste país e atividades para a geração de critérios para o seu manejo sustentável. Tem como missão promover, coordenar, apoiar e realizar atividades voltadas para a conscientização sobre a biodiversidade e sua

---

<sup>67</sup> Para mais informações, acessar <http://www.nbii.gov/termination/index.html> (acessado em 01/09/2012)

conservação e utilização sustentável para benefício da sociedade mexicana. A CONABIO foi concebida como uma organização de pesquisa aplicada, que promove a pesquisa básica, compila e gera informações sobre biodiversidade, desenvolve a capacitação de recursos humanos na área da informática para biodiversidade e é fonte pública acessível de informação e conhecimento para toda a sociedade. Gera inteligência sobre o capital natural do México, servindo de ponte entre a academia, o governo e a sociedade, e procura promover a ideia que a população local é quem deve realizar as ações para a conservação e manejo da biodiversidade. (SARUKHÁN K. e SECO M., 2012).

Uma de suas ações principais é o desenvolvimento e manutenção do Sistema Nacional de Informação sobre Biodiversidade (SNIB)<sup>68</sup>, conforme previsto no artigo 80º, Seção V da Lei Geral de Equilíbrio Ecológico e Proteção Ambiental. O SNIB está sendo desenvolvido para obter, selecionar, agregar valor e sistematizar os dados sobre a biodiversidade do México, disseminando as informações para os diferentes usuários. Uma vez disponibilizados os dados e informações, o segundo passo é desenvolver métodos de análise de dados resultando em produtos para a tomada de decisão, como a distribuição de espécies e o monitoramento de possíveis alterações dos ecossistemas, entre outros. O terceiro passo é a divulgação da informação e do conhecimento que tem por objetivo a geração de inteligência para a tomada de decisão em conservação e uso da biodiversidade.

**Governança**<sup>69</sup>. A CONABIO é uma comissão interministerial de caráter permanente que envolve 10 ministérios. A coordenação é do Ministério do Meio Ambiente e Recursos Naturais (SEMARNAT). Participam dela também os seguintes ministérios: Agricultura, Pecuária, Desenvolvimento Rural, Pesca e Alimentação (SAGARPA), Desenvolvimento Social (SEDESOL), Economia (SE), Educação Pública (SEP), Energia (Sener), Finanças e Crédito Público (SHCP), Relações Exteriores (SRE), Saúde (SSA) e Turismo (Ministério do Turismo). Sua equipe é composta por um Coordenador Nacional e um grupo de cerca de 250 funcionários.

A CONABIO é o representante do governo mexicano na CDB, sendo ponto focal do SBSTTA (*Subsidiary Body on Scientific, Technical and Technological Advice* da CDB) e do GBIF. Além da estratégia global, a CONABIO coordena várias ações que incluem o estabelecimento de estratégias nacionais para vários temas como: Biodiversidade do México;

---

<sup>68</sup> Para mais informações sobre o SNIB, acessar <http://www.conabio.gob.mx/institucion/snib/doctos/acerca.html> (acessado em 19/01/2013)

<sup>69</sup> Para mais informações, acessar <http://www.conabio.gob.mx> (acessado em 19/01/2013)

Espécies invasoras; Estratégia Mexicana para a Conservação Vegetal (EMCV); Estratégias Estaduais de Biodiversidade; Restauração e Compensação Ambiental; e Corredor Biológico Mesoamericano do México. Os países participantes dessa iniciativa incluem México, Belize, Costa Rica, El Salvador, Guatemala, Honduras, Nicarágua e Panamá.

Além de organizar e disseminar a base informacional sobre a biodiversidade do México, a CONABIO está promovendo o desenvolvimento de estratégias estaduais e está atuando de forma integrada com outros países no desenvolvimento de um corredor biológico.

**Provedor de Dados.** A CONABIO trabalha com dados de acervos de coleções biológicas do país e do exterior (repatriação de dados). Trabalha também com catálogos de autoridades taxonômicas ou bases de dados com os nomes científicos de espécies, usados para o controle de qualidade dos dados do SNIB. Financia projetos para preencher lacunas de dados (recursos para digitação de dados e realização de novos inventários) e de conhecimento (produção de análises e publicações). A CONABIO também trabalha com imagens de satélites e sensoriamento remoto, imagens e mapas.

**Usuário-alvo.** Os dados são disponibilizados de forma livre e aberta, sendo importante fonte de informação para o desenvolvimento científico. Mas o principal foco da CONABIO é o uso dos dados para a tomada de decisão. Além de procurar gerar conhecimento para fundamentar políticas públicas, tem como meta subsidiar as decisões da sociedade sobre a sua biodiversidade. Para tanto, dispõe de material para educação voltado a professores do ensino infantil, fundamental e médio.

**Política de dados:** o SNIB explicita a sua política sobre a distribuição de informação em seu site<sup>70</sup>. A política geral inclui: (i) respeito aos termos do convênio estabelecido com o provedor de dados e a legislação vigente sobre propriedade intelectual; (ii) a menos que existam restrições explícitas por parte do provedor, os dados serão disponibilizados ao público; (iii) ao autor ou responsável serão dados os devidos créditos; (iv) será solicitado ao usuário citar a fonte dos dados, solicitar permissão expressa para redistribuir os dados, sendo que os mesmos não podem ser usados em atividades prejudiciais à biodiversidade.

Em relação aos projetos financiados pela CONABIO, a política do SNIB indica que, independentemente do tipo de restrição que houver, a CONABIO pode utilizar os dados para

---

<sup>70</sup> Ver o texto completo no endereço <http://www.Conabio.gob.mx/institucion/snib/doctos/politicas.html> (acessado no dia 15/01/2012).

análises internas e outras análises que tenham como objetivo assessorar o governo. Os dados originais não publicados podem ter uma restrição temporal de até três anos, período que garante prioridade ao autor para a publicação dos resultados de sua pesquisa. Dados com restrições permanentes serão analisados caso a caso e, se alguma restrição for aprovada, ela será avaliada periodicamente. Dados sem restrições, após conclusão satisfatória do projeto e análise de qualidade dos produtos, serão disseminados de forma livre e aberta.

**Tecnologia:** A CONABIO utiliza padrões e protocolos compatíveis com a rede global (GBIF). As instruções para a conformação das bases de dados com o SNIB indicam que os dados de observação devem ser enviados seguindo a estrutura do padrão Darwin Core usando Excel, Access ou arquivo de texto delimitado. Arquivos com o padrão são disponibilizados. Para as coleções biológicas é determinado como padrão usar o *software* Biotica, desenvolvido pela CONABIO ou um modelo compatível com o SNIB (bases de dados relacionais), que inclui os campos obrigatórios.

**Resultados:** O SNIB disponibiliza cinco milhões de registros com dados sobre espécimes de coleções do México e do exterior, além de mais de três mil temas cartográficos<sup>71</sup>, sendo dois mil e quatrocentos de acesso público. Também disponibiliza imagens de sensoriamento remoto (satélites e fotos aéreas) e fotografias e ilustrações (SARUKHÁN K. e SECO M., 2012).

Para atender a demanda de um público mais amplo, o sistema disponibiliza mais de mil e quatrocentas fichas técnicas de espécies. Disponibiliza também um sistema de informação sobre espécies prioritárias (manejo e conservação de espécies em risco com valor biológico) e avaliação do estado atual de 1.370 espécies, além de informações sobre organismos vivos modificados e sobre espécies invasoras.

O portal de geoinformação disponibiliza 1.450 mapas de biodiversidade como monitoramento de manguezais; distribuição potencial de espécies; espécies ameaçadas e prioritárias para conservação, entre outros.

Disponibiliza também sistemas de monitoramento de mudanças de cobertura do solo por sensoriamento remoto, monitoramento oceânico, monitoramento de aves e participação cidadã, monitoramento de manguezais e de incêndios florestais. Apresenta análises de risco de espécies invasoras e de organismos vivos modificados e define as espécies e as áreas onde o México é o centro de origem e de diversidade.

---

<sup>71</sup> Inclui temas ambientais, infraestrutura e sócio-econômicos.

Possui aplicativos para o georreferenciamento de dados e algoritmos para a produção de informação relevante sobre o tema Redução de Emissões para o Desmatamento e Degradação - REDD (*Reduce Emissions for Deforestation and Degradation*) e disponibiliza também uma série de relatórios em resposta às demandas da CDB.

### *3.2.1. Pontos Fortes da Conabio*

Um dos pontos fortes é o mandato oficial e a sua estrutura como comissão interministerial de caráter permanente envolvendo 10 ministérios. Com isso, a participação das instituições públicas que possuem dados de interesse para a CONABIO passa a ser política de Estado. Graças à sua estrutura de governança, é possível à CONABIO assumir a liderança na definição da política de estado sobre biodiversidade de maneira coordenada entre os diferentes ministérios.

A possibilidade de utilizar recursos privados para desenvolver suas ações e ao mesmo tempo contar com recursos governamentais estáveis é um ponto forte que torna possível estabelecer estratégias e ações de longo prazo.

O fato de a CONABIO ser ao mesmo tempo uma agência financiadora de projetos e a instituição responsável por receber, armazenar e disponibilizar os dados resultantes desses projetos, torna o controle e avaliação dos resultados mais eficiente. Assim, além de integrar dados de coleções biológicas do país e do exterior, a CONABIO identifica lacunas de dados e informações e lança editais para preencher estas lacunas, dando apoio à digitação de dados, a levantamentos em campo etc..

Antes de integrar os dados à rede, os mesmos são analisados em relação à sua completude e qualidade taxonômica. Registros incompletos ou com possíveis erros de digitação são, em princípio, descartados, o que beneficia o uso de dados de qualidade para a tomada de decisão.

O caráter utilitário do sistema para a tomada de decisão é mais um ponto forte. Como a CONABIO é o ponto focal do governo para as convenções e acordos internacionais e nacionais sobre biodiversidade, ela tem como influir diretamente em toda a cadeia de informação de forma bastante pragmática, da coleta de dados até a sua transformação em conhecimento e ação.

Como a CONABIO tem como princípio que a população local é quem deve realizar as ações para a conservação e manejo da biodiversidade, outro ponto forte é a informação disponibilizada para públicos não especializados, com especial destaque ao material para professores do ensino infantil ao médio.

### 3.2.2. Pontos Fracos da Conabio

Um problema para a Conabio é ter uma tecnologia consolidada, desenvolvida ao longo dos últimos 20 anos, quando adotou padrões e protocolos próprios e estratégias centralizadoras. Há 20 anos, a opção mais viável para integrar dados de sistemas diversos era através da padronização do *software* e do uso de uma estrutura comum de dados. Para as coleções biológicas, a CONABIO recomenda o uso do *software* por eles desenvolvido (*Biotica*) ou o uso do mesmo modelo de banco de dados orientado a objetos. Ao término do projeto, os dados de observação necessariamente devem ser enviados em planilha Excel ou Access já formatadas nos campos DarwinCore. Essa estratégia não dá liberdade ao provedor de usar o *software* que conhece e pode demandar do provedor de dados um maior conhecimento em computação. Outro problema é que, aparentemente, não há um mecanismo de atualização simples que possa ser feita de forma autônoma pelo provedor, correndo o risco do provedor perder o controle dos seus próprios dados. Com isso, os dados podem ficar desatualizados, principalmente em relação à taxonomia.

A estratégia de só aceitar dados completos e sem erros certamente propicia uma maior segurança em relação à sua qualidade. No entanto, corre-se o risco desses dados nunca serem corrigidos e incorporados ao sistema. A disponibilização *on-line* de dados com erros e/ou incompletos os torna visíveis e, portanto, passíveis de análise, o que por sua vez permite a elaboração de estratégias para melhorar a qualidade dos acervos. Dados incompletos, considerados de menor qualidade para determinados usos, podem ser fundamentais para a definição de estratégias para melhorar a qualidade dos mesmos. É preferível aceitar todos os dados e desenvolver filtros para garantir que os usuários possam recuperar somente os dados que atendam à sua demanda.

### 3.3. A REDE *SPECIESLINK*

A rede *speciesLink*, brasileira, nasceu como estratégia do Programa Biota-Fapesp. Um dos meios para atingir o objetivo do Programa era informatizar todos os acervos e coleções científicas do Estado de São Paulo, e estabelecer uma rede de informação em biodiversidade entre todas as instituições envolvidas com a pesquisa e conservação da biodiversidade no

Estado<sup>72</sup>. O projeto speciesLink teve por objetivo desenvolver um sistema distribuído de informação, de acesso livre e aberto, sobre espécies e espécimes da fauna, flora e microbiota, associado a um sistema de previsão de distribuição geográfica de espécies. Além da integração dos dados das coleções biológicas do estado, o sistema deveria também integrar os dados do SinBiota, o sistema de informação *on-line* desenvolvido para receber e disseminar os dados gerados pelos projetos do Programa Biota-Fapesp (Processo Fapesp 2001/02175-5).

**Governança:** Em linhas gerais, trata-se de uma rede colaborativa na qual o Centro de Referência em Informação Ambiental - CRIA<sup>73</sup> é responsável pelo desenvolvimento e manutenção do sistema de informação e cada provedor é responsável pelo seu acervo e pela qualidade dos dados enviados à rede. A parceria é formalizada através de um protocolo de intenções<sup>74</sup> no qual o CRIA, além da manutenção da rede, se compromete a garantir a integridade dos dados e das informações enviadas pelo provedor, não alterando o conteúdo. Também garante o reconhecimento da origem dos dados. É atribuição de cada provedor enviar os dados não sensíveis à rede indicando as restrições ao uso, se houver. O provedor de dados tem total autonomia e as ferramentas necessárias para remover todos os dados sob a sua custódia quando quiser. Portanto, todo controle dos dados é do provedor. Quando existe um projeto em andamento que utiliza a rede *speciesLink* como base informacional, normalmente existe um comitê gestor onde o CRIA tem assento. O comitê tem um papel fundamental na articulação da rede e na definição de novos aplicativos e produtos.

**Provedor de dados:** O principal provedor de dados são as coleções biológicas. O sistema, mesmo quando era exclusivamente financiado pela Fapesp, sempre esteve aberto à participação de coleções de outros estados e do exterior que queiram repatriar os dados das coletas realizadas no Brasil. Diferentemente dos grandes acervos das coleções da Europa e Estados Unidos, as coleções biológicas no Brasil são, na sua grande maioria, coleções de pequeno porte. A rede *speciesLink* em agosto de 2011, integrava os dados de 207 coleções e subcoleções biológicas do

---

<sup>72</sup> Os objetivos, meios e produtos definidos para o Programa Biota-Fapesp foram publicados on-line em 1999 e continuam disponíveis no endereço <http://www.biota.org.br/info/metad> e também na Fapesp no endereço <http://www.biota-fapesp.net/meios.html> (acessado em 19/01/2013).

<sup>73</sup> O CRIA é uma associação civil, sem fins lucrativos que tem como ação referencial dotar governo e sociedade de informações necessárias para o estabelecimento de prioridades e ações para a promoção do desenvolvimento sustentável. Entre outras ações, oferece apoio à comunidade científica e tecnológica na organização, estruturação e disseminação de seus dados e informações. Veja o Estatuto Social disponível na página <http://www.cria.org.br/about>

<sup>74</sup> O protocolo está disponível no endereço <http://splink.cria.org.br/files/protocolodeintencoes.doc> (acessado em 02/09/2012)

Brasil, e, dessas, mais da metade (111) eram coleções mantidas em universidades. Nessa data, 88% das coleções participantes da rede possuíam acervos menores que 100 mil registros. Juntas, essas coleções representavam 26% do acervo total das coleções brasileiras participantes da rede *speciesLink*, mas eram responsáveis por cerca de 50% dos dados disponíveis *on-line*<sup>75</sup>. As coleções menores, via de regra, têm mais facilidade na informatização de seus acervos e entre elas existem coleções com grande especificidade geográfica e taxonômica, que são muito importantes para o conhecimento da biodiversidade do país. Por outro lado, as coleções menores têm problemas de pessoal e infraestrutura e muitas não são formalmente reconhecidas como coleções por suas instituições. O sistema também integra em menor número (cerca de 6% do total) dados de observação e dados compilados por especialistas.

**Usuário Alvo:** Os principais usuários alvo são a comunidade científica e gestores de políticas públicas.

**Política de Dados:** Cada provedor de dados tem liberdade para definir as condições para utilização dos dados de sua instituição, sendo que todo dado enviado à rede é de acesso livre e aberto.

**Tecnologia:** O início dos trabalhos de desenvolvimento da rede *speciesLink* em 2001 coincidiu com o início do desenvolvimento da rede GBIF. Ao invés de desenvolver padrões e protocolos próprios, a equipe do CRIA optou por trabalhar com as equipes do GBIF e das Universidades de Kansas e Califórnia nos Estados Unidos. O resultado foi o desenvolvimento do modelo de dados DarwinCore (DwC) e do protocolo de comunicação DiGIR que agora vem sendo substituído pelo protocolo TAPIR. O uso de sistemas e padrões não proprietários, adotados internacionalmente, é fundamental para uma infraestrutura aberta de dados.

A arquitetura da rede é necessariamente distribuída, dado o grande número de coleções biológicas do país e das amostras coletadas no Brasil que estão depositadas em coleções do exterior. Uma arquitetura distribuída pressupõe que cada provedor de dados tenha um servidor instalado na rede (com conectividade rápida e estável) e disponível 24 horas por dia, 7 dias por semana. No Brasil, a maioria das coleções biológicas tem problemas com sua infraestrutura física (*hardware* e redes) e com pessoal, principalmente em relação a técnicos em informática. Para viabilizar a participação dessas coleções de maneira autônoma, foi desenvolvido um aplicativo

---

<sup>75</sup> Os dados sobre as coleções participantes da rede, seus acervos e o número de registros *on-line* podem ser pesquisados na página do sistema de gerenciamento da rede - <http://splink.cria.org.br/manager>

(*spLinker*<sup>76</sup>) que é responsável pelo mapeamento dos campos utilizados pela coleção para o padrão DarwinCore e pelo envio dos dados não sensíveis a um servidor regional (*cache node*). Com isso o provedor, além de continuar usando o *software* que atende às suas necessidades e que ele sabe usar, mantém todo o controle sobre os dados enviados à rede. Assim, para participar da rede, a coleção não precisa manter uma equipe e infraestrutura em informática nem depende de uma linha de comunicação rápida e estável. Um diagrama da arquitetura da rede é apresentado na Figura 3.2 a seguir.

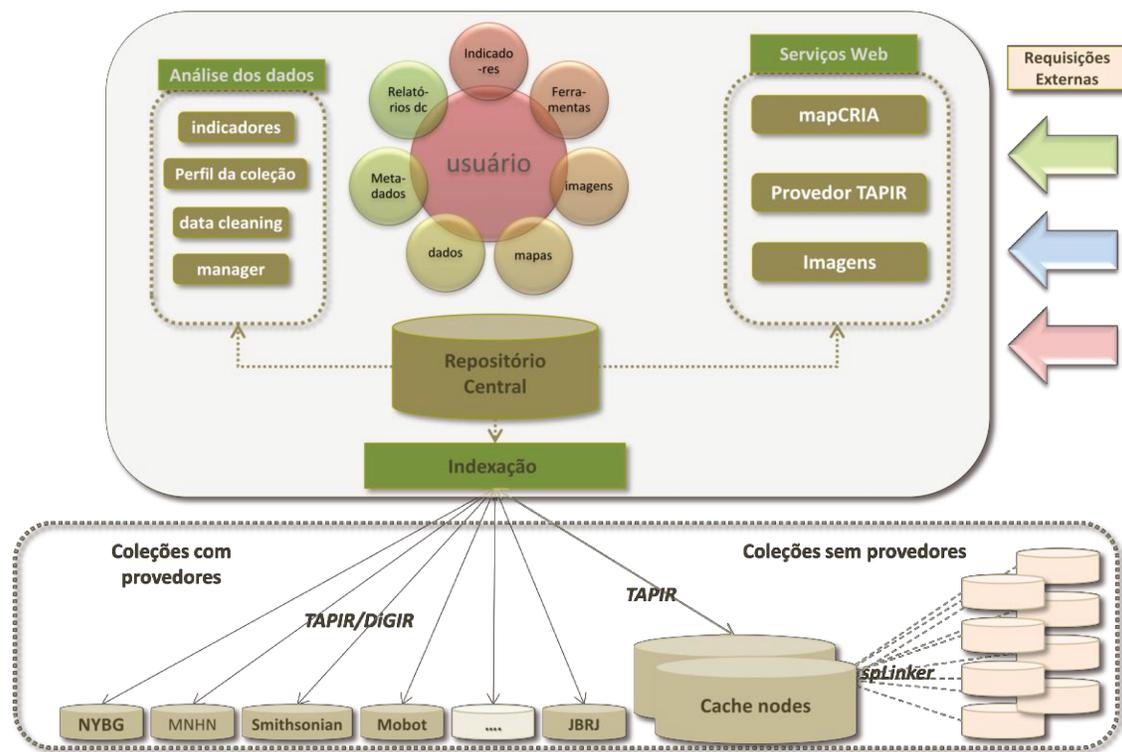


Figura 3.2. Diagrama da arquitetura da rede *speciesLink*  
 Fonte: CRIA (2011)

As coleções atualizam seus dados em seus provedores ou nos *cache nodes* da rede através do uso do *software spLinker*. Novos dados são indexados toda noite e depositados em um repositório central, o que possibilitou o desenvolvimento de aplicativos para analisar os dados (p.ex. *data cleaning*<sup>77</sup> e indicadores<sup>78</sup>), e de diferentes interfaces para visualizar o sistema, como a

<sup>76</sup> Para mais informações acessar <http://splink.cria.org.br/splinker>

<sup>77</sup> Para conhecer melhor os aplicativos denominados *data cleaning* acessar a página <http://splink.cria.org.br/dc> e escolha alguma coleção (página acessada no dia 22 de agosto de 2011)

<sup>78</sup> Para conhecer melhor os aplicativos denominados *indicadores*, acessar a página <http://splink.cria.org.br/indicators> e escolha as opções disponíveis (página acessada no dia 22 de agosto de 2011)

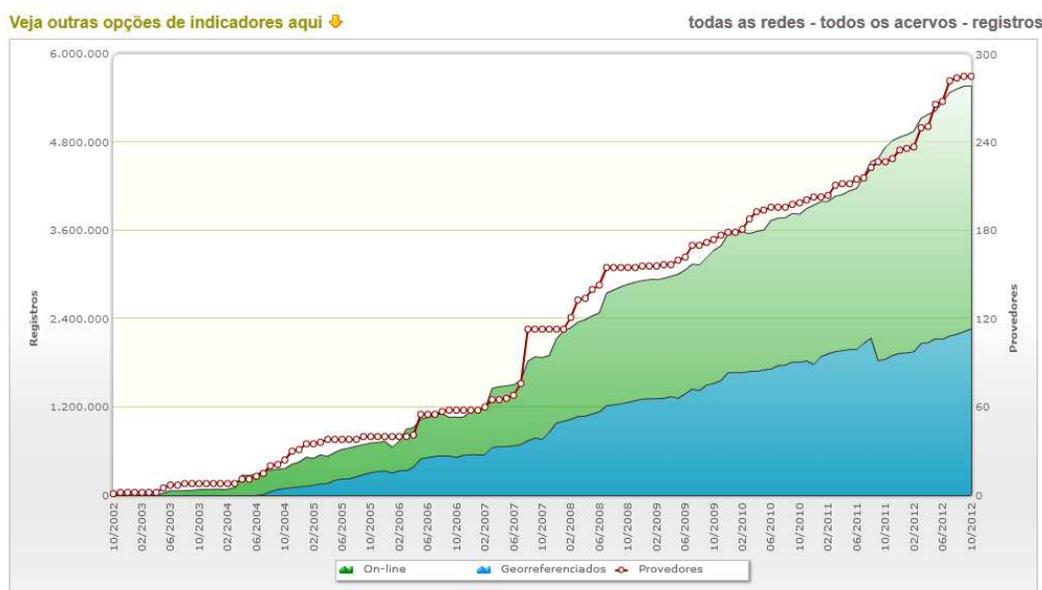
apresentação de redes temáticas. Foram também desenvolvidos alguns serviços *web* (serviço de mapas, provedor dos dados Tapir e um serviço de imagens) disponíveis para aplicativos internos e externos ao CRIA.

O conceito de serviços *web* é muito importante e nada mais é do que um método de comunicação em rede entre dois dispositivos eletrônicos. O W3C (*World Wide Web Consortium*) define o serviço *web* como sendo um sistema desenhado para dar suporte a interações interoperáveis máquina-máquina sobre uma rede. Essa tecnologia abre um leque de possibilidades de uso dos dados por diferentes sistemas, inclusive de diferentes áreas de interesse. Os aplicativos desenvolvidos para a rede *speciesLink* procuram atender às demandas e necessidades do seu público-alvo. Mas dados de ocorrência de espécies são importantes para outras áreas de conhecimento que, dessa forma podem usar essa infraestrutura pública via serviços (acesso máquina- máquina) e se concentrar na integração de dados e no desenvolvimento dos aplicativos que atendem às demandas de seu público-alvo. Essa tecnologia abre novas possibilidades e uma nova estratégia para sistemas de informação científica com a agregação de diferentes serviços para produzir diferentes produtos.

**Resultados.** A rede *speciesLink* no dia 18 de outubro de 2012 disponibilizava mais de 5,6 milhões de registros *on-line* procedentes de 278 coleções e subcoleções biológicas (5,3 milhões de registros) e 10 coleções de dados de observação (cerca de 334 mil registros). Apesar do término do apoio da Fapesp para o desenvolvimento e manutenção da rede em 2005, o sistema provou ser dinâmico e de interesse para outros projetos. A rede usa um padrão de dados reconhecido internacionalmente e o provedor não precisa alterar o seu *software* ou aprender a usar novos sistemas para integrar seus dados à rede. Ao longo dos anos, novas redes temáticas (p.ex. rede SICol para coleções microbianas, Polinizadores e o INCT Herbário Virtual da Flora e dos Fungos) e redes geográficas (p. ex. Rede Paranaense TAX*on-line* e o PPBio Amazônia Ocidental) foram propostas e obtiveram apoio de diferentes agências de fomento. Essas redes possuem seus próprios *websites*, organizam a informação de acordo com os seus interesses, mas o sistema informacional dos dados dos acervos biológicos é a rede *speciesLink*.

A Figura 3.3 apresenta a média mensal do número de registros disponíveis *on-line*, o número de registros georreferenciados e o número de provedores. Pode-se observar um crescimento expressivo após o término do projeto da Fapesp em outubro de 2005, o que reflete o apoio de outros projetos e sua arquitetura escalável, mas também reflete uma mudança cultural da

comunidade de coleções biológicas que passou a compreender melhor as vantagens do compartilhamento aberto de dados *on-line*.



Histórico do envio e retirada de dados da rede. São apresentadas as médias mensais, tanto do número total de registros on-line, como também do número de registros georreferenciados. A linha vermelha traz o número mensal de provedores de dados (coleções biológicas ou de dados).

Atualizado em 13/10/12 04:02

Figura 3.3. Evolução do número de registros *on-line* da rede *speciesLink*  
 Fonte: Indicadores *speciesLink* (outubro, 2012)

Além dos dados textuais, os herbários estão enviando imagens das amostras de seus acervos. São mais de 200 mil imagens<sup>79</sup>, sendo que mais de 14 mil são imagens em alta resolução de *tipos*<sup>80</sup>. O acesso às imagens das amostras é importante para a identificação de material e pode também ser usado para fins didáticos.

Associado ao conteúdo e qualidade, a disponibilização dos dados em formatos úteis e utilizáveis determina a usabilidade de um sistema. Ao longo dos anos, os produtos ou saídas dos dados foram sendo aprimorados e diversificados. A interface de busca e recuperação dos dados permite a visualização dos resultados como texto, mapas, gráficos, imagens e planilhas. Também foram desenvolvidos aplicativos e ferramentas como indicadores, relatórios de *data cleaning* e o openModeller<sup>81</sup>, um ambiente computacional para a geração de modelos de distribuição potencial de espécies.

<sup>79</sup> Dados de novembro de 2012.

<sup>80</sup> Um espécime *tipo* é o espécime selecionado para servir de referência quando uma espécie é nomeada pela primeira vez.

<sup>81</sup> Para mais informações sobre a ferramenta openModeller, ver <http://openmodeller.cria.org.br>

O aumento da quantidade e da usabilidade dos dados tem como consequência o aumento do acesso aos dados da rede *speciesLink*. Além dos indicadores usuais de acesso como *hits*, visitas e *bandwidth*, para a rede *speciesLink* existe um contador que monitora o número de registros que o usuário visualiza. A tabela a seguir apresenta esses números.

Tabela 3.1. Acesso aos dados da rede *speciesLink*

Ano	2007	2008	2009	2010	2011	Crescimento de 2007 a 2011
Nº. de Registros <i>on-line</i>	2.131.278	2.939.987	3.551.813	4.012.333	4.889.678	131%
Nº. de Registros visualizados (download)	2.742.558	14.310.948	20.211.348	92.546.559	256.496.993	9.252%

Fonte: estatísticas rede *speciesLink* (2012)

Outro desenvolvimento importante é a possibilidade do usuário enviar comentários sobre o dado que está visualizando, diretamente ao curador. Isso viabiliza um processo de interação direta entre o usuário e o provedor dos dados, que passa a poder contribuir para com a correção de erros e até a identificação do material.

### 3.3.1. Pontos Fortes da rede *speciesLink*

O primeiro fator importante que merece destaque é a arquitetura. Assim como para a rede GBIF, uma característica importante é a sua escalabilidade, ou seja, a arquitetura permite a inclusão de novos provedores de dados sem que seja necessário realizar novos desenvolvimentos.

O controle do provedor sobre seus dados é outro ponto forte que também diz respeito à arquitetura. Para compartilharem os seus dados com a rede, as coleções podem continuar utilizando o *software* de sua preferência, determinando que dado é público ou sensível e indicando possíveis restrições ao seu uso. Essa estratégia tecnológica foi determinante para o sucesso da rede. A arquitetura permite que cada coleção tenha sua própria política de dados, tornando desnecessária a adoção de uma política única, que dificilmente seria consensual.

Ainda em relação à arquitetura, o uso dos mesmos padrões e protocolos adotados pela rede GBIF foi outro fator positivo uma vez que facilitou a estratégia de repatriação de dados, hoje responsável por cerca de 12% dos registros disponíveis *on-line*.

O desenvolvimento de redes temáticas também provou ser estratégico. Aumenta o foco, envolvendo grupos de interesse comum que lançam um novo olhar sobre os dados, criando novas demandas e gerando novos produtos.

Um outro ponto forte são os aplicativos desenvolvidos para auxiliar as coleções biológicas na identificação e correção de erros e para agregar valor aos dados (p.ex. os aplicativos para georreferenciamento automático).

O sistema de busca, recuperação e visualização dos dados em diferentes formatos aumenta a utilidade e usabilidade dos dados, constituindo mais um ponto forte da rede.

O desenvolvimento de serviços *web* associados aos produtos da rede amplia o uso dos dados por outros sistemas, valorizando ainda mais o trabalho das coleções biológicas na identificação, catalogação, manutenção e disseminação dos dados e informações de seus acervos.

Por fim, a centralização da infraestrutura mais complexa onde existe uma equipe de informática altamente especializada que dá apoio para viabilizar a participação das coleções biológicas do país na rede reduz o custo em comparação com sistemas que demandam das coleções, infraestrutura e competência em informática. Pelo fato de estar essa equipe em uma instituição privada, sem fins lucrativos e de interesse público, ela tem maior agilidade e capacidade de estabelecer parcerias, ao mesmo tempo em que mantém o seu caráter público. Trata-se também de uma instituição que as coleções biológicas consideram “neutra”, já que não compete com elas pelos mesmos recursos.

### *3.3.2. Pontos Fracos da rede speciesLink*

A maior fragilidade da rede é quanto à sua sustentabilidade. A manutenção e o desenvolvimento dessa infraestrutura pública de dados dependem de recursos de projetos de curta duração, não dispendo de recursos de longo prazo para manutenção de uma equipe mínima e de sua infraestrutura básica. Várias coleções biológicas participantes, principalmente aquelas instaladas em universidades, não têm apoio nem o reconhecimento institucional necessários para garantir a qualidade dos acervos e a digitação de seus dados.

A segunda fragilidade é a falta de estratégias e planejamento de longo prazo. Como não existem mecanismos de financiamento de longo prazo, as estratégias e produtos acabam sendo determinados por projetos de curta duração.

Outro debate necessário é sobre a natureza jurídica do CRIA (Centro de Referência em Informação Ambiental). Trata-se de uma associação civil, sem fins lucrativos, qualificada pelo

Ministério da Justiça como uma Organização da Sociedade Civil de Interesse Público – OSCIP<sup>82</sup>. Se por um lado, a manutenção de uma e-infraestrutura sobre biodiversidade requer agilidade, capacidade de estabelecer parcerias, capacidade de absorver rapidamente mudanças tecnológicas, capacidade de gestão e uma administração ágil – e essas características estão presentes em uma OSCIP –, por outro lado não dispõe de recursos financeiros estáveis e contínuos para garantir a sua manutenção. Resumindo, o CRIA presta um serviço de interesse público, não lucrativo e não dispõe de recursos públicos de longo prazo para a sua manutenção.

Recentemente o CNPq solicitou junto à RNP (Rede Nacional de Ensino e Pesquisa, uma OS – Organização Social<sup>83</sup>) que fosse autorizado ao CRIA hospedar no IDC (Internet Data Center) em regime de *colocation*, seus equipamentos utilizados para disseminar de forma livre e aberta os dados e informações sobre biodiversidade. Essa solicitação foi aceita e toda a infraestrutura pública de dados, hoje hospedada no CRIA, será em breve transferida ao IDC da RNP, o que dará maior segurança ao sistema e aos dados.

Resta ainda uma terceira fragilidade, a ausência de um comitê gestor representativo dos principais atores envolvidos com a rede, entre eles os provedores de dados, os usuários alvo, o CRIA, a RNP, os agentes financiadores (públicos e privados) e o MCTI. Outros ministérios possuem importantes acervos e têm grande interesse no uso dos dados como o Ministério da Saúde, do Meio Ambiente e da Educação e também poderiam participar. A composição desse comitê deve ser muita bem estudada para ser ao mesmo tempo representativa e ágil.

---

<sup>82</sup> OSCIP é uma qualificação decorrente da lei 9.790 de 23/03/99, também conhecida como Lei do Terceiro Setor. É o reconhecimento legal e oficial das ONGS, principalmente pela transparência administrativa que a legislação exige. Para obtenção da qualificação como OSCIP, o CRIA se enquadra no seguinte objetivo social: *defesa, preservação, conservação do meio ambiente e promoção do desenvolvimento sustentável*.

<sup>83</sup> OS É uma qualificação que pode ser outorgada pelo Poder Executivo às pessoas jurídicas sem fins lucrativos que a pleitearem e cumprirem os requisitos legais para obtê-la. Um dos requisitos é a participação, no órgão colegiado de deliberação superior, de representantes do Poder Público.

Tabela 3.2. Resumo dos principais pontos levantados sobre as três infraestruturas de dados analisadas

Itens	GBIF	CONABIO	speciesLink
<b>Motivação</b>	Resultante do fórum de Megaciência da OCDE, estabelecida por governos	Criada com o objetivo de coordenar as ações e estudos relacionados ao conhecimento e à preservação das espécies biológicas do México	Desenvolvimento inicial motivado pelo programa Biota/Fapesp. Após 2005 ganhou escopo nacional.
<b>Início</b>	2001	Março de 1992	Outubro de 2001
<b>Provedor de Dados</b>	Membros do GBIF ou por eles determinados	Coleções biológicas do país e do exterior e projetos financiados para preencher lacunas.	Coleções biológicas do país e do exterior (94% dos dados) e bancos de dados de observação.
<b>Usuário-Alvo</b>	Comunidade Científica (principal) e tomadores de decisão	Principal usuários alvo é o Governo Mexicano, mas também foca em professores do ensino infantil ao médio.	Comunidade científica e gestores de políticas públicas.
<b>Política de dados</b>	Acesso livre e aberto	Dados podem ter acesso restrito, mas sempre estão disponíveis para análises internas.	Acesso livre e aberto.
<b>Tecnologia</b>	Rede distribuída de provedores e um servidor central mantido pelo Secretariado. Uso de padrões e protocolos internacionalmente aceitos	Modelo centralizado. Provedores deve usar o software Biotica ou enviar os dados em planilha Excel com a estrutura do DarwinCore	Rede distribuída de provedores e um servidor central mantido pelo CRIA. Uso de padrões e protocolos aceitos internacionalmente. Provedores podem utilizar o software de sua preferência. Serviços web implementados.
<b>Resultados</b>	390 milhões de registros on-line; mais de 70% dados de observação, predominantemente aves. 75% dos dados são do hemisfério norte	5 milhões de registros, 3 mil temas cartográficos, imagens de sensoriamento remoto, fotografias, 1.370 fichas técnicas de espécies, relatórios e material didático.	5,6 milhões de registros on-line; aplicativos de data cleaning, indicadores, gráficos, mapas, imagens, modelagem de nicho ecológico. Mecanismos que possibilitam ao usuário interagir com o provedor de dados.

<b>Itens</b>	<b>GBIF</b>	<b>CONABIO</b>	<b>speciesLink</b>
<b>Governança</b>	Memorando não vinculativo com governos e organizações intergovernamentais e internacionais. Conselho de Administração com membros votantes e países e instituições associadas; Comitê Executivo, Secretariado	Comissão interministerial de caráter permanente que envolve 10 ministérios.	Protocolo de intenção entre provedores e o CRIA; dependência exclusiva de projetos; comitê gestor, quando existe, é por projeto.
<b>Principais pontos fortes</b>	Modelo colaborativo e financeiramente sustentável; agenda científica; natureza jurídica; foco; arquitetura distribuída e escalável; estímulo a estruturação de nós de países; uso de padrões e protocolos internacionais.	Mandato oficial; estrutura como comissão interministerial envolvendo 10 ministérios; possibilidade de captar recursos privados e contar com recursos governamentais estáveis; caráter utilitário do sistema; informação para públicos não especializados.	Arquitetura distribuída e escalável; provedor com total controle sobre os seus dados; uso de padrões e protocolos adotados internacionalmente; redes temáticas usando um banco de dados comum; aplicativos e serviços web; complexidade em termos da informática está no CRIA (e não no provedor).
<b>Principais pontos fracos</b>	Metas numéricas vs qualidade dos dados; poucos nós estruturados	Tecnologia centralizada e rigidez no software utilizado pelos provedores de dados; não tratamento dos dados com problemas de qualidade e completude.	Sustentabilidade financeira; ausência de estratégias de longo prazo (sempre baseadas em projetos); Ausência de um comitê gestor representativo.

### 3.4. POLÍTICAS E ESTRATÉGIAS PARA INFRAESTRUTURAS ELETRÔNICAS DE DADOS SOBRE BIODIVERSIDADE NO BRASIL

O objetivo da análise, ao indicar os pontos fortes e fracos de cada uma das iniciativas estudadas, é identificar as variáveis que podem afetar, de forma positiva ou negativa, uma estratégia para infraestruturas eletrônicas de dados sobre biodiversidade no Brasil. A formulação de uma estratégia deve considerar essas variáveis ao determinar um plano de ação que possa potencializar ou mitigar os seus efeitos. A análise aqui apresentada deve ser considerada como ponto de partida para um debate maior, envolvendo outros atores, principalmente gestores e formuladores de políticas públicas.

A concepção do GBIF como um modelo colaborativo, financeiramente sustentável e com uma forte agenda científica deu o foco necessário para uma governança participativa. Sua arquitetura distribuída, com um nó central no Secretariado, garante a autonomia para que cada país possa desenvolver sua própria infraestrutura de acordo com a sua capacidade e interesses locais. Trata-se de uma importante oportunidade para o Brasil participar de forma ativa na rede global.

O ponto fraco em relação ao número pequeno de nós técnicos e a constatação de que a contribuição dos países latino americanos é pequena, é mais uma questão da política local, de entender como prioridade o desenvolvimento de e-infraestruturas sobre biodiversidade para o desenvolvimento sustentável local e de compreender as vantagens de participar de uma iniciativa global.

Não basta ao país ter acesso a dados e informações disponibilizadas em infraestruturas internacionais. A organização local dos dados e a sua disponibilização *on-line* com foco nas suas necessidades contribui para uma maior organização da comunidade científica e viabiliza estudos de problemas locais, aumentando a importância do tema “biodiversidade” nas agendas científicas e políticas de um país. Essa organização também promove o desenvolvimento científico com a possibilidade do uso e reuso dos dados e da integração de dados de diferentes fontes e áreas de conhecimento. Contribui também para estudos e para a resolução de problemas globais, ao disponibilizar e integrar dados, informações e conhecimentos locais nas infraestruturas globais.

A CONABIO percebeu a importância de estruturar o seu próprio sistema de informação há vinte anos. Tem em seu mandato e estrutura interministerial o seu ponto forte, mas esse desenho institucional dificilmente poderá ser reproduzido em países como o Brasil que tem várias

instituições já estabelecidas desempenhando parte das atividades da CONABIO, como o CRIA (rede *speciesLink*), o INPE (imagens de satélite) e o IBGE (mapas), entre muitos outros. No Brasil, dada a diversidade de instituições e da grande distribuição das coleções biológicas, é necessário trabalhar em rede.

Como melhorar a efetividade da associação do conhecimento à ação é um enorme desafio para as e-infraestruturas. E esse desafio é ainda maior quando o público-alvo é mais amplo e inclui as populações locais. A CONABIO tem alcançado excelente resultados no uso dos dados para a tomada de decisão, produzindo análises de fácil compreensão para públicos não especializados. A estratégia de produzir material didático para professores do ensino infantil ao médio usando os dados da e-infraestrutura também parece ser um caminho interessante para o Brasil considerar.

Assim, e experiência da Conabio, do GBIF e de muitas outras iniciativas não avaliadas nesta tese, representam uma oportunidade de estudo e troca de informações e conhecimento que deve ser explorada pela infraestrutura brasileira.

A rede *speciesLink* utiliza padrões e protocolos compatíveis com a rede GBIF, o que significa que o Brasil poderá participar da rede de forma plena. Outra característica estratégica da rede distribuída é a autonomia dos provedores em relação a seus dados, determinando o que é de acesso público e pode ser compartilhado. Porém é fundamental estudar possíveis modelos de governança e de sustentabilidade de longo prazo.

Além da integração de dados de coleções biológicas, as e-infraestruturas representam um importante componente da política científica para dados que já nascem digitais. A *National Science Foundation* (NSF), agência de fomento dos Estados Unidos, solicita em seus editais que toda proposta apresente um plano para a gestão e disseminação dos dados<sup>84</sup>. Divulgar os resultados de uma pesquisa *on-line* com os dados associados é certamente um importante passo, mas não é suficiente. Se não houver uma infraestrutura capaz de receber os dados e prover um sistema integrado de busca, recuperação e visualização desses dados, corre-se o risco do dado ser depositado de forma isolada, sem qualquer tipo de integração com outros sistemas e, ainda, de não estar mais disponível após o término do projeto.

Muitas universidades norte-americanas estão criando repositórios institucionais para facilitar o depósito de conteúdo digital e, dessa forma, garantir a sua preservação e a

---

<sup>84</sup> Para mais informações, acessar <http://www.nsf.gov/bfa/dias/policy/dmp.jsp> (acessado em 02/09/2012)

possibilidade de compartilhamento. Essas infraestruturas certamente facilitam o trabalho do pesquisador que, na proposta apresentada ao NSF, precisa somente garantir que os dados serão depositados no repositório da universidade. Mas essa estratégia não é a mais adequada em áreas que necessitam do compartilhamento maciço de dados e aplicativos, como é o caso da biodiversidade. Para essas áreas é fundamental estabelecer e-infraestruturas capazes de receber os dados e se responsabilizar por toda a cadeia de atividades, desde o depósito até a disseminação dos dados, inclusive possibilitando a integração desses dados com os de outras áreas de conhecimento.

É, comparativamente, a diferença entre exigir que cada pesquisador publique o seu currículo *on-line* e que cada pesquisador publique o seu currículo *on-line* na plataforma Lattes. A integração dos dados em uma plataforma única abre um leque de possibilidades de uso e análise que seria impossível de realizar com dados disponibilizados de forma isolada.

No Brasil, o Programa Biota-Fapesp exigiu o depósito dos dados resultantes do financiamento do programa em um sistema desenvolvido com recursos de um projeto aprovado no contexto do programa (SinBiota). No entanto, a estratégia foi incompleta por não ter garantido a sustentabilidade financeira da infraestrutura, tratando-a como um projeto comum, com início, meio e fim, e não como uma atividade central do Programa.

Infraestruturas públicas de dados devem ser tratados como importantes serviços científicos e necessitam de estratégias de financiamento de longo prazo. Em abril de 2012 o MCTI lançou o projeto do Sistema de Informação sobre a Biodiversidade Brasileira (SIBBr). É uma oportunidade para discutir modelos de governança e implementar mecanismos de apoio de longo prazo às e-infraestruturas sobre biodiversidade.

A definição de uma estratégia para o estabelecimento e manutenção de longo prazo de uma infraestrutura eletrônica de dados sobre biodiversidade, de acesso livre e aberto, inclui estratégias para o seu uso, não só para pesquisas científicas, mas para realimentar e acompanhar políticas e estratégias públicas relativas ao meio ambiente e à própria infraestrutura.

O Capítulo 4 a seguir, analisa o uso da infraestrutura brasileira, a rede *speciesLink*, para a elaboração de estratégias e políticas públicas em biodiversidade. São apresentados três potenciais usos: (i) para o monitoramento e acompanhamento de programas e projetos e para a valorização de serviços científicos; (ii) na identificação de lacunas de dados e conhecimento; e (iii) na definição de estratégias de apoio à informatização dos acervos de coleções biológicas do país.

## CAPÍTULO 4. E-INFRAESTRUTURAS E POLÍTICAS PÚBLICAS

Fomentar o desenvolvimento e a manutenção de infraestruturas de dados sobre biodiversidade de acesso público para armazenar, organizar, recuperar e disseminar *on-line* dados e informações de qualidade geradas com financiamento público parece ser, como já indicado nesse trabalho anteriormente, uma estratégia importante para monitorar e compreender a biodiversidade em âmbitos globais, nacionais e locais. Mas é verdade também, que sem competência para se apropriar de dados e informações disponibilizadas nessas e-infraestruturas, as iniciativas de desenvolvimento de bancos de dados abertos não vão além da disponibilidade de números que nada ou pouco representam.

Assim, além de desenvolver e aplicar uma estratégia clara de disseminação livre e aberta de dados sobre biodiversidade e de financiar o desenvolvimento e a manutenção contínua das e-infraestruturas, é importante o uso desses sistemas de informação na elaboração de políticas públicas para melhorar a qualidade, confiabilidade e completude dos dados e informações e para promover uma maior interação do conhecimento também via redes digitais.

Este capítulo analisa, a partir da e-infraestrutura brasileira, a rede *speciesLink*, três (potenciais) usos da referida rede. A primeira análise recai sobre o uso dos dados para monitoramento e acompanhamento de programas e projetos e para serviços científicos, a partir de indicadores gerados por aplicativos que utilizam os dados disponíveis *on-line*.

A segunda análise associa os dados de diferentes fontes de informação gerando um diagnóstico *on-line* das possíveis lacunas de dados ou de conhecimento da flora e dos fungos do Brasil. A partir desse resultado, propõe o uso de aplicativos em bases dinâmicas de dados para subsidiar a análise de especialistas na orientação de novas pesquisas e na caracterização e monitoramento do status de conservação de espécies.

A terceira análise coloca no foco as coleções biológicas, que são as provedoras de dados primários sobre biodiversidade para as redes digitais. A análise parte de um levantamento junto às coleções, realizado no âmbito desta tese, e da avaliação das informações disponíveis *on-line*, buscando fazer um diagnóstico com a participação ativa das coleções biológicas visando promover uma maior interação na definição de ações prioritárias e no acompanhamento dos resultados.

Procura-se, com esses exemplos, mostrar a oportunidade, viabilidade e importância de usar e-infraestruturas em biodiversidade, não só para ampliar o acesso e a usabilidade dos dados para o desenvolvimento científico, mas para os processos de monitoramento e avaliação de programas e projetos e para o desenvolvimento de estratégias de fomento à pesquisa pelo poder público. Esse uso das e-infraestruturas também depende da qualidade e confiabilidade dos dados e das informações e das estruturas digitais que sustentam essas iniciativas.

A metodologia usada baseou-se em diferentes métodos e ferramentas que serão comentadas quando da apresentação de cada exemplo.

#### 4.1. MONITORAMENTO E ACOMPANHAMENTO DE PROGRAMAS, PROJETOS E SERVIÇOS CIENTÍFICOS UTILIZANDO AS E-INFRAESTRUTURAS

Apesar dos grandes avanços nos estudos sobre sistemas de avaliação de programas de C&T, principalmente nos Estados Unidos onde a pressão política é muito grande e a liberação de verbas é cada vez mais dependente da avaliação da qualidade, relevância e produtividade da pesquisa, o método básico de avaliação ainda é o uso de painéis de especialistas (COZZENS, 2012). O uso de aplicativos capazes de sintetizar a informação e produzir gráficos e relatórios de fácil compreensão pode constituir uma importante ferramenta para gerar indicadores para subsidiar processos de monitoramento e avaliação externa por especialistas de programas e projetos.

O uso dos dados on-line para o acompanhamento dos projetos associados à rede *speciesLink* nasceu primeiro da necessidade de monitorar o status e a evolução de cada provedor de dados participante da rede. O objetivo era ter uma ferramenta para facilmente identificar os provedores com problemas com a atualização de seus dados para buscar soluções. Essa ferramenta, denominada *network Manager*<sup>85</sup> é utilizada até hoje e mostra a relação de todos os provedores de dados (*on-line* e *off-line*), cidade e estado de origem, *software* usado, tamanho do acervo, quantidade de registros *on-line*, quantidade de registros georreferenciados e data da última atualização.

---

<sup>85</sup> A ferramenta está disponível no endereço <http://smlink.cria.org.br/manager>

Para estruturar a tabela on-line, a ferramenta acessa os metadados<sup>86</sup> e demais informações necessárias de cada provedor nos bancos de dados dos *providers* ou *cache nodes* (*Tapir* ou *DiGIR* (Figura 3.2)), em tempo real. Essa tabela era somente disponível para acesso interno.

A partir de 2007, a arquitetura da rede passou a contar com um repositório central de dados, tornando factível desenvolver outros aplicativos a partir da avaliação dos registros disponíveis *on-line*.

Para especificar o conjunto de aplicativos que poderiam contribuir para monitorar as atividades da rede e para motivar a participação das coleções, foram estudadas as análises apresentadas por diversos autores no livro “Coleções Biológicas de Apoio ao Inventário, Uso Sustentável e Conservação da Biodiversidade” (PEIXOTO, 2003) e nos diversos documentos de apoio para o desenvolvimento da proposta de “Diretrizes e Estratégias para a Modernização de Coleções Biológicas Brasileiras e a Consolidação de Sistemas Integrados de Informação sobre Biodiversidade” (EGLER e SANTOS, 2006)<sup>87</sup>.

Foram enviados e-mails para os curadores e responsáveis pelos 76 provedores ativos da rede em julho de 2007, solicitando sugestões em relação a possíveis indicadores de desempenho de interesse para as coleções. Também foram estudados os relatórios de atividades apresentados pelos bolsistas responsáveis pela digitação de dados no projeto Fapesp 2001/02175-5 e pelos estagiários contratados para o projeto *Brazilian Bee Data* (contrato no. 2007-02, GBIF).

O resultado do estudo e da contribuição dos provedores foi a identificação de indicadores de evolução da rede e a definição dos parâmetros que melhor caracterizam uma coleção biológica. Após esse estudo foram elaboradas especificações técnicas para cada indicador desejado (basicamente conjunto de dados e produtos desejados). Com base nessas especificações, os aplicativos foram desenvolvidos pela equipe do CRIA em linguagem Perl e os gráficos são produzidos utilizando o software FusionCharts.

Esse trabalho é dinâmico, uma vez que a interação com os provedores de dados é contínua. Outra fonte de ideias são as demais infraestruturas de dados sobre biodiversidade como o GBIF.

---

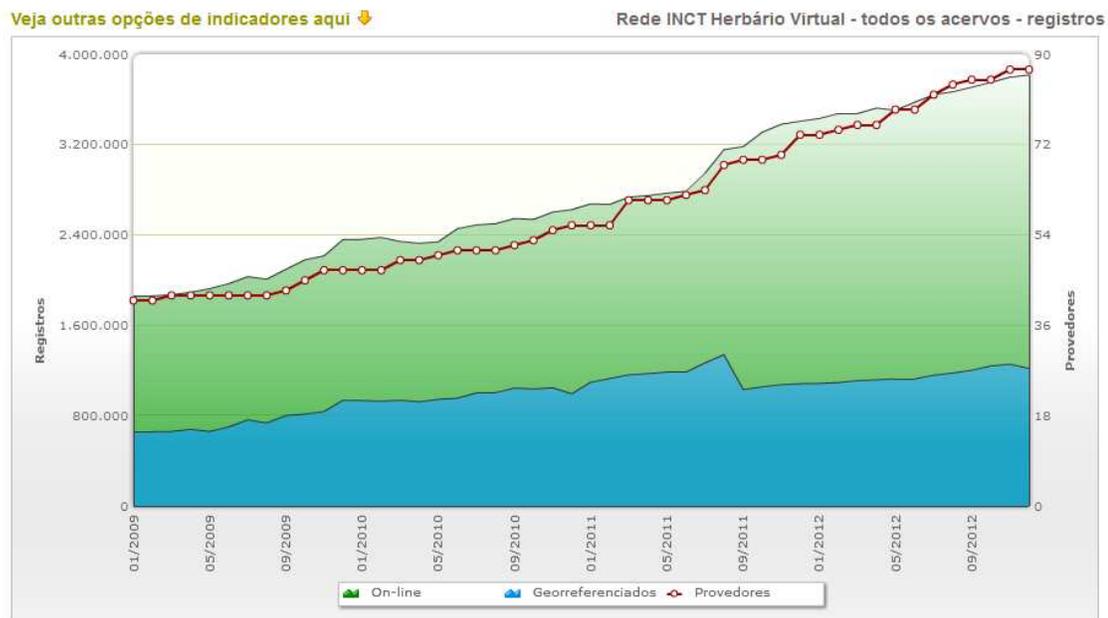
<sup>86</sup> Metadados são dados sobre os dados. No caso são o acrônimo e localidade de cada coleção, software utilizado e o tamanho do acervo

<sup>87</sup> A maioria desses documentos de apoio não foi publicada, mas os originais estão disponíveis on-line no endereço <http://www.cria.org.br/cgee/col>

#### 4.1.1. Monitoramento da E-Infraestrutura

A seguir são apresentados alguns exemplos de interesse para o monitoramento de redes, projetos e provedores. Todas as figuras apresentadas são imagens da tela do computador.

Além de poder visualizar os dados segundo diferentes critérios para toda a rede *speciesLink*, o usuário pode analisar os dados por projeto, como, por exemplo, o INCT Herbário Virtual da Flora e dos Fungos. Como os gráficos são produzidos dinamicamente essa ferramenta pode ser usada para monitorar o andamento do projeto. Um exemplo é apresentado na Figura 4.1.



Histórico do envio e retirada de dados da rede. São apresentadas as médias mensais, tanto do número total de registros on-line, como também do número de registros georreferenciados. A linha vermelha traz o número mensal de provedores de dados (coleções biológicas ou de dados).

Atualizado em 27/12/12 04:12

Fonte: ALCB, ASE, BHCB, BOTU, CARTSA, CEPEC, CGMS, CNMT, CRI, CVRD, EAC, EAFM, EAN, ESA, F\_Botany\_BR, FLOR, FUEL, Funed-Pol, FURB, HAS, HCF, HERBAM, HFSL, HFSL-Fungos, HPL, HRCB, HST, HTSA, HUCPE, HUICS, HUEFS, HUEG, HUEM, HUESB, HUESC, HUFU, HUTO, HVASF, IAC, ICN, INPA-Carpoteca, INPA-Fungos, INPA-Herbario, INPAw, IPA, IRAI, JBRJ\_RB, JBRJ\_RBdna, JBRJ\_RBw, JPB, MAC, MBM, MBML-Herbario, MFS, MIRR, MNHN\_P\_PC\_BR, MOBOT\_BR, MOSS, MPUC, NMNH-Botany\_BR, NYBG\_BR, OUPR, PEUFR, R, RON, R-Tipos, SLUI, SP, SP-Algae, SP-Bryophyta, SPF, SPF-Algae, SPFW, SPSF, TEPB, UB, UEC, UFACPZ, UFG, UFP, UFP-Carpoteca, UFRN, UFRN-Fungos, UPCB, URM, VIES, XILO TSA

Figura 4.1. Diagrama *on-line* com a evolução do número de registros e coleções integrantes do INCT Herbário Virtual da Flora e dos Fungos. Fonte: rede *speciesLink* (2012)

O diagrama apresenta a evolução dos dados ao longo do projeto, entre janeiro de 2009 e dezembro de 2012. Pode-se rapidamente visualizar que o número de registros *on-line* saltou de 1,8 milhão para 3,8 milhões e o número de coleções participantes, de 41 para 87. Pode-se comparar esses resultados com as metas propostas para avaliar o andamento do projeto.

Esse é apenas um dos quatorze gráficos produzidos por aplicativos e disponíveis *on-line*. Esses aplicativos e relatórios poderiam ser estudados e adaptados pelas agências de fomento para

serem utilizados como indicadores de acompanhamento de projetos. Outros exemplos são apresentados a seguir.

O gráfico *regiões brasileiras* mostra o número de registros no banco de dados da rede *speciesLink* por região geográfica do Brasil, para todos os registros da rede ou para determinada rede, grupo taxonômico ou coleção. O aplicativo também calcula e apresenta o número de registros por km<sup>2</sup>, por região. Esse número representa a densidade de coleta, parâmetro muito utilizado na botânica (PRANCE, 1977; SHEPHERD, 2003; BARBOSA e PEIXOTO, 2003). Tem por objetivo comparar o nível de conhecimento por região e oferecer mais um elemento de análise para auxiliar os pesquisadores e tomadores de decisão na definição das regiões prioritárias para novas coletas ou para a inclusão de novos dados na infraestrutura. A Figura 4.2 mostra a densidade de coleta para plantas por região, baseada nos dados disponíveis *on-line*.

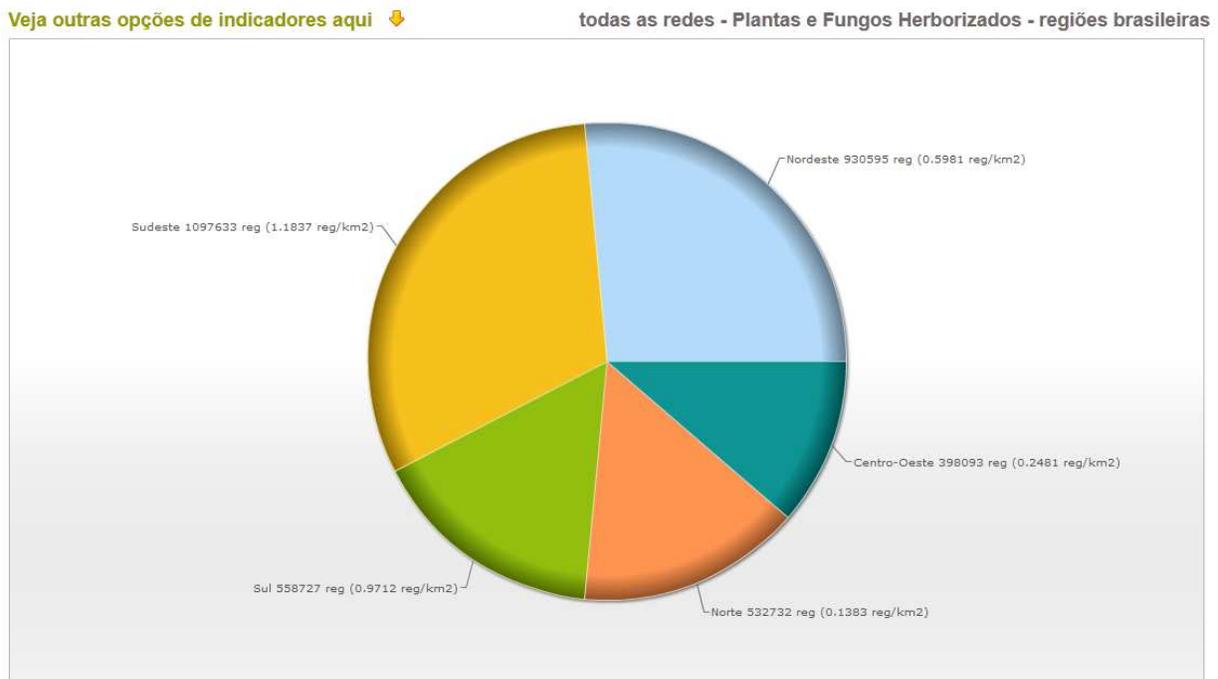


Figura 4.2. Índice de espécimes por km<sup>2</sup>. Fonte: Rede *speciesLink* (2012)

Nesse exemplo, em números absolutos, a região com mais registros *on-line* é o Sudeste, seguido pelo Nordeste, Sul, Norte e Centro-Oeste. Mas, em relação à densidade de coleta a ordem, é Sudeste, Sul, Nordeste, Centro-Oeste e Norte. Com a visualização dos dados *on-line*, cabe ao gestor apoiar a realização de novas coletas nas regiões menos coletadas ou a digitação e integração de novos dados à rede.

Outro aplicativo apresenta o número de coleções e subcoleções por estado brasileiro (Figura 4.3). O objetivo é dar maior visibilidade ao grau de regionalização da rede e identificar os estados onde deve haver um esforço maior para integrar novos acervos ou ainda identificar os estados que necessitam de mais investimentos em coleções biológicas. Trata-se de mais um elemento importante na definição de programas de apoio a coleções biológicas pelo poder público.

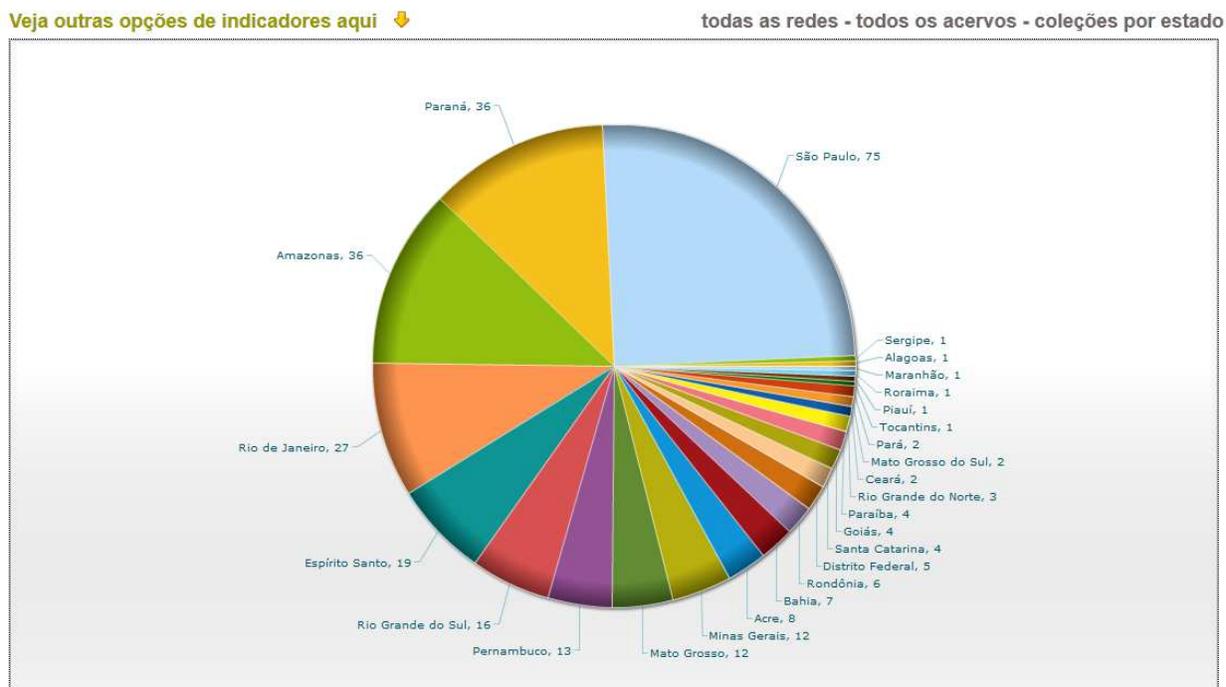


Figura 4.3. Aplicativo que mostra o número de coleções e subcoleções por estado brasileiro participante da rede *speciesLink*. Fonte: Rede *speciesLink* (2012)

No escopo do projeto INCT Herbário Virtual da Flora e dos Fungos, foi desenvolvido o indicador *índice de atualização*. Um aplicativo classifica as coleções em três categorias: as que atualizaram os seus dados há mais de um ano; entre seis meses e um ano; e há menos de seis meses. Esse indicador apresenta o índice de atualização para toda a rede ou para determinado grupo de provedores de dados. O que significa um “bom” índice de atualização não foi estabelecido. É evidente que a meta é ter todos os provedores atualizando seus dados constantemente, ou seja, todos na categoria *há menos de seis meses*. Assim, esse parâmetro por enquanto é comparativo. Se considerarmos como bom, um elevado índice de coleções que atualizaram os seus dados nos últimos 12 meses, o índice de atualização para os grandes grupos

é: Animais 58%, Plantas e Fungos Herborizados 75% e Microorganismos 67%<sup>88</sup>. Se avaliarmos só os acervos nacionais participantes do projeto INCT, que dispõem de recursos para digitação, esse índice sobe para 81%. Esses números mostram a importância dos projetos que fomentam a digitação dos dados e, mais importante, a articulação das coleções, a definição de estratégias e o acompanhamento dos resultados.

O CRIA utiliza essa informação para entrar em contato com as coleções que não atualizaram seus dados há mais de um ano, a fim de verificar se estão com algum problema técnico. O Comitê Gestor do INCT utiliza essa informação para avaliar os pedidos de renovação de bolsas para digitação de dados. A frequência na atualização dos dados do acervo *on-line* poderia ser mais um elemento na avaliação de projetos de apoio às coleções ou até do trabalho desenvolvido pelo curador.

Em fevereiro de 2011, foi desenvolvido o aplicativo que mostra a dependência de cada estado brasileiro nos dados mantidos por instituições de outros estados<sup>89</sup>. A tabela é gerada automaticamente no dia 02 de cada mês. Os dados apresentados na Tabela 4.1, atualizada em 01/12/2012, indicam, por exemplo, que o estado da Bahia possui sete coleções integradas à rede *speciesLink*, contribuindo com 8,36% dos dados totais da rede. É o quarto estado do país que mais contribui com dados para a rede, mas 43% dos dados das amostras coletadas na Bahia estão depositados em coleções de outros estados ou do exterior. Significa que o estado da Bahia, em dezembro de 2012, tinha uma dependência de 43% em dados mantidos e disponibilizados por coleções de outros estados ou do exterior. Esses valores mostram a importância do trabalho em rede.

---

<sup>88</sup> Dados de setembro de 2012

<sup>89</sup> Disponível no endereço <http://smlink.cria.org.br/indicators/stateGraph>

Tabela 4.1. Contribuição vs dependência de registros por estado.

UF [nCol]	AC	AL	AM	AP	BA	CE	DF	ES	GO	MA	MG	MS	MT	PA	PB	PE	PI	PR	RJ	RN	RO	RR	RS	SC	SE	SP	TO	Total	Cont(%)
AC [7]	11723	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	11724	0.24%
AL [1]	0	19522	38	20	1015	138	58	30	65	38	71	8	57	335	163	892	114	29	51	723	6	15	12	22	626	371	67	24486	0.50%
AM [35]	12182	50	172571	1801	2358	341	1381	491	1707	1548	1626	340	13032	27924	173	645	213	1934	1082	55	16363	19692	1132	152	8	915	692	280408	5.71%
AP [0]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00%
BA [7]	143	2430	812	195	325094	6530	3285	5712	5945	1544	13573	1369	2434	1388	2231	7398	4613	6303	3960	1691	194	178	1487	1056	1856	7618	1397	410436	8.36%
CE [2]	5	481	44	63	2410	35844	380	64	435	1899	244	159	163	661	902	1165	3538	161	213	1441	31	90	328	21	191	200	56	51189	1.04%
DF [4]	77	101	969	196	9768	119	45820	295	30215	1674	21094	2613	11251	1341	67	436	1134	3529	719	164	103	381	200	157	214	6523	3452	142612	2.91%
ES [18]	275	411	1324	59	9539	54	325	211932	974	440	7167	207	769	1636	105	238	109	1054	3087	1197	339	10	2308	305	261	12007	252	256384	5.22%
GO [3]	2	0	23	3	102	15	469	5	17669	46	401	111	121	52	0	10	6	75	33	2	13	9	67	31	0	124	770	20159	0.41%
MA [1]	0	0	2	0	0	0	0	0	0	838	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	840	0.02%
MG [11]	386	181	1154	299	2750	613	1417	3825	3405	364	150527	1593	1445	1644	55	549	416	4169	2260	137	545	204	2234	1883	58	7561	440	190114	3.87%
MS [1]	5	0	30	3	82	5	37	19	331	35	247	23076	912	47	1	10	6	652	11	3	0	23	351	54	0	290	6	26236	0.53%
MT [11]	1	0	136	0	4	20	0	29	5	8	103	807	27388	969	1	0	17	3	30	0	217	0	87	57	0	194	7	30083	0.61%
PA [2]	0	0	23	2	0	0	0	0	0	56	0	0	0	970	0	0	0	0	0	0	2	0	0	0	0	0	0	1053	0.02%
PB [4]	15	1116	402	1	2887	589	325	172	389	61	178	1	289	70	49977	4990	308	100	135	2675	85	4	215	19	1038	831	43	66915	1.36%
PE [12]	405	9814	9828	7167	24973	7119	1401	592	1200	13457	2187	279	1786	2921	8728	104562	5158	1124	987	4634	6256	1621	1161	291	9626	3186	752	231215	4.71%
PI [1]	16	426	21	2	523	285	483	31	453	302	212	41	43	61	19	246	6919	47	9	254	7	7	0	1	36	201	70	10715	0.22%
PR [34]	879	441	11228	438	22321	695	3117	13454	9013	651	30302	10676	15008	3381	493	1585	213	330646	10971	678	3968	361	14683	27313	257	25352	1475	539599	10.99%
RJ [24]	7820	1413	27798	2893	44662	5560	11314	26477	20975	2958	68485	5311	20041	25023	2441	7521	3109	16468	157037	1353	4014	2034	6549	14408	1578	29638	4029	520909	10.61%
RN [3]	0	560	7	1	1184	3094	261	11	82	336	206	23	30	13	242	1139	515	152	7	20196	8	94	64	24	191	50	81	28571	0.58%
RO [6]	3	49	1749	0	45	4	4	0	18	1	92	0	526	38	1	32	7	3	4	0	12190	160	1	0	2	0	17	14946	0.30%
RR [1]	51	0	597	0	28	0	1	3	35	0	42	4	3	338	0	1	16	0	0	0	2	5779	5	0	0	5	0	6910	0.14%
RS [16]	1360	509	4756	519	1729	501	230	934	4553	276	3003	608	6257	3534	145	968	176	3739	870	335	1764	542	162801	21240	52	3216	248	224865	4.58%
SC [3]	102	30	310	60	1282	40	305	186	367	36	1309	467	375	484	73	185	8	5934	603	81	76	850	8100	65578	8	1053	39	87941	1.79%
SE [1]	0	491	1	0	1002	255	161	2	38	83	187	16	86	13	249	943	124	212	3	217	4	58	7	19	19696	236	36	24139	0.49%
SP [74]	6131	3642	59789	4134	44960	4966	8673	24084	22555	9288	95804	20058	42659	29815	3461	19261	4235	41475	28035	5046	11537	3563	27542	23280	3397	614847	6484	1168721	23.81%
TO [1]	0	0	0	0	1	0	33	0	42	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	860	942	0.02%
EX [8]	24782	1311	68972	8073	70063	2896	22812	8092	34317	9804	62644	5719	22436	39171	1922	5314	1344	32893	28776	455	9193	6991	15619	19208	624	29616	2920	535967	10.92%
Total	66363	42978	362584	25929	568782	69683	102292	296440	154788	45743	459710	73486	167111	141829	71449	158090	32298	450702	238883	41337	66918	42666	244953	175119	39719	744034	24193	4908079	100%
Dep(%)	82%	55%	52%	100%	43%	49%	55%	29%	89%	98%	67%	69%	84%	99%	30%	34%	79%	27%	34%	51%	82%	86%	34%	63%	50%	17%	96%		

Fonte: rede *speciesLink* (01/12/2012)

Hoje existem 14 gráficos disponíveis para cada rede, tipo de acervo e ainda para cada provedor de dados. Apenas para ilustrar sua importância como ferramenta na gestão pública, é apresentado mais um parâmetro, a *informatização*. A Figura 4.4 mostra o número de registros *on-line* georreferenciados e não georreferenciados e o número de registros *off-line* para os herbários brasileiros participantes do INCT. O número de registros *off-line* é obtido pelo número total de exsicatas informado por cada herbário, menos o número de registros *on-line*.

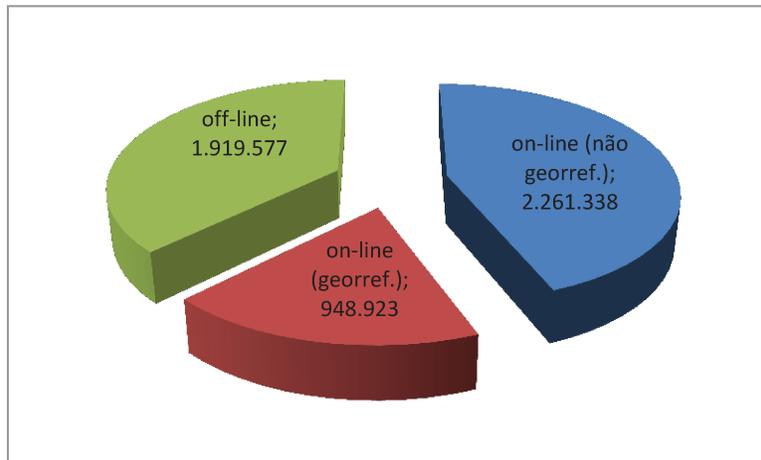


Figura 4.4. Indicador de informatização dos acervos dos provedores nacionais do INCT Herbário Virtual da Flora e dos Fungos. Fonte: Rede *speciesLink* (2012)

O trabalho apresentado por Peixoto e Lima (2005) indica um custo médio de R\$ 1,20 para a digitação de uma etiqueta de herbário (sem considerar custos indiretos e de infraestrutura). Utilizando esse número como parâmetro, multiplicando-o pelo número de registros *off-line* indicado na Figura 4.4, tem-se que seria necessário investir mais de 2,3 milhões de reais só na digitação dos dados de material já depositado nos herbários participantes da rede. Esse valor precisaria ser atualizado, mas é possível rapidamente obter o número de amostras que ainda não foram informatizadas e, dessa forma, quantificar uma demanda concreta por recursos para a digitação dos dados.

#### 4.1.2. Valorização dos Serviços Científicos de Curadores e Taxonomistas

Monitorar a evolução das atividades das coleções biológicas como provedores de dados à infraestrutura representa uma ferramenta importante para identificar dificuldades e definir as ações para a resolução de problemas. Mas, representa também uma oportunidade para qualificar, quantificar e valorizar esse trabalho. Esse trabalho é dinâmico, uma vez que a atualização dos

dados é contínua. Todos os gráficos apresentados a seguir são produzidos dinamicamente por aplicativos e disponibilizados *on-line*.

Existe um conjunto de nove aplicativos para acompanhar o processo de compartilhamento dos dados e para caracterizar os acervos. Além da evolução do número de registros *on-line* e do grau de informatização, foram desenvolvidos aplicativos para visualizar a representatividade das famílias<sup>90</sup>, os principais coletores<sup>91</sup>, a distribuição do número de registros por ano de coleta, o número de registros por país da coleta, o número de registros por estado brasileiro onde a coleta foi realizada, o material-tipo e o número de registros por região brasileira<sup>92</sup>.

Os indicadores apresentados pela e-infraestrutura caracterizam o acervo com base apenas nos dados disponíveis *on-line*. Acervos 100% digitalizados, ou próximo a isso, têm nos indicadores uma importante ferramenta de descrição do acervo, enquanto o perfil dos acervos em processo de digitação, muitas vezes reflete a estratégia adotada para a digitação dos dados (por família, por tipo etc.). A Figura 4.5 mostra dois indicadores para uma determinada coleção: o compartilhamento *on-line* dos dados; e, o número de amostras coletadas por ano.

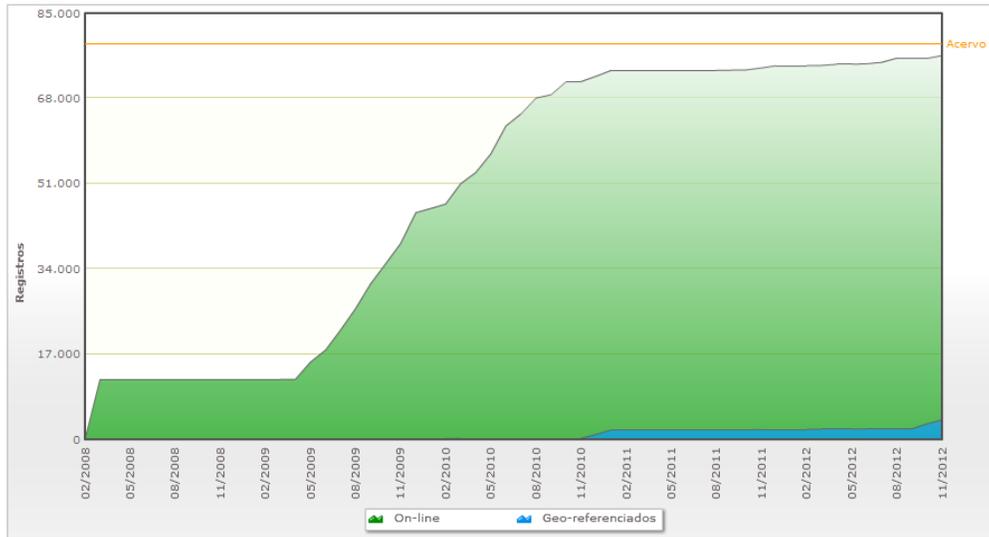
---

<sup>90</sup> Na hierarquia taxonômica, família agrupa um conjunto de gêneros.

<sup>91</sup> *Coletor* é o técnico responsável pela coleta de material no campo

<sup>92</sup> Um exemplo de perfil do acervo *on-line* está disponível no endereço <http://splink.cria.org.br/manager/detail?setlang=pt&resource=BHCB>

histórico de movimentação de dados



Histórico do envio e retirada de dados da rede. São apresentadas as médias mensais, tanto do número total de registros on-line, como também do número de registros georreferenciados. A linha vermelha traz o número mensal de provedores de dados (coleções biológicas ou de dados).

Atualizado em 18/11/2012 15:26

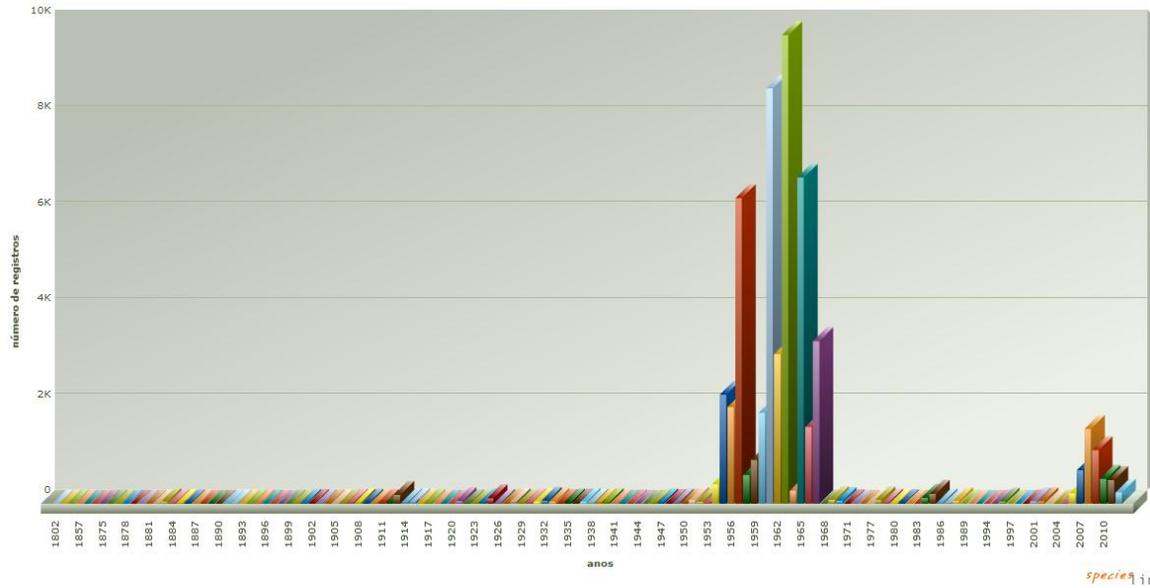


Figura 4.5. Gráficos *on-line* sobre o histórico de movimentação de dados e do número de registros por ano de coleta de uma mesma coleção biológica. Fonte: rede *speciesLink* (2012)

O exemplo da Figura 4.5 mostra claramente muita atividade nas décadas de 1950 e 1960 em relação à coleta e ampliação do acervo e, mais recentemente, a partir de 2007, com novas coletas e com a digitação e disponibilização dos dados *on-line*. Esses gráficos mostram uma coleção dinâmica com um trabalho de curadoria efetivo e que provavelmente, a partir de 2007, recebeu recursos para novos desenvolvimentos.

O sistema também produz um relatório sobre a qualidade dos dados, indicando possíveis erros de digitação ou incompletudes e monitora o trabalho da curadoria na correção de seus dados. Produz também relatórios de acesso aos dados por usuários da rede. O conjunto desses aplicativos revela o trabalho do curador. Um acervo bem curado, do ponto de vista da e-infraestrutura, está on-line, as amostras estão identificadas, os nomes são válidos, as informações estão completas e os dados estão georreferenciados. A evolução contínua do número de registros disponibilizados e da qualidade e completude dos dados retrata uma coleção ativa e bem curada.

Esses relatórios poderiam gerar um índice de produtividade para ser usado na avaliação do trabalho do curador responsável pelos dados servidos pela coleção à infraestrutura.

O mesmo pode ser dito quanto ao trabalho do taxonomista. Se não, vejamos. Uma abordagem multidimensional na avaliação de programas de C&T em rede foi objeto de tese que apresentou um estudo sobre a avaliação do Programa Biota-Fapesp (CASTRO, 2011). Algumas análises e conclusões apresentadas chamam a atenção em relação a possíveis usos dos dados e aplicativos de e-infraestruturas. Por exemplo, o estudo realizou um levantamento junto aos pesquisadores que apontaram a falta de apoio à taxonomia como o principal ponto negativo do Programa. A crítica é em relação ao não reconhecimento da atividade de identificação de espécies e da geração de listas de espécies por taxonomistas como produção científica. Essa crítica é especialmente importante quando mundialmente se reconhece que a taxonomia é um fator determinante na geração de conhecimento sobre biodiversidade. Há três décadas o número de taxonomistas em atividade no mundo está declinando enquanto a demanda por serviços taxonômicos está aumentando (KIM e BYRNE, 2006).

A e-infraestrutura apresenta ferramentas que poderiam ser utilizadas como indicadores de produtividade dos taxonomistas que prestam serviços de identificação de material às coleções biológicas. A valorização desse trabalho certamente estimularia essa atividade, melhorando o número e a qualidade dos dados disponíveis *on-line*.

Outras ferramentas desenvolvidas com recursos do projeto INCT - Herbário Virtual da Flora e dos Fungos, viabilizam, para algumas amostras, a prestação desse serviço à distância. Servidores de imagens permitem às coleções biológicas disponibilizarem *on-line* as imagens de suas amostras com os dados textuais associados. O herbário do Jardim Botânico de Nova Iorque, entre outros, disponibiliza as imagens de material coletado no Brasil e identificado apenas até família. Especialistas poderão acessar os dados sobre a coleta do material e visualizar a imagem

utilizando os serviços *on-line* que permitem fazer zoom para visualizar mais detalhes, fazer medições ou até alterar a cor da imagem (Figura 4.6).

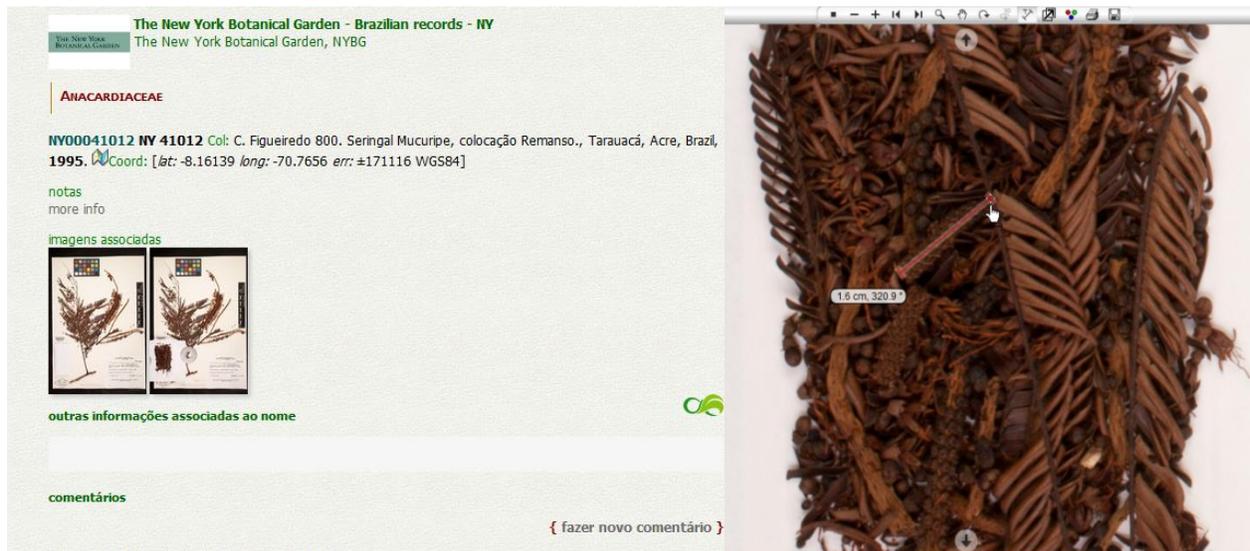


Figura 4.6. Ficha de material não identificado e o acesso à imagem. Fonte: Rede *speciesLink* (2012)

O especialista pode analisar a imagem e enviar a sua contribuição, ou o seu comentário, à coleção, por exemplo, com a determinação do nome da espécie. É possível também comparar imagens de diferentes amostras (p.ex. material-tipo com material não identificado) para facilitar o processo de identificação do material à distância.

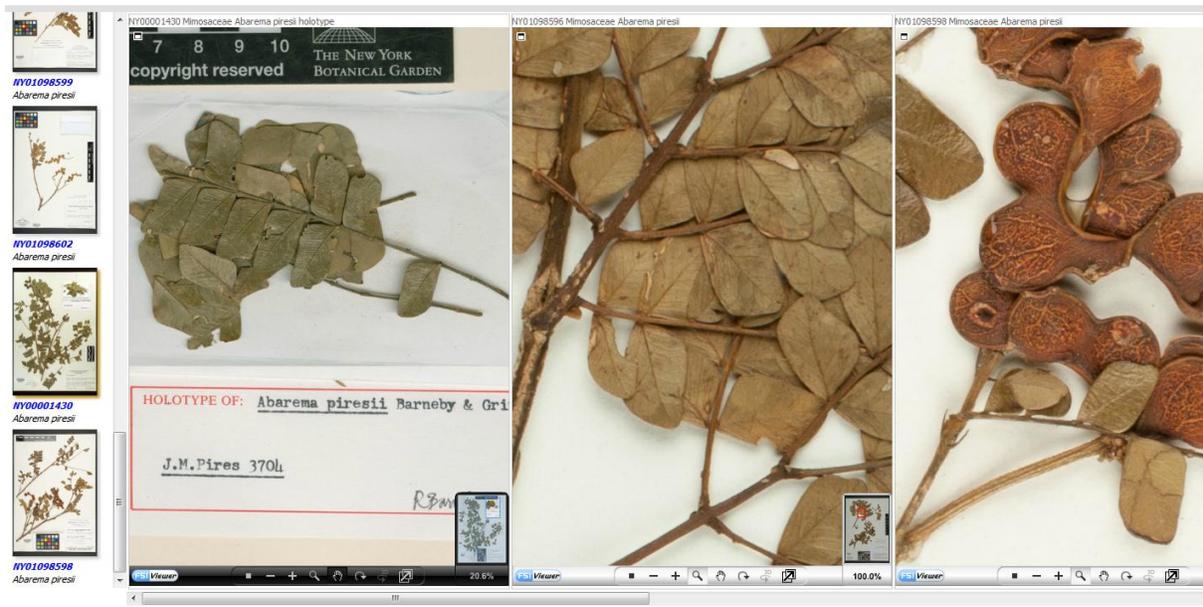


Figura 4.7. Ferramenta para comparar imagens de diferentes amostras. Fonte: Rede *speciesLink* (2012)

Essas ferramentas, e inúmeras outras que estão disponíveis ou serão desenvolvidas, permitem um alcance muito maior dos serviços dos taxonomistas que precisam ser reconhecidos e valorizados.

Os gráficos apresentados na Figura 4.8 foram produzidos dinamicamente. Foi realizada uma busca pelo nome de um taxonomista no campo “determinador” (pessoa responsável pela identificação da amostra) e foram produzidos dois gráficos: (1) herbários com amostras determinadas pelo taxonomista e (2) número de determinações por ano do taxonomista. Ambos ilustram o trabalho realizado e poderiam ser usados como critério de avaliação e valorização do serviço prestado, compondo um índice de produtividade do taxonomista.

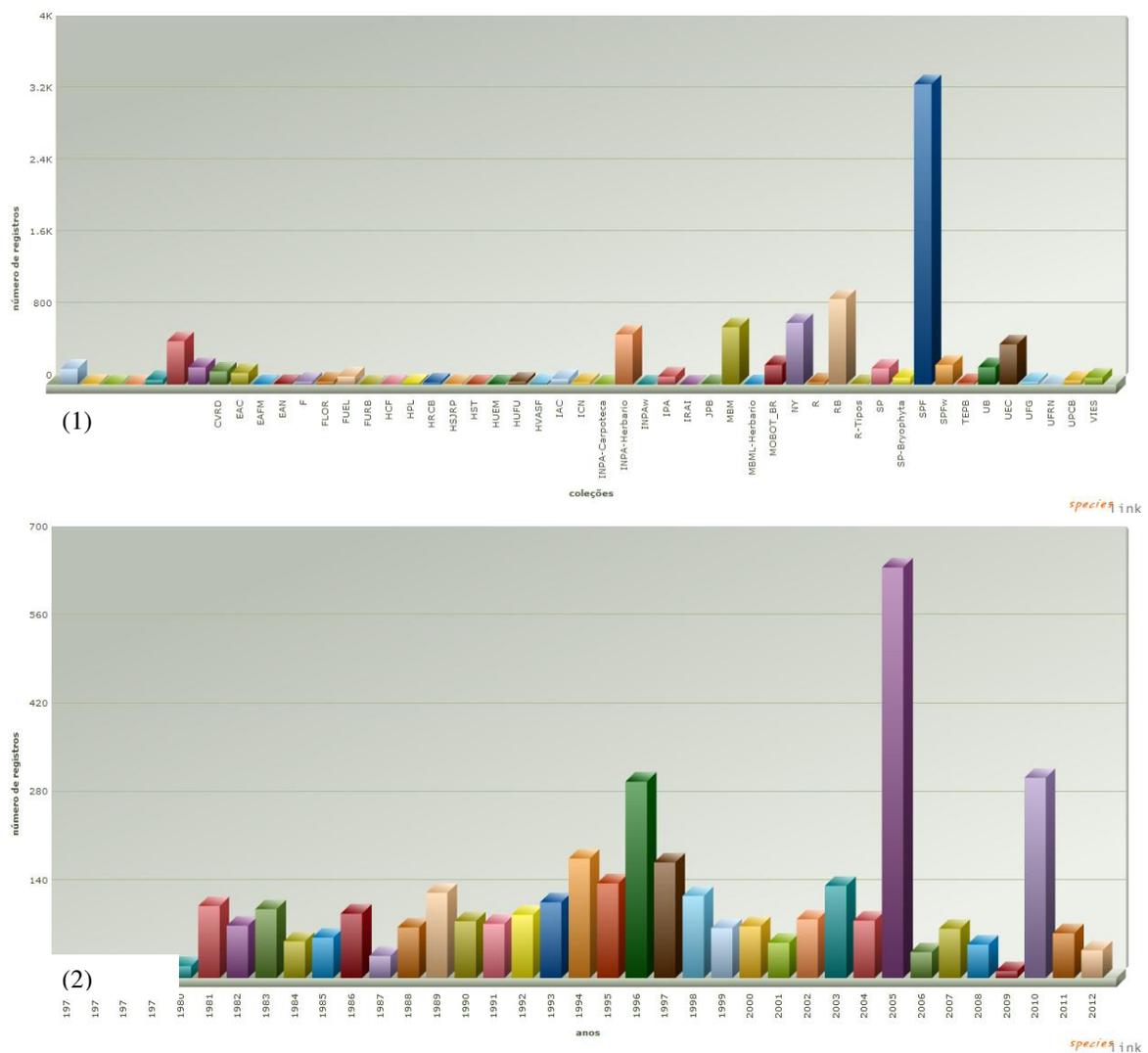


Figura 4.8. Gráficos produzidos dinamicamente indicando (1) os herbários que usaram os serviços do taxonomista e (2) o número de determinações realizadas pelo especialista por ano. Fonte: rede *speciesLink* (2012)

O uso da e-infraestrutura para a geração de indicadores de produtividade do taxonomista como prestador de serviços científicos deve ser estudado por especialistas para a sua implementação. É importante lembrar que todos os dados da rede speciesLink estão disponíveis via serviços web. Assim, a criação dos indicadores pelo poder público, em princípio, não depende do desenvolvimento de novos aplicativos pela rede.

#### 4.2. USO DA E-INFRAESTRUTURA NA DEFINIÇÃO DE ESTRATÉGIAS PARA PREENCHER LACUNAS DE DADOS E CONHECIMENTO EM BIODIVERSIDADE

Em seu trabalho sobre a avaliação da inovação em políticas públicas, CASTRO (2011) chama a atenção para o uso pela Secretaria Estadual do Meio Ambiente do mapa de áreas prioritárias para conservação e recuperação da biodiversidade que sintetiza informações de especialistas a partir dos dados disponibilizados na rede *speciesLink*. Indica ser esse um elo entre a caracterização e a conservação da biodiversidade e políticas públicas e *um produto que suplanta os resultados de um único projeto podendo ser considerado um produto transversal aos projetos do Programa*. A interação entre agentes do governo, a comunidade científica e os mantenedores de e-infraestruturas pode propiciar o desenvolvimento de produtos úteis e utilizáveis para processos de tomada de decisão e formulação de estratégias e políticas para biodiversidade.

O trabalho apresentado a seguir procurou responder a uma pergunta do projeto INCT Herbário Virtual da Flora e dos Fungos de como a infraestrutura de dados poderia ser usada para definir os grupos taxonômicos ou áreas geográficas prioritários para orientar novas coletas.

A simples disponibilidade dos dados *on-line* e a possibilidade de plotá-los em uma base cartográfica mostra possíveis lacunas de conhecimento geográfico. A Figura 4.9 mostra os registros de plantas com coordenadas geográficas que ocorrem no Brasil e aqueles que ocorrem no Estado da Bahia.

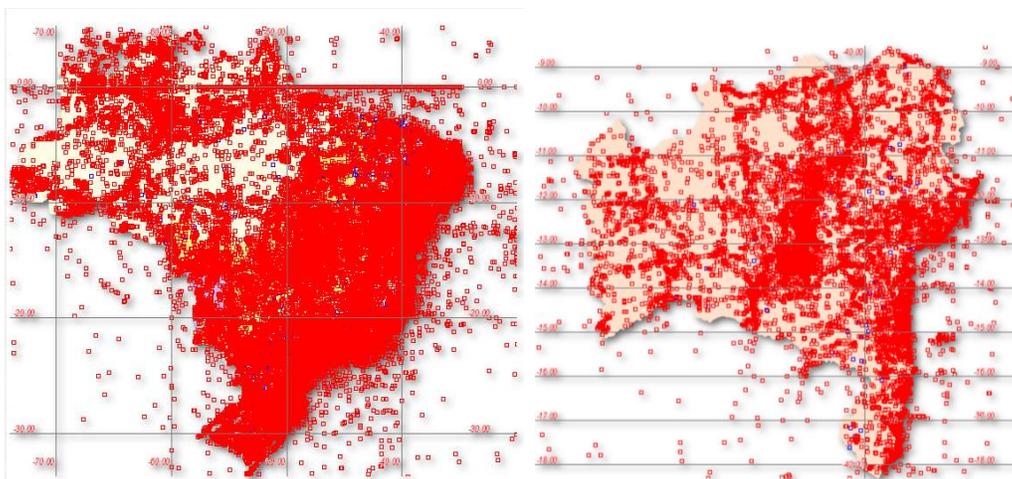


Figura 4.9. Visualização dos pontos de ocorrência de espécies de plantas do INCT Herbário Virtual da Flora e dos Fungos no Brasil e no Estado da Bahia. Fonte: Rede *speciesLink* (2012)

Percebe-se claramente que há necessidade de definição de estratégias, políticas e ações em diferentes escalas, do local ao nacional, para ampliar a base de conhecimento sobre biodiversidade. Mas o fato de existir a e-infraestrutura, com dados e ferramentas de acesso livre e aberto, disponíveis a qualquer pessoa interessada via *web*, permite essa visão multi-escala.

Mas não basta plotar pontos de ocorrência de espécies em uma base cartográfica. É necessário qualificar essa informação, procurando responder o que coletar e onde.

Em 2010, o CNPq lançou o programa SISBIOTA Brasil – Sistema Nacional de Pesquisa em Biodiversidade. Um de seus eixos temáticos é a ampliação do conhecimento da biodiversidade. O primeiro edital lançado em 2010 apresentou como uma das três chamadas: *Sínteses e Lacunas do Conhecimento da Biodiversidade Brasileira*. O INCT Herbário Virtual da Flora e dos Fungos apresentou proposta que foi aprovada (Processo CNPq nº 563342/2010-2). A proposta incluía a elaboração de um relatório de lacunas de informação da flora brasileira em herbários. A ideia original, dentre outras ações do projeto, era divulgar os nomes da Lista de Espécies da Flora do Brasil que não possuíam nenhum registro na rede *speciesLink*.

Com uma visão também voltada ao uso dos dados para o processo de tomada de decisão, foi decidido que, além do nome, seria importante considerar a coordenada geográfica na análise. Assim, o objetivo inicial do aplicativo foi ampliado graças ao estudo realizado no âmbito desta tese.

O processo de concepção do aplicativo teve início com um estudo da estrutura e qualidade dos dados sobre ocorrência de espécies de plantas no Brasil disponíveis na rede *speciesLink*. Os dois elementos fundamentais dos dados são o nome científico e as coordenadas geográficas.

Muitos registros apresentam erros de grafia e falta de atualização da nomenclatura. Os dados de plantas da rede *speciesLink* foram “filtrados” com os dados da Lista de Espécies da Flora do Brasil<sup>93</sup>, que em setembro de 2012 tinha 43.278 nomes aceitos de espécies e 29.146 sinônimos. O uso da lista como referência de certa forma “limpa” os dados da rede *speciesLink*, uma vez que, ao usar a lista como filtro, só são considerados os nomes válidos e seus sinônimos. A busca é fonética, para também recuperar a maioria dos registros com erros de grafia, mas os registros com nomes que não constam da Lista Oficial são descartados. Isso significa que podem estar sendo descartados bons dados cujos sinônimos ainda não constam da Lista do Brasil.

Com relação às coordenadas geográficas, somente 31% dos dados de plantas *on-line* são georreferenciados na origem<sup>94</sup>. Para aumentar o número de registros georreferenciados, em 2009 foi desenvolvido um aplicativo que atribui coordenadas geográficas ao dado com base no município informado. O uso do aplicativo eleva o percentual de dados georreferenciados de 31 para 73%, mas a sua utilidade depende sempre da precisão requerida pelo usuário. Ambas as coordenadas, na origem e por aplicativo, foram consideradas pelo sistema desenvolvido.

Novamente pensando no uso do sistema por um público mais amplo, foi considerado estratégico agregar mais valor ao produto a ser gerado pelo aplicativo, associando ao nome de cada espécie informações sobre possível endemismo<sup>95</sup> (Lista de Espécies da Flora do Brasil) e status de conservação<sup>96</sup> (Listas do Ministério do Meio Ambiente e da Fundação Biodiversitas).

Em relação ao status de conservação, a lista oficial é a Lista Nacional das Espécies da Flora Brasileira Ameaçadas de Extinção do MMA<sup>97</sup>, mas a Lista da Flora Brasileira Ameaçada de Extinção, resultante do estudo realizado no âmbito do Convênio IBAMA/Fundação Biodiversitas (nº. 46/2002), traz os critérios de avaliação de acordo com a IUCN<sup>98</sup> (IUCN, 2001). A lista do MMA traz 472 espécies em seu Anexo I (espécies prioritárias para políticas públicas) e 1079 espécies em seu Anexo II (espécies deficientes em dados), enquanto a Lista da Biodiversitas traz 1.495 espécies classificadas como ameaçadas de extinção e 2.513 classificadas como deficientes

---

<sup>93</sup> Disponível no endereço <http://floradobrasil.jbrj.gov.br/2012>

<sup>94</sup> Dados georreferenciados na origem são aqueles cujas coordenadas geográficas são preenchidas pelo provedor de dados.

<sup>95</sup> O produto gerado indica se a espécie só ocorre no Brasil

<sup>96</sup> O produto gerado indica se a espécie está na lista de espécies ameaçadas de extinção

<sup>97</sup> Instrução Normativa MMA nº 06, de 23 de setembro de 2008

([http://www.mma.gov.br/estruturas/179/\\_arquivos/179\\_05122008033615.pdf](http://www.mma.gov.br/estruturas/179/_arquivos/179_05122008033615.pdf))

<sup>98</sup> A lista mundial de espécies ameaçadas é coordenada pela IUCN (International Union for Conservation of Nature) e está disponível no endereço [www.iucnredlist.org](http://www.iucnredlist.org)

em dados. Para este estudo foram incluídos os nomes das espécies dos Anexos I e II do MMA e da lista de espécies ameaçadas de extinção da Biodiversitas.

Um problema também identificado nas listas do MMA e da Biodiversitas foi a existência de nomes com erros de grafia, sinônimos e nomes de espécies sem correspondentes na Lista de Espécies da Flora do Brasil. Erros de grafia e sinonímia foram resolvidos, mas as espécies sem correspondência na Lista do Brasil não foram consideradas.

Com base neste estudo, as seguintes fontes de dados foram selecionadas para compor o sistema: (i) A rede *speciesLink* para os registros de ocorrência de plantas no Brasil; (ii) A Lista de Espécies da Flora do Brasil, 2012 para a relação de nomes válidos, seus sinônimos e dados sobre endemismo; (iii) A Lista Nacional das Espécies da Flora Ameaçadas de Extinção do MMA (Anexos I e II) e a Lista da Flora Brasileira Ameaçada da Fundação Biodiversitas para a relação de espécies ameaçadas.

Uma vez especificados todos os requerimentos - entrada de dados e produtos resultantes -, o sistema foi desenvolvido pela equipe do CRIA, sendo avaliado e aprimorado por duas bolsistas pós-doc, uma com doutorado em botânica e outra em biologia vegetal. O protótipo foi lançado para testes no dia 25/11/2011, somente para o comitê gestor do INCT Herbário Virtual. Um segundo protótipo foi lançado em abril de 2012 para um público mais amplo – os curadores participantes do Herbário Virtual, e o sistema foi oficialmente lançado para acesso público no dia 17 de julho de 2012.

O produto resultante é o sistema *on-line* denominado *Lacunas de conhecimento da flora e dos fungos do Brasil*<sup>99</sup>. Apresenta um relatório do status dos dados disponíveis na rede *speciesLink* para cada espécie citada como nome aceito na Lista de Espécies da Flora do Brasil.

O aplicativo também produz uma análise por grupo taxonômico, família ou gênero, com as espécies distribuídas em 4 categorias: sem registros (zero); entre 1 a 5 registros; entre 6 a 20 registros; e, com mais de 20 registros. Tem por objetivo rapidamente visualizar as espécies sem dados ou com insuficiência de dados. A classificação por número de registros segue a lógica da utilização de dados da ocorrência de espécies para modelagem do seu nicho ecológico e potencial distribuição geográfica, ferramenta muito usada em processos de tomada de decisão. O número de pontos de ocorrência necessários para produzir um bom modelo depende de cada espécie, mas, em geral, com até 5 pontos, obtém-se um modelo preliminar, entre 6 a 20 pontos um modelo

---

<sup>99</sup> Disponível no endereço <http://lacunas.inct.florabrasil.net>

exploratório e com mais de 20 pontos um modelo com uso potencial para a tomada de decisão. Quanto maior o número de dados de qualidade, melhor o modelo resultante (WISZ, HIJMANS, *et al.*, 2008).

Mesmo em se tratando de uma avaliação preliminar dos dados disponíveis *on-line*, no sistema *Lacunas* o usuário tem várias opções de análise. O resultado por grupo taxonômico, família ou gênero pode ser apresentado somente considerando o nome aceito ou incluindo seus sinônimos e a busca pode ser exata ou fonética. O usuário poderá ainda filtrar os dados de acordo com o tipo de coordenada geográfica (informada pela coleção ou atribuída por um aplicativo), podendo excluir as coordenadas inconsistentes<sup>100</sup> ou ainda selecionar somente coordenadas distintas.

A Tabela 4.2 mostra os resultados obtidos para o grupo das Angiospermas, com 31.718 espécies (nomes aceitos). Os resultados apresentados para os diferentes critérios de busca são o número de espécies classificadas de acordo com o número de registros disponíveis *on-line*.

Tabela 4.2. Resultado da avaliação dos dados para espécies de Angiospermas

Critérios de busca		No. de especies por categoria e % em relação ao total							
		zero	%	com 1 a 5 registros	%	com 6 a 20 registros	%	com mais de 20 registros	%
1	busca por nomes aceitos, grafia exata, com ou sem coordenadas	3.090	10	7.406	23	7.543	24	13.679	43
2	busca por nomes aceitos, grafia exata, com qualquer coordenada	4.502	14	8.070	25	7.387	23	11.759	37
3	busca por nomes aceitos, grafia exata, com coordenadas consistentes de origem	8.945	28	9.597	30	6.788	21	6.388	20
4	busca por nomes aceitos, grafia exata, com coordenadas consistentes e distintas de origem	9.019	28	10.817	34	6.684	21	5.198	16
5	busca fonética com sinônimos, com ou sem coordenadas	2.429	8	7.048	22	7.628	24	14.613	46
6	busca fonética com sinônimos, com qualquer coordenada	3.792	12	7.835	25	7.545	24	12.546	40
7	busca fonética com sinônimos, coordenadas consistentes de origem	8.244	26	9.564	30	7.053	22	6.857	22
8	busca fonética com sinônimos, coordenadas totais consistentes	4.733	15	7.597	24	7.440	24	11.948	38

<sup>100</sup> É importante enfatizar que se trata de uma avaliação preliminar dos dados disponíveis e que não considera a precisão da coordenada geográfica. Em termos das coordenadas, a ferramenta checa uma única inconsistência – se a coordenada não cai no município indicado. Um importante trabalho, que tem por objetivo avaliar os dados de espécies da família Passifloraceae para a produção de modelos, é apresentado por Giovanni *et al.* (2012), que aplica filtros muito mais restritivos às coordenadas geográficas.

Critérios de busca		No. de espécies por categoria e % em relação ao total							
		zero	%	com 1 a 5 registros	%	com 6 a 20 registros	%	com mais de 20 registros	%
9	busca fonética com sinônimos, coordenadas consistentes e distintas de origem	8.308	26	10.848	34	6.975	22	5.587	18
10	busca fonética com sinônimos, coordenadas totais consistentes e distintas	4.746	15	9.916	31	8.256	26	8.800	28

Fonte: Lacunas (setembro, 2012)

O critério de busca n.º. 4 na Tabela 4.2 apresenta os resultados para a opção mais restritiva, que é a busca no banco de dados da rede *speciesLink* apenas considerando os nomes aceitos da Lista de Espécies da Flora do Brasil, e com grafia exata, com coordenadas geográficas consistentes e distintas informadas pela coleção. Nesse caso, os dados estão taxonomicamente atualizados e com a grafia correta e o sistema só está considerando os registros cujas coordenadas são informadas pela coleção e caem no município indicado. Tem-se ainda que o sistema só está contabilizando as coordenadas distintas. Isso significa que, se existem 10 registros de uma mesma espécie com a mesma coordenada, o sistema irá contar como sendo um único registro. Usando esses critérios de busca e seleção, o número de espécies sem nenhum registro representa 28% do número total de espécies de Angiospermas e apenas 16% das espécies possuem dados suficientes para produzir modelos de distribuição de nicho ecológico com uso potencial para a tomada de decisão.

Analisando o resultado adotando como critério o número 10 onde são incluídos sinônimos e usados aplicativos de busca fonética e georreferenciamento automático por município, mas ainda são selecionados somente os registros com coordenadas consistentes e distintas, esses números passam a ser 15% das espécies sem qualquer registro *on-line* e 28% das espécies com dados suficientes para produzir modelos de distribuição de nicho ecológico com uso potencial para a tomada de decisão. Esse resultado mostra o impacto que as ferramentas podem ter no uso e análise dos dados.

Para avaliar melhor o aumento da usabilidade dos dados através do uso de aplicativos, na Tabela 4.2 o número de espécies com mais de 20 registros, ao adotarmos como critério de busca a busca pelo nome aceito e exato, é 13.679. Se utilizarmos um aplicativo que é capaz de associar os sinônimos ao nome aceito e realizar uma busca fonética, esse número passa a ser 14.613. Significa que a ferramenta (busca fonética mais a inclusão de sinônimos) aumentou o número de espécies com mais de 20 pontos em 6,8%. É importante lembrar que a geração do sistema

Lacunas parte dos dados dos nomes de espécies já filtrados pela Lista de Espécies da Flora do Brasil. Esse primeiro filtro descarta cerca de 27% dos registros de plantas da rede *speciesLink*. Uma análise dos registros não utilizados, indicou que 69% são registros identificados até família ou até gênero e 31% é material identificado, mas que não têm correspondência na Lista de Espécies da Flora do Brasil. Pode tratar-se de sinônimos ainda não incluídos na Lista ou erros de grafia que o algoritmo fonético não consegue identificar.

Adotando como critério a busca fonética dos nomes aceitos mais os sinônimos e comparando as coordenadas consistentes na origem com as coordenadas consistentes totais, ou seja, incluindo as coordenadas atribuídas pelo aplicativo (georreferenciamento por município), o número de espécies com mais de 20 pontos de ocorrência salta de 6.857 para 11.948, ou seja, um aumento de 74,2%. Se restringirmos a busca somente para as coordenadas distintas, sem o uso do aplicativo tem-se 5.587 (coordenadas consistentes, distintas na origem) e 8.800 registros com coordenadas distintas totais, ou seja, o aplicativo aumenta a oferta de dados em 57,5%.

O uso dos dados com coordenadas geográficas por município depende da precisão requerida pelo usuário. Para os casos onde essa precisão é adequada, o impacto da ferramenta é muito expressivo. Além de aumentar a usabilidade dos dados, as ferramentas também demonstram que ainda é necessário investir na correção de erros de digitação, na atualização taxonômica e no georreferenciamento dos dados.

Em 2012 o sistema *Lacunas* foi gerado em maio e setembro. A Figura 4.10 mostra os gráficos dos relatórios gerados para Angiospermas usando os critérios de busca fonética com sinônimos, coordenadas totais consistentes e distintas, em maio e em setembro de 2012.

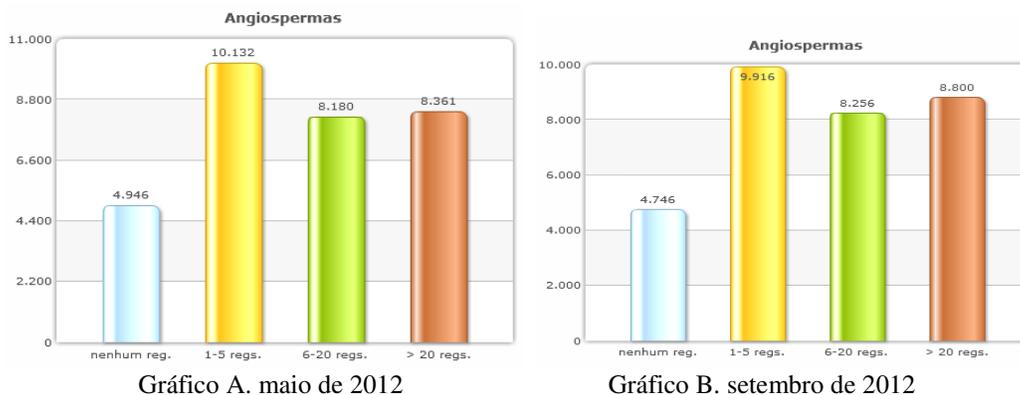


Figura 4.10. Gráficos com o número de espécies de Angiospermas agrupadas nas quatro categorias de quantidade de registros *on-line* gerados em maio (gráfico A) e em setembro (gráfico B) de 2012. Fonte: Lacunas (2012)

O número total de espécies de Angiospermas em maio de 2012 era 31.619 e em setembro 31.718. Tanto a Lista do Brasil como os dados do Herbário Virtual são atualizados constantemente e novos acervos são incorporados, alterando a quantidade e qualidade dos dados disponíveis. Significa que, para o processo de tomada de decisão, é importante sempre monitorar os dados e atualizar os diagnósticos.

Além da análise por grandes grupos como Angiospermas, são também apresentadas análises por família ou gênero sendo possível visualizar a relação das espécies e o respectivo número de registros. Permite aos especialistas por família ou gênero identificarem lacunas taxonômicas ou geográficas de dados e conhecimento.

A Figura 4.11 mostra o relatório produzido pelo aplicativo para a família Alstroemeriaceae, evidenciando as espécies endêmicas do Brasil (E da Lista do Brasil), as que estão no Anexo I (I) ou II (II) da portaria do MMA e na lista da Fundação Biodiversitas, apresentando os critérios de ameaça: EX extinta; EW extinta na natureza; CR criticamente em perigo; EN em perigo; e VU vulnerável.

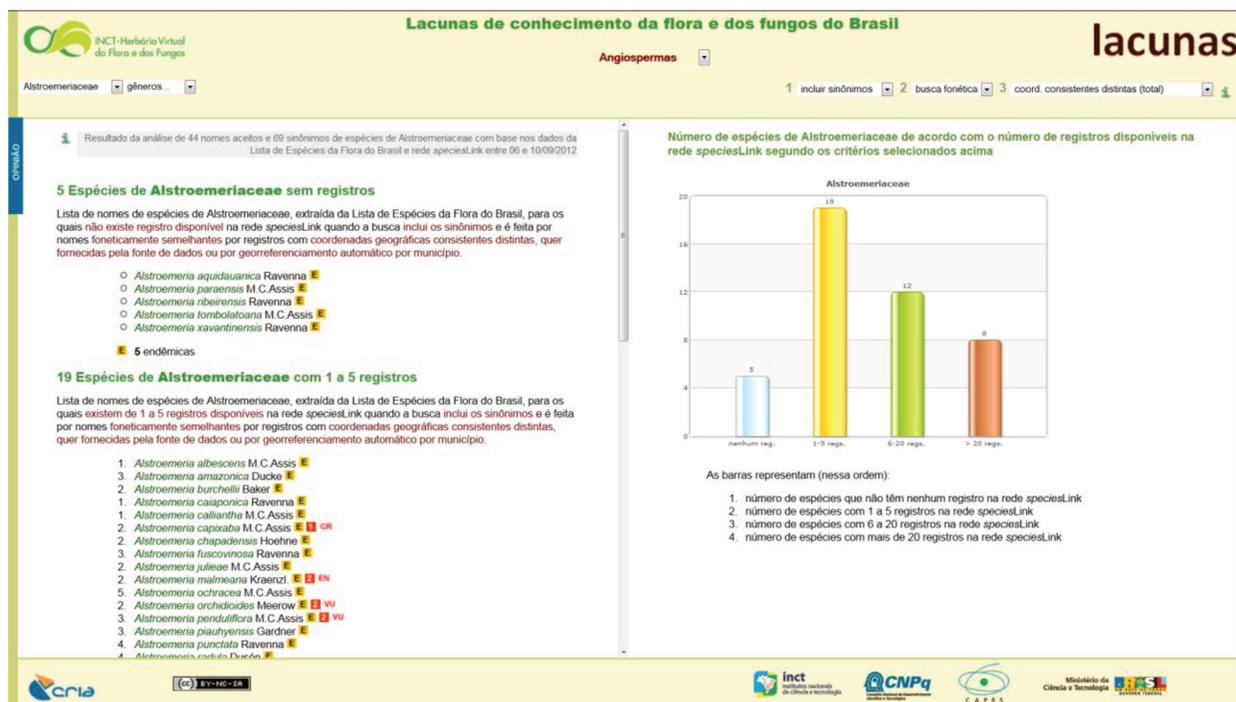


Figura 4.11. Vista parcial do relatório do status dos dados disponíveis *on-line* para as espécies da família Alstroemeriaceae. Fonte: Lacunas (2012)

O sistema *Lacunas*, também apresenta relatórios específicos para as espécies ameaçadas de extinção, incluídas nos Anexos I e II da Instrução Normativa MMA nº 06<sup>101</sup>. Os relatórios desses Anexos apresentam a relação das espécies classificadas nos mesmos grupos (sem registros, 1-5, 6-20, e >20) com base em dois critérios: busca fonética incluindo sinônimos e todos os registros (com ou sem coordenadas) e busca fonética incluindo sinônimos e coordenadas consistentes distintas (Figura 4.12).

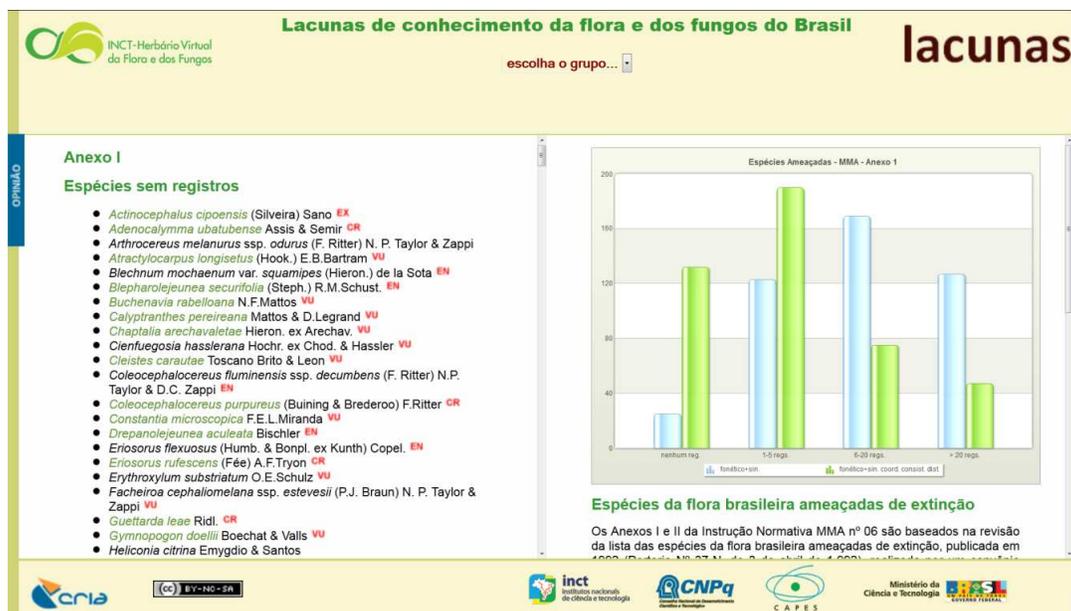


Figura 4.12. Relatório do aplicativo *Lacunas* para o Anexo I das espécies da flora brasileira ameaçadas de extinção. Fonte: Lacunas (2012)

É interessante avaliar o relatório do status dos dados disponíveis *on-line* para as espécies do Anexo II que, segundo a portaria do MMA, apresentam deficiência de dados. O relatório gerado para todas as espécies do Anexo II apresenta 38 espécies sem qualquer registro e mais 187 espécies sem registros com coordenadas geográficas; 438 espécies com 1 a 5 pontos distintos de ocorrência; 224 espécies com 6 a 20 pontos distintos de ocorrência - para as quais poder-se-ia desenvolver modelos exploratórios; e 107 espécies com mais de 20 pontos com coordenadas distintas, para as quais teoricamente podem ser desenvolvidos modelos consistentes de distribuição de espécies que são úteis para processos de tomada de decisão<sup>102</sup>.

<sup>101</sup> A Instrução Normativa nº. 6 de 23 de setembro de 2008, que reconhece como espécies da flora brasileira ameaçadas de extinção aquelas do Anexo I e com deficiência de dados aquelas do Anexo II está disponível no endereço [http://www.mma.gov.br/estruturas/179/arquivos/179\\_05122008033615.pdf](http://www.mma.gov.br/estruturas/179/arquivos/179_05122008033615.pdf)

<sup>102</sup> Dados de 10/09/2012

Ao clicar no nome da espécie, o usuário tem acesso ao mapa do Brasil com a sua distribuição informada pelo especialista da Lista do Brasil (estados coloridos), além dos pontos de ocorrência provenientes dos dados dos herbários participantes da rede *speciesLink*. O relatório apresenta o número de registros por ano de coleta, outra informação importante para avaliar o status de conservação da espécie. São também apresentadas várias tabelas sobre os dados disponíveis nos herbários, permitindo que o pesquisador possa avaliar a informação na origem e, se necessário, validar o dado analisando a amostra disponível.

Somente para citar um exemplo da importância de uma infraestrutura pública de dados para processos decisórios e para um maior controle social, a Figura 4.13 apresenta o relatório com o status dos dados gerado pelo sistema *Lacunas* para a espécie *Attalea funifera*, a *piaçava*, utilizada na confecção de vassouras e escovas. Trata-se de uma espécie endêmica do Brasil, considerada vulnerável na avaliação da Fundação Biodiversitas e incluída no Anexo II do MMA como deficiente em dados. Para o MMA, *com deficiência de dados* são aquelas cujas informações (distribuição geográfica, ameaças/impactos e usos, entre outras) são ainda deficientes, não sendo possível enquadrá-las com segurança na condição de ameaçadas.



Figura 4.13. Trechos do relatório sobre o status dos dados disponíveis para a espécie *Attalea funifera* Mart.  
Fonte: Lacunas (setembro de 2012).

Para esta espécie, o sistema *Lacunas* apresenta 55 registros *on-line*, sendo 25 pontos com coordenadas distintas (17 na origem e 8 por aplicativo). Essa espécie tem coletas de 1915 a 2009. Trata-se, portanto, de uma espécie relativamente bem coletada, com coletas recentes e para a qual se conhece o uso. Aparentemente, ela não deveria estar no Anexo II por não ser deficiente em dados. No entanto, a apresentação do status dos dados disponíveis *on-line* para uma espécie por si só não estabelece o seu status de conservação. O relatório apresentado pelo sistema *Lacunas* deve sempre ser avaliado por especialistas.

O cruzamento dos dados de diversas fontes e a produção de relatórios com mapas e gráficos de fácil interpretação precisa ser mais bem explorada pelas e-infraestruturas e agências de fomento. O sistema *Lacunas* é útil para subsidiar a definição, por especialistas, de áreas e grupos taxonômicos prioritários para coleta, de dados de acervos a serem integrados à rede, e ainda, para a determinação ou contestação do status de conservação de espécies.

#### 4.3. OS PROVEDORES DE DADOS: DIAGNÓSTICO E DESAFIOS PARA O FORTALECIMENTO DE E-INFRAESTRUTURAS

No Simpósio Latino-Americano de Coleções Biológicas e Biodiversidade: Conhecimento e Gestão organizado pelo Instituto Oswaldo Cruz (Fiocruz) em maio de 2012, a Dra. Carol Roetzel Butler, Chefe das Coleções do Museu Nacional de História Natural Smithsonian dos Estados Unidos, apresentou o processo de planejamento e avaliação interna da referida instituição (*Collections Assessment and Planning*<sup>103</sup>). O primeiro passo foi o estabelecimento de um comitê responsável pela avaliação de todas as coleções do Museu que, juntas, perfazem um total 127 milhões de espécimes e artefatos. As coleções foram organizadas em 266 grupos com sistemas de gestão comuns. A missão da instituição foi revisitada e o papel e a importância das coleções estabelecidas.

O grupo desenvolveu uma estrutura analítica com 13 elementos agrupados em quatro áreas: estado (armazenamento, conservação, acessibilidade); informação (identificação, metadados, registros eletrônicos, documentação); importância (missão, escopo, singularidade, uso) e divulgação (educação, exposição). A situação de cada elemento foi avaliada para cada coleção e classificada como (1) inaceitável, ausente; (2) ruim; (3) aceitável, padrão; (4) melhor; e (5) ideal.

---

<sup>103</sup> Relatório interno não publicado

A área “importância” foi apresentada graficamente com os valores atribuídos a cada uma das demais áreas, identificando assim as coleções de grande importância com problemas de armazenamento, informação e divulgação. Através dessa representação gráfica, foram priorizadas as coleções e definidas ações mais objetivas. Trata-se de um processo objetivo, transparente e participativo, através do qual foram definidas as prioridades lógicas e justificáveis. A implementação das decisões foi monitorada e o resultado do processo foi amplamente divulgado. Conseguiu-se, dessa forma, reduzir a competitividade e aumentar a colaboração e o trabalho em equipe.

A experiência realizada pelo Smithsonian serviu de inspiração para o trabalho apresentado a seguir, realizado no âmbito desta tese, a partir de consulta junto aos curadores e responsáveis pelas coleções biológicas que compartilham seus dados com a rede *speciesLink*. O objetivo foi testar um modelo de diagnóstico e definição de ações e estratégias de apoio às coleções biológicas que são provedores de dados da rede *speciesLink*.

Quando o formulário foi aplicado, a rede *speciesLink* integrava os dados de 268 coleções e subcoleções de 75 instituições. Diferentemente do Smithsonian, não existe uma missão institucional comum a ser cumprida, mas existem objetivos comuns, que são: melhorar a qualidade e a relevância dos acervos e promover a disseminação *on-line* dos dados de forma livre e aberta.

#### *4.3.1. Metodologia para Diagnóstico das Coleções Biológicas Provedoras de Dados*

Procurou-se estabelecer uma metodologia que envolvesse as coleções, utilizasse os dados da e-infraestrutura e fosse simples de implementar. As atividades incluíram (i) a elaboração de um questionário padrão; (ii) a estruturação do questionário *on-line*; (iii) o convite às coleções da rede *speciesLink* para o preenchimento do questionário; (iv) a tabulação das respostas; (v) a avaliação dos dados *on-line*; e, (vi) a análise e apresentação dos resultados.

##### *4.3.1.1. Elaboração do questionário padrão*

Para a elaboração do questionário, além da estratégia desenvolvida pelo Smithsonian, foram também introduzidos itens destacados pelas sociedades de zoologia, botânica e microbiologia em 25 documentos básicos e notas técnicas<sup>104</sup> utilizados para a elaboração da publicação *Diretrizes e estratégias para modernização de coleções biológicas brasileiras e a*

---

<sup>104</sup> Disponíveis na página <http://www.cria.org.br/cgee/col>

*consolidação de sistemas integrados de informação sobre biodiversidade* (EGLER e SANTOS, 2006). Foram também analisados os relatórios de atividades da rede *speciesLink* (2003, 2004, 2005, 2006)<sup>105</sup> e os relatórios do INCT – Herbário Virtual da Flora e dos Fungos (2009, 2010)<sup>106</sup>.

O objetivo foi a elaboração de um questionário de fácil preenchimento e interpretação, para uma primeira avaliação do cenário geral das coleções. O questionário, disponível no Anexo 1, apresenta os seguintes tópicos:

- A. Identificação do entrevistado (indicando que essa informação não será divulgada) e caracterização da instituição.
- B. Caracterização da Coleção: grupo taxonômico e tamanho do acervo.
- C. Avaliação da Coleção

Com base nos estudos dos documentos indicados, foram determinadas cinco áreas (recursos humanos, infraestrutura física, informatização, atividades e governança) englobando 22 elementos que se encontram relacionados a seguir.

- Recursos Humanos: curadoria, taxonomia / sistemática; apoio técnico; pesquisa;
- Infraestrutura física: recebimento e preparo de material, armazenamento, acesso e manipulação de material, segurança;
- Informatização: processo de documentação, equipe, equipamento, serviços Internet;
- Atividades: gestão da coleção, intercâmbio, pesquisa, ensino, serviços; e,
- Governança: reconhecimento institucional, fixação de recursos humanos, orçamento anual, planejamento estratégico, apoio externo a projetos.

Foi solicitado aos especialistas que as respostas fossem dadas a partir da situação atual da coleção. Para cada elemento, foram apresentadas as seguintes opções de resposta: (i) muito fraco ou ausente; (ii) fraco ou insuficiente; (iii) médio ou suficiente; (iv) forte ou bom; e, (v) muito forte ou muito bom.

#### D. Ações Prioritárias

Foi solicitado aos especialistas que indicassem o grau de prioridade dos itens a seguir.

- Ampliação do Acervo
- Fixação de Recursos Humanos

---

<sup>105</sup> Disponíveis na página <http://splink.cria.org.br/reports>

<sup>106</sup> Disponíveis na página <http://inct.florabrasil.net/producao>

- Formação de Recursos Humanos
- Informatização e Limpeza de Dados
- Interação com Outras Coleções
- Investimento na Capacitação em Gestão da Coleção
- Melhoria da Infraestrutura
- Planejamento
- Reconhecimento Institucional

Para cada item, as opções de resposta foram: prioridade máxima; prioritário; medianamente prioritário; baixa prioridade; não é prioritário.

#### 4.3.1.2. Estruturação do questionário on-line

Foi desenvolvido um formulário para preenchimento *on-line*, sendo que as respostas, quando enviadas, eram salvas como arquivo texto. Foram atribuídos os seguintes valores numéricos às respostas:

- Muito fraco ou ausente = 1
- Fraco ou insuficiente = 2
- Médio ou suficiente = 3
- Forte ou bom = 4
- Muito forte ou muito bom = 5

Para o bloco das Ações Prioritárias, os valores numéricos às respostas foram:

- Prioridade máxima = 1
- Prioritário = 2
- Medianamente prioritário = 3
- Baixa prioridade = 4
- Não é prioritário = 5

Os arquivos-texto foram combinados usando o comando *CAT* do *Linux* e abertos em uma planilha Excel para análise.

#### 4.3.1.3. Levantamento das informações junto às coleções biológicas

No dia 04 de junho de 2012, foi enviado um *e-mail* acompanhado da carta da orientadora apresentando a aluna e o objeto de estudo (disponíveis nos Anexos 2 e 3) a 323 pessoas

(diretores, curadores e responsáveis por coleções) de 268 coleções e subcoleções de 75 instituições. Todos os contatos foram obtidos dos metadados de cada coleção participante da rede *speciesLink*<sup>107</sup>. Foi dado um prazo até o dia 15 de junho de 2012 para o preenchimento do questionário, sendo que 17 *e-mails* voltaram com problemas (usuário inexistente, cota excedida, etc.). No dia 13 de junho, após o recebimento de 49 questionários preenchidos, foi enviado *e-mail* a todos que não responderam, novamente solicitando o preenchimento do questionário. O resultado final foi a resposta de 99 especialistas.

#### 4.3.1.4. *Tabulação dos dados*

Os dados foram tabulados em uma planilha Excel, sendo calculada a mediana e a frequência das respostas a cada elemento. Os resultados foram organizados em tabelas.

#### 4.3.1.5. *Avaliação dos dados on-line*

No modelo adotado pelo Museu Nacional de História Natural Smithsonian dos Estados Unidos, por ser uma avaliação interna, foi possível estabelecer a importância relativa dos diferentes acervos para a instituição. Para criar o conceito de acervos prioritários em rede, pode-se criar um painel de especialistas ou até desenvolver aplicativos que calculem a singularidade dos acervos a fim de que o resultado seja avaliado por especialistas. Para o presente estudo foi adotado como elemento prioritário o tamanho do acervo e o grau de informatização, sendo esses números tabulados com a mediana dos elementos avaliados no item “informatização” e plotados em um gráfico apresentado no item 4.3.3 a seguir.

### 4.3.2. *Resultados*

As tabelas originais com as respostas das coleções biológicas encontram-se no Anexo 4 sem a identificação das coleções ou dos responsáveis pelo seu preenchimento.

#### 4.3.2.1. *Caracterização das instituições e coleções*

49% das respostas foram de coleções mantidas em universidades públicas, 31% em instituições de pesquisa públicas, 9% em universidades privadas, 2% em empresas de direito público, 3% em empresas de direito privado e 6% de outras instituições como órgãos estatais e instituições sem fins lucrativos. 44% das respostas foram de herbários, 41% de coleções

---

<sup>107</sup> É importante frisar que essa informação é pública e tem o consentimento dos responsáveis pelas coleções para estar disponível *on-line*.

zoológicas e 15% de coleções microbianas. Em relação ao tamanho do acervo, 9% são coleções com até mil registros, 19% entre mil a cinco mil registros, 51% entre cinco mil e cinquenta mil registros, 7% entre 50 e 100 mil registros, 4% entre 100 a 200 mil registros, 6% entre 200 a 500 mil registros, 1% entre 500 a um milhão de registros e 3% com acervos com mais de um milhão de registros (Tabela 4.3).

Tabela 4.3. Tamanho dos acervos das coleções que preencheram o formulário

Tamanho do Acervo	Botânica	Zoologia	Microbiologia	Total	%
até um mil registros	1	1	7	9	9
de um mil a 5 mil registros	5	6	8	19	19
de 5 mil a 50 mil registros	25	25	50	51	51
de 50 mil a 100 mil registros	5	2	7	7	7
de 100 mil a 200 mil registros	4		4	4	4
de 200 mil a 500 mil registros	2	4	6	6	6
de 500 mil a um milhão de registros	1		1	1	1
acima de um milhão de registros		3	3	3	3
<b>Total</b>	<b>43</b>	<b>41</b>	<b>15</b>	<b>99</b>	

#### 4.3.2.2. Avaliação das Respostas

Como o objetivo deste estudo era o de propor um método participativo para a definição de prioridades e não propriamente definir uma estratégia de apoio às coleções biológicas do país, optou-se por analisar apenas os questionários das coleções botânicas por serem mais representativos do número e das características dos acervos *on-line*. Foram recebidas 43 respostas de 39 coleções de 18 estados brasileiros e do Distrito Federal, uma amostra bastante significativa do universo da rede *speciesLink* que, na data da distribuição do questionário, integrava os dados de 70 herbários do país. Obteve-se, portanto, uma taxa de retorno de 56%.

A maioria dos herbários que respondeu ao questionário (82%) tem entre mil e 100 mil registros, 16% têm mais que 100 mil e apenas 2% possuem acervos com menos de mil registros. Essa proporção é semelhante ao conjunto de herbários participantes da rede *speciesLink* onde 79% possuem entre mil e 100 mil registros, 15% mais de 100 mil registros e 6% menos de mil registros de amostras em seu acervo.

A seguir, são apresentadas as sínteses das respostas das coleções botânicas. Para cada critério analisado foi calculada a frequência das respostas e a mediana.

## 1. Recursos Humanos

A Tabela 4.4 apresenta a frequência das respostas enviadas pelas coleções sobre a disponibilidade de pessoal qualificado e em número adequado para os cargos indicados.

Tabela 4.4. Frequência das respostas sobre a disponibilidade e qualificação de recursos humanos.

Recursos Humanos	Frequência das Respostas						
	Muito fraco ou ausente	Fraco ou insuficiente	Médio ou suficiente	Forte ou bom	Muito forte ou muito bom	Sem resposta	Mediana
<b>Curadoria</b>	1	7	21	9	4	1	<b>3</b>
%	2	16	49	21	9	2	
<b>Taxonomia</b>	3	22	15	0	2	1	<b>2</b>
%	7	51	35	0	5	2	
<b>Apoio Técnico</b>	9	20	9	4	1	0	<b>2</b>
%	21	47	21	9	2	0	
<b>Pesquisa</b>	0	17	20	4	2	0	<b>3</b>
%	0	40	47	9	5	0	

De acordo com as respostas dos herbários, as maiores limitações com relação a recursos humanos são apoio técnico e taxonomia, com respectivamente 68% e 58% das respostas entre muito fraco (ou ausente) e fraco (ou insuficiente) e um valor de mediana igual a 2. Os recursos humanos para as atividades de curadoria e pesquisa são suficientes (ou bons) e muito bons, respectivamente para 80% e 60% dos herbários.

Com base nas respostas enviadas pelos herbários pode-se concluir que é necessário manter bolsas de apoio técnico e o programa de visita de especialistas para identificação de material, atualmente mantido pelo INCT Herbário Virtual da Flora e dos Fungos, como modo de suprir a demanda por taxonomistas.

## 2. Infraestrutura Física

O item infraestrutura física procurou avaliar as condições das salas de recebimento e preparo de material, armazenamento, acesso e manipulação e segurança. A Tabela 4.5 traz a frequência das respostas enviadas pelas coleções e respectiva porcentagem por elemento avaliado e o valor da mediana.

Tabela 4.5. Frequência das respostas sobre infraestrutura física

Infraestrutura	Frequência das Respostas					Mediana
	Muito fraco ou ausente	Fraco ou insuficiente	Médio ou suficiente	Forte ou bom	Muito forte ou muito bom	
<b>Preparo</b>	2	9	20	7	5	<b>3</b>
%	5	21	47	16	12	
<b>Armazenamento</b>	0	7	18	14	4	<b>3</b>
%	0	16	42	33	9	
<b>Acesso</b>	0	6	21	9	7	<b>3</b>
%	0	14	49	21	16	
<b>Segurança</b>	4	11	13	12	3	<b>3</b>
%	9	26	30	28	7	

A infraestrutura física em geral está satisfatória, com uma mediana igual a 3 para todos os itens avaliados. No entanto, 26% dos herbários avaliaram como muito fraca ou fraca a infraestrutura para recebimento e preparo de material; 16% como fraca a infraestrutura para armazenamento, 14% como fraca a infraestrutura para acesso e manipulação das amostras e 35% como fraca ou muito fraca a infraestrutura de segurança. Assim, apesar da mediana ser satisfatória, ainda existem coleções com problemas principalmente em relação à segurança. O resultado indica que investimentos nessa área ainda são necessários.

### 3. Informatização

Os dois itens anteriores, recursos humanos adequados e suficientes e infraestrutura satisfatória garantem a qualidade e organização do acervo. O item *informatização* tem por objetivo identificar se existem deficiências importantes para o compartilhamento dos dados *on-line*. É importante lembrar que o questionário foi somente enviado às coleções que já disponibilizam seus dados à rede, ou seja, nos quais já não havia resistências políticas ou culturais ao compartilhamento aberto de dados. Em relação à informatização, o questionário procura levantar se existem problemas com infraestrutura e pessoal.

Tabela 4.6. Frequência das respostas sobre informatização

Informatização		Frequência das Respostas					Mediana
		Muito fraco ou ausente	Fraco ou insuficiente	Médio ou suficiente	Forte ou bom	Muito forte ou muito bom	
<b>Documentação</b>		2	12	17	8	4	<b>3</b>
	%	5	28	40	19	9	
<b>Equipe</b>	<b>Digitação</b>	9	17	10	4	3	<b>2</b>
	%	21	40	23	9	7	
	<b>Digitalização</b>	15	17	7	2	2	<b>2</b>
	%	35	40	16	5	5	
	<b>Apoio</b>	15	12	12	2	2	<b>2</b>
	%	35	28	28	5	5	
<b>Equipamento</b>	<b>Digitação</b>	2	14	12	13	2	<b>3</b>
	%	5	33	28	30	5	
	<b>Digitalização</b>	20	10	7	4	2	<b>2</b>
	%	47	23	16	9	5	
<b>Internet</b>		1	9	19	12	2	<b>3</b>
	%	2	21	44	28	5	

O percentual da avaliação muito fraco (ou ausente) e fraco (ou insuficiente) é:

- Equipe de digitalização (imagens): 75%
- Equipamento para digitalização: 70%
- Equipe de apoio em informática: 63%
- Equipe de digitação: 61%
- Equipamento para digitação: 38%
- Documentação: 33%
- Internet: 23%

Tem-se que os grandes gargalos aparentemente são equipe e equipamento para a digitalização das imagens do acervo, seguidos pela equipe de apoio em informática e de digitação dos dados. 62% das respostas indicaram suficiência em equipamento para digitação dos dados, 67% indicaram como satisfatória a rotina de documentação e 77% indicaram como satisfatório o acesso à Internet. Esse último certamente reflete a efetividade da ação da RNP.

As respostas indicam que é necessário prover mão de obra para digitação e digitalização (imagens) do acervo e apoio técnico em informática, além de equipamento para a digitalização das imagens.

#### 4. Atividades

O item “Atividades” procura caracterizar a atuação da coleção com relação à sua gestão, atividade de intercâmbio, pesquisa; ensino, serviços de atendimento à sua demanda interna e externa. Aos responsáveis pela coleção foi solicitado atribuir um conceito entre muito fraco ou ausente a muito forte ou muito bom, para os quais são atribuídos valores numéricos entre 1 e 5, respectivamente.

Tabela 4.7. Frequência das respostas sobre atividades dos herbários

Atividades dos herbários	Frequência das Respostas						
	Muito fraco ou ausente	Fraco ou insuficiente	Médio ou suficiente	Forte ou bom	Muito forte ou muito bom	Sem resposta	Mediana
<b>Gestão</b>	1	5	23	11	3	0	<b>3</b>
%	2	12	53	26	7	0	
<b>Intercâmbio</b>	1	12	17	7	4	2	<b>3</b>
%	2	28	40	16	9	5	
<b>Pesquisa</b>	0	16	19	6	2	0	<b>3</b>
%	0	37	44	14	5	0	
<b>Ensino</b>	1	8	21	11	2	0	<b>3</b>
%	2	19	49	26	5	0	
<b>Serviços Internos</b>	3	7	18	11	4	0	<b>3</b>
%	7	16	42	26	9	0	
<b>Serviços Externos</b>	3	8	20	7	4	1	<b>3</b>
%	7	19	47	16	9	2	

As atividades em geral são bem avaliadas. Muitos ressaltaram, como comentários, o atendimento aos cursos de graduação e pós-graduação e aos alunos desde a iniciação científica ao doutorado. O percentual indicando uma avaliação entre muito fraco e fraco para cada atividade é menor do que aqueles apresentados nos itens anteriores:

- Pesquisa: 37%
- Intercâmbio: 30%
- Serviços externos: 26%
- Serviços internos: 23%

- Ensino: 21%
- Gestão: 14%

A avaliação mais positiva foi atribuída à gestão do herbário, seguida do ensino, atendimento à demanda interna, demanda externa, intercâmbio e pesquisa.

Essa avaliação, ainda que à distância e superficial, indica o caráter importante da prestação de serviços à comunidade, interna e externa, atividades normalmente não valorizadas nos sistemas de avaliação vigentes.

## 5. Governança

Uma avaliação sobre governança procurou levantar informações sobre o reconhecimento institucional, a fixação de recursos humanos, o orçamento anual, o planejamento estratégico e o apoio externo a projetos apresentados pelos herbários.

Tabela 4.8. Frequência das respostas sobre governança.

Governança	Frequência das Respostas						Mediana
	Muito fraco ou ausente	Fraco ou insuficiente	Médio ou suficiente	Forte ou bom	Muito forte ou muito bom	Sem avaliação	
Reconhecimento Institucional	8	11	10	10	4		<b>3</b>
%	19	26	23	23	9		
Fixação de RH	13	15	11	3	1		<b>2</b>
%	30	35	26	7	2		
Orçamento Anual	17	13	9	2	2		<b>1</b>
%	40	30	21	5	5		
Planejamento Estratégico	4	11	24	4			<b>3</b>
%	9	26	56	9	0		
Apoio Externo a Projetos	4	17		10		12	2
%	9	40	0	23	0	28	

Os percentuais apresentados como fracos (insuficientes) ou muito fracos (ausentes) são:

- Orçamento anual: 70%
- Fixação de recursos humanos: 65%
- Reconhecimento institucional: 45%
- Planejamento estratégico: 35%

- Apoio externo a projetos: 49% sendo que 28% não responderam

O resultado corresponde à expectativa de que esse é um dos principais problemas dos herbários, constatando que, para a maioria, o orçamento não atende às suas necessidades e existem dificuldades quanto à fixação de recursos humanos. Tem-se ainda que 45% dos herbários não têm reconhecimento institucional adequado. Considerando esses problemas de governança, surpreendeu o índice de 63% das respostas que consideraram o planejamento estratégico suficiente ou bom.

Outro ponto importante a ser analisado é a dificuldade em conseguir apoio externo a projetos. Uma das causas pode ser a ausência de mecanismos de valorização da prestação de serviços científicos (p.ex. identificação de material, curadoria, compartilhamento de dados), fazendo com que pesquisadores que acumulam a função de curadoria não tenham tempo para se dedicar às publicações e acabem tendo suas propostas rejeitadas em sistemas competitivos com base na análise de seus currículos.

## 6. Ações Prioritárias

Foram selecionadas as principais ações indicadas como prioritárias nos documentos utilizados nesse estudo para avaliar a opinião das coleções.

Tabela 4.9. Priorização das ações das coleções

Ações Prioritárias	Frequência das Respostas						
	Prioridade Máxima	Prioritário	Mediana-mente Prioritário	Baixa Prioridade	Não é Prioritário	Não Respondeu	Mediana
Ampliação do Acervo (%)	13	15	11	2	1	1	2
	30	35	26	5	2	2	
Fixação de Recursos Humanos (%)	18	14	5	4	1	1	2
	42	33	12	9	2	2	
Formação de Recursos Humanos (%)	11	20	7	3		2	2
	26	47	16	7	0	5	
Informatização e Limpeza de Dados (%)	17	19	4	1		2	2
	40	44	9	2	0	5	
Interação com Outras Coleções (%)	8	18	14	2		1	2
	19	42	33	5	0	2	

Ações Prioritárias	Frequência das Respostas						
	Prioridade Máxima	Prioritário	Medianamente Prioritário	Baixa Prioridade	Não é Prioritário	Não Respondeu	Mediana
Investimento na Capacitação em Gestão da Coleção (%)	7	15	15	5		1	2
Melhoria da Infraestrutura (%)	16	35	35	12	0	2	
Planejamento (%)	17	10	8	5	1	2	2
Reconhecimento Institucional (%)	40	23	19	12	2	5	
	6	23	11	2		1	2
	14	53	26	5	0	2	
	8	20	7	5	2	1	2
	19	47	16	12	5	2	

A análise da mediana indica todos os itens como sendo prioritários. Ao avaliar a frequência das respostas *prioridade máxima* ou *prioritário* por item, tem-se:

- Informatização e limpeza de dados: 84%
- Fixação de recursos humanos: 74%
- Formação de recursos humanos: 72%
- Planejamento: 67%
- Ampliação do acervo e reconhecimento institucional: 65%
- Melhoria da infraestrutura: 63%
- Interação com outras coleções: 60%
- Investimento na capacitação em gestão da coleção: 51%

Alguns itens não são da competência das agências financiadoras, mas mostram diretrizes claras para o apoio institucional. É importante observar que o item “informatização e limpeza de dados” foi considerado o mais prioritário dentre todas as opções, o que fortalece a ação da e-infraestrutura.

#### 4.3.3. Proposta de Aplicação

O objetivo do levantamento junto às coleções biológicas foi realizar um diagnóstico dos provedores de informações para uma e-infraestrutura (*specieslink*), visando maior efetividade nesse trabalho participativo e interativo. O diagnóstico é importante para conhecer a diversidade das coleções, suas necessidades e dificuldades, fornecendo os elementos necessários para

priorizar determinadas linhas de apoio. Ao estabelecer estratégias a partir das coleções biológicas que estão compartilhando os dados não sigilosos de seus acervos *on-line*, o poder público premia e estimula essas coleções a continuarem com esse trabalho.

O estudo apresentado indica, por exemplo, que o edital para a rede *speciesLink* deveria incluir bolsas ou pagamento de serviços de terceiros para digitação dos dados e digitalização das amostras, passagens e diárias para o programa de visita de especialistas, equipamento de digitalização, entre outros. Pode-se concluir que a avaliação das coleções apresenta demandas claras para o financiamento de agências de fomento.

O cruzamento com os dados *on-line* permite avaliar a contribuição real e potencial de cada acervo à e-infraestrutura. A avaliação do grau de informatização versus tamanho do acervo pode indicar a necessidade de ações induzidas por parte das agências de fomento. Adicionando a esses dois componentes (tamanho e grau de informatização) o valor da mediana para o critério Informatização pode-se avaliar a situação de cada acervo em relação à atividade informatização e à meta que é 100% dos dados (não sensíveis) disponíveis *on-line*. Ao plotar o resultado em um gráfico (Figura 4.14) tem-se para cada herbário que respondeu ao questionário o tamanho do acervo (tamanho da bola) e o percentual dos dados *on-line* versus a mediana do critério Informatização, lembrando que 1 e 2 são respectivamente muito fraco e fraco e valores a partir de 3 representam uma condição satisfatória.

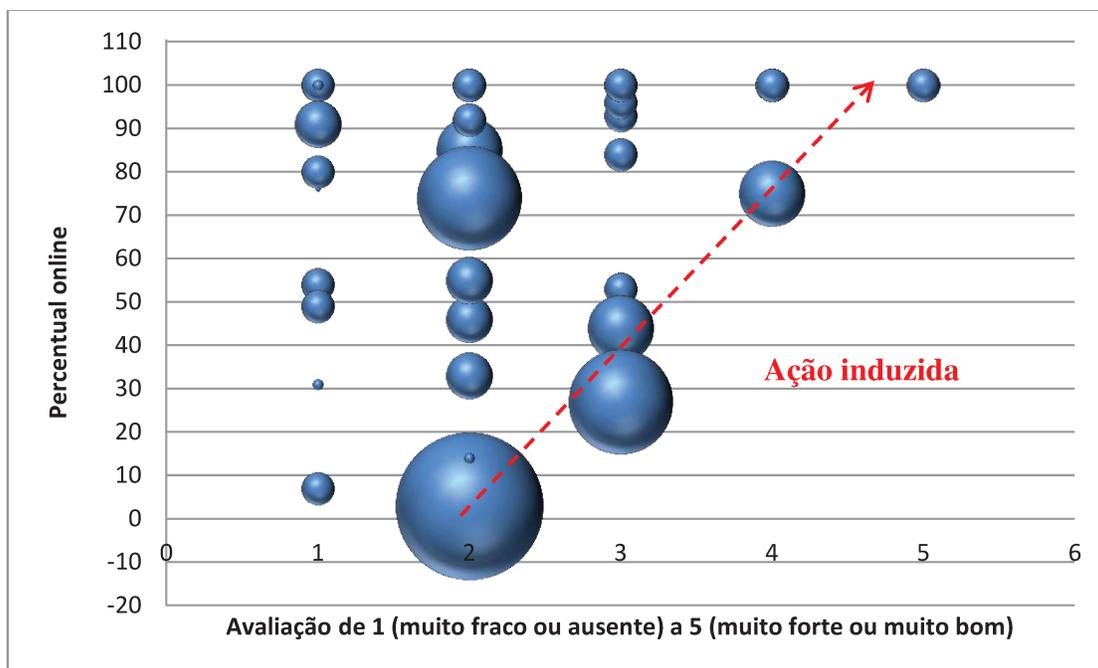


Figura 4.14. Gráfico do percentual dos dados *on-line* versus a avaliação do critério *Informatização* por herbário com o tamanho da bola proporcional ao tamanho do acervo.

A representação gráfica permite uma observação rápida do status dos herbários em relação ao critério *Informatização*. Existem três coleções maiores que avaliaram como fraco o critério *Informatização* em suas análises, sendo que a maior tem um percentual de dados *on-line* menor que 10. Visando transformar a informação existente, mas não disponível, rapidamente em informação *on-line*, os acervos maiores com baixo nível de informatização poderiam ser objeto de uma ação induzida visando a digitação e disponibilização dos dados de seus acervos *on-line*. A seta vermelha na Figura 4.14 procura ilustrar a meta a ser atingida pela maior coleção da rede através de uma ação induzida, que é melhorar a atividade de informatização e ter 100% dos dados não sensíveis do acervo disponíveis *on-line*.

Com a avaliação da situação e das necessidades das coleções com os dados *on-line*, além de definir os editais e ações induzidas que podem modificar o cenário atual, é possível monitorar o progresso das coleções e a eficácia da ação.

Esse é apenas um exemplo do tipo de análise que é possível realizar cruzando a avaliação da coleção com os dados *on-line*. Esse sistema pode ser mais sofisticado com o desenvolvimento de outros aplicativos para estabelecer prioridades no uso dos recursos públicos como, por exemplo, a avaliação da singularidade dos acervos.

Além dos editais abertos e ações induzidas buscando fomentar a informatização dos acervos e a melhoria da qualidade dos dados, é importante apoiar ações de articulação e coordenação, como as dos Institutos Nacionais de Ciência e Tecnologia, incentivando programas colaborativos e não competitivos, visando um objetivo comum que é o de disponibilizar dados não sensíveis e de qualidade sobre a ocorrência de espécies em uma infraestrutura eletrônica de acesso livre e aberto.

#### 4.4. DISCUSSÃO

As infraestruturas eletrônicas de dados sobre biodiversidade estão sendo fomentadas, principalmente visando o desenvolvimento de projetos de ciência global. Com o desenvolvimento das TICs é possível não só o uso e reuso de dados, informações e aplicativos pelas disciplinas que os geraram, mas por outras áreas de conhecimento, ou até, integrando diferentes áreas de conhecimento. Os trabalhos apresentados neste capítulo cumprem um dos objetivos da tese ao mostrar possíveis usos da infraestrutura pelo poder público no monitoramento e acompanhamento de projetos e programas, na valorização de serviços científicos providos por curadores e taxonomistas, na identificação de lacunas de dados e conhecimento, e no desenvolvimento de estratégias de fomento para coleções biológicas provedoras de dados e informações para as redes.

Nos exercícios e exemplos apresentados, procurou-se demonstrar que as e-infraestruturas representam uma oportunidade para o desenvolvimento e uso de indicadores para subsidiar os sistemas de avaliação e monitoramento de programas e projetos. Assim, uma política de acesso livre e aberto aos dados e de apoio de longo prazo às e-infraestruturas pode também subsidiar sistemas de avaliação.

O impedimento taxonômico é um tema proposto pelo Brasil junto à CDB que motivou a criação da Iniciativa Mundial sobre Taxonomia. No entanto, não existem indicadores para valorizar esse serviço científico e, dessa forma incentivar o trabalho dos taxonomistas na determinação das espécies, atividade essa essencial para garantir a utilidade e qualidade da informação disseminada. Procurou-se mostrar que, com base nos dados disponibilizados pelas e-infraestruturas, é possível desenvolver aplicativos capazes de monitorar e quantificar o trabalho dos especialistas envolvidos com a identificação de material biológico, que assim, pode ser valorizado como um serviço científico. O mesmo foi demonstrado em relação ao papel do curador, responsável pela curadoria da coleção e pelo envio dos dados do acervo à e-

infraestrutura. Sugere-se assim o estudo e aprimoramento desses aplicativos pelas agências de fomento para a sua implementação como indicador de produtividade desses especialistas.

Castro (2011) indica a importância do lançamento de editais pelo Programa Biota-Fapesp, buscando induzir pesquisas em determinadas áreas de interesse. Afirma que um programa restrito à demanda espontânea de projetos gera dificuldades no cumprimento de seus objetivos maiores. Nessa linha, o segundo exemplo apresentado aponta o uso da e-infraestrutura para identificar lacunas de dados e conhecimento, baseando-se somente nas informações disponíveis *on-line* para subsidiar a análise de especialistas e assim definir prioridades e estratégias de ação visando aumentar a disponibilidade e a qualidade dos dados *on-line*. Possíveis ações incluem: (i) identificar e integrar novos acervos do país e do exterior à infraestrutura de dados; (ii) melhorar a qualidade taxonômica e geográfica dos dados disponíveis *on-line*; (iii) digitar os dados dos acervos das coleções biológicas do país que ainda não estão nesse formato; e, (iv) realizar novas coletas considerando as lacunas geográficas e taxonômicas identificadas.

A relação apresentada segue uma lógica de custo. É normalmente mais barato integrar novos acervos já digitais à infraestrutura e melhorar a qualidade dos dados, seguida pela digitação dos dados. O custo maior normalmente é aquele que envolve expedições científicas para a realização de novas coletas. Todas as atividades são essenciais, mas ao avaliar a informação disponível *on-line* é possível direcionar melhor a ação e os recursos necessários para preencher as lacunas identificadas.

O status dos dados também representa um importante instrumento na definição do status de conservação de espécies ameaçadas. Tradicionalmente, cabe ao especialista levantar dados e preencher fichas para avaliar o grau de ameaça de cada espécie e seu status de conservação, seguindo critérios pré-estabelecidos. A ferramenta *Lacunas* não dispensa a análise do especialista. Permite uma avaliação baseada nos dados publicamente disponíveis *on-line*, análise que pode ser validada ou contestada por outros atores. Pode também ser útil para o monitoramento da evolução do status dos dados (conhecimento) sobre a espécie. Pode ainda ser integrada ou associada a outros sistemas *on-line* para instrumentar o especialista na sua análise. Certamente torna o processo de tomada de decisão mais transparente e dinâmico.

O terceiro caso procurou demonstrar a importância dos provedores de dados e de estratégias de apoio articuladas com as coleções biológicas. A realização de um diagnóstico participativo e transparente permite a definição de ações consensuais e o seu acompanhamento

via resultados refletidos na e-infraestrutura. Promove uma maior cooperação visando um objetivo comum, ao invés da competição por recursos, ao tornar todo processo transparente.

As e-infraestruturas sobre a diversidade de espécies são um importante elemento da comunicação científica. Para o país, é fundamental desenvolver e manter sua própria infraestrutura de dados, apropriando-se da informação e fazendo uso dela nas diferentes áreas de conhecimento e nas diferentes esferas de tomada de decisão. Procurou-se enfatizar o seu uso pelas agências de fomento, que devem se capacitar para se apropriar dos dados e ferramentas disponibilizadas, além de criar novas demandas por novos dados ou composição de dados e desenvolvimento de aplicativos. Para tanto, é fundamental que a política de dados sobre a diversidade de espécies tenha como regra o seu acesso livre e aberto, com um tratamento diferenciado para dados sigilosos ou sensíveis.

## CONCLUSÕES

Não é sem razão que Castells (1999) considera a revolução das tecnologias de informação e comunicação como um evento histórico da mesma importância da Revolução Industrial do século XVIII, chegando mesmo a induzir a um padrão de descontinuidade nas bases materiais da economia, sociedade e cultura. É nessa mesma linha, também, que o relatório do *National Science Board* da *National Science Foundation*, perante a evidência do surgimento raro de novas abordagens à essência da pesquisa, indica a tecnologia da informação e as coleções digitais de dados respectivamente como promotor e centro dessa mudança (NSB, 2005).

Existe uma demanda da ciência aberta contemporânea, pela qual não só os dados precisam ser preservados e disponibilizados, mas os processos, as análises e métodos também, e, se possível, em tempo real, proporcionando uma maior transparência e reprodutibilidade dos resultados (REICHMAN, JONES e SCHILDHAUER, 2011). É esse o lugar que a Internet está ocupando: o local onde bilhões de pessoas compartilham seus conhecimentos, recursos e capacidade de computação (TAPSCOTT e WILLIAMS, 2010). Trata-se de um novo tempo, em que não é mais suficiente o paradigma tradicional de compartilhamento de dados e resultados científicos através apenas de publicações em livros e revistas especializadas (SCHOFIELD, EPPIG, *et al.*, 2010).

Além da tecnologia e da necessidade científica de compartilhar dados, métodos e análises em diferentes escalas e disciplinas, existe uma necessidade de uso de uma infraestrutura de dados e aplicativos para processos de tomada de decisão, também em diferentes escalas, do local até global. Citando como exemplo os objetivos de desenvolvimento do milênio (*Millennium Development Goals* das Nações Unidas) – compromissos assumidos pelo Brasil – e as metas universais de desenvolvimento sustentável (*Universal Sustainable Development Goals* SDGs), são necessários dados precisos, adequados e disponíveis para gestores, tomadores de decisão e para o público (SACHS, 2012).

Essa é a motivação da tese: procurar compreender a influência e o impacto das tecnologias da informação e comunicação na circulação do conhecimento científico e avaliar o seu efeito na elaboração de políticas públicas em biodiversidade.

Os avanços das TICs tornaram possível o envolvimento de mais atores na gênese do conhecimento, atores de diferentes disciplinas, especialidades, instituições, localidades, países,

culturas e realidades sociais. Essa possibilidade é fundamental para o tema desenvolvimento sustentável.

No caso do Brasil, trata-se de um país de dimensões continentais que ocupa várias zonas climáticas (trópico úmido no Norte, o semiárido no Nordeste e áreas temperadas no Sul) que levam à formação de várias zonas biogeográficas distintas ou biomas (Floresta Amazônica, Cerrado, Caatinga, Mata Atlântica, Pantanal e Pampa), além da zona costeira com recifes, dunas, manguezais, lagoas, estuários e pântanos. Essa variedade biogeográfica reflete a enorme diversidade da flora, fauna e microbiota do país.

Como país megadiverso, o Brasil necessita de ações efetivas e urgentes tanto para diminuir a perda da biodiversidade, como também para recuperar áreas degradadas e, dessa forma, procurar garantir a manutenção dos serviços ambientais tão vitais para o bem-estar humano. Como ferramenta, o governo brasileiro dispõe de medidas de incentivo e de regulamentação do uso dos recursos naturais pelos setores privados e públicos e a ação direta sobre as Unidades de Conservação sob o seu controle, tanto as de proteção integral como as de uso sustentável, nas diferentes esferas – federal, estadual e municipal. Necessita, portanto, de estratégias multiescala e de curto a longo prazos, visando um desenvolvimento ambientalmente sustentável, socialmente justo e economicamente viável.

Em julho de 2011, o Conselho Assessor em Ciência e Tecnologia da Presidência dos Estados Unidos (PCAST – *President's Council of Advisors on Science and Technology*) apresentou um relatório denominado Sustentando o Capital Ambiental: Protegendo a Sociedade e a Economia (*Sustaining Environmental Capital: Protecting Society and the Economy*). O relatório indica que o impacto e a eficácia da ação do governo podem ser melhorados através do bom uso dos dados, modelos e conhecimentos disponíveis e através de ações focadas visando atualizar e aumentar esse conhecimento, modelos e bases de dados. As recomendações apresentadas envolvem um esforço para encontrar maneiras de fazer melhor uso do conhecimento existente, para apoiar a geração de novos conhecimentos essenciais, e para expandir o uso da informática (PCAST, 2011).

Assim, certamente será um diferencial se o Brasil desenvolver políticas públicas para aumentar o conhecimento sobre a biodiversidade, garantindo, ao mesmo tempo, a manutenção permanente e o desenvolvimento contínuo de uma infraestrutura eletrônica de dados, informações e aplicativos, de acesso livre e aberto. A manutenção e o desenvolvimento contínuo de tais

infraestruturas eletrônicas é essencial tanto para a pesquisa em biodiversidade como para os processos de tomada de decisão.

Acredita-se numa política de acesso livre e aberto a dados sobre biodiversidade, na qual determinados tipos de dados que necessitam de uma estratégia diferenciada são exceção – e não o contrário – ou seja, a proteção ou sigilo dos dados e informações geradas como regra e a disseminação aberta como exceção. Essa posição foi demonstrada ao longo da tese mas, para reforçá-la, é interessante compará-la com uma política restritiva, como a da Medida Provisória 2.052, de 29 de junho de 2000, que regulamenta o acesso a recursos genéticos no país. A Medida Provisória (MP) define como acesso ao patrimônio genético a “*obtenção de amostra de componente do patrimônio genético para fins de pesquisa científica, desenvolvimento tecnológico ou bioprospecção, visando à sua aplicação industrial ou de outra natureza*”. Com a inclusão do termo *ou de outra natureza*, quase toda a ciência biológica é incluída, seja ela de aplicação industrial ou não. A MP só exclui, em seu artigo 3º, o patrimônio genético humano. Foi uma lei elaborada com a preocupação de evitar a biopirataria, sem tratar de forma adequada a promoção do desenvolvimento científico e o aumento do conhecimento sobre a biodiversidade do Brasil. O acesso e uso ficam, a priori, circunscritos a uma lógica de criminalização (SALLES, TAVARES, *et al.*, 2007) e de propriedade industrial, visando a repartição dos benefícios econômicos que os produtos obtidos da biodiversidade brasileira eventualmente possam gerar para as indústrias.

Ao longo dos anos, vários textos para discutir um anteprojeto de lei têm sido apresentados buscando viabilizar as atividades de coleta, transporte e remessa de material biológico para atividades de pesquisa científica e tecnológica (SETTE e VAZOLLER, 2010), mas infelizmente não houve avanços. Os mecanismos criados para garantir a preservação dos recursos genéticos coletados, que são depositados em instituições *fiéis depositárias*, têm por objetivo permitir o rastreamento do patrimônio genético acessado por instituição devidamente autorizada, visando à repartição de benefícios (AZEVEDO e SILVA, 2005). Os acervos têm um tratamento semelhante aos acervos envolvidos com depósitos de patentes, não havendo qualquer política ou estratégia para a disseminação dos dados e informações geradas. Assim, por causa da bioprospecção – sem dúvida um importante, mas pequeno segmento do conhecimento de componentes da biodiversidade – toda a pesquisa visando inventariar e aprofundar o conhecimento sobre a diversidade biológica do país foi muito dificultada e, assim, prejudicada.

Esse exemplo mostra a importância de ter como regra, para biodiversidade, o acesso livre e aberto aos dados, informações e conhecimentos, tratando as exceções em separado.

Uma questão comum levantada em algumas reuniões, que inclusive foi durante alguns anos um dos argumentos para a não adesão do país ao GBIF, é – Por que o país deve investir em uma infraestrutura aberta de dados se os dados da rede GBIF já são de acesso público e aberto?, - ou ainda – Como país megadiverso, não seria vantajoso manter os dados em sigilo para assegurar uma vantagem competitiva?.

A tese procurou mostrar que não basta ao país ter acesso a dados e informações disponibilizadas em infraestruturas internacionais. A organização local dos dados e a sua disponibilização *on-line* com foco nas suas necessidades contribui para uma maior organização da comunidade científica e viabiliza estudos de problemas locais, aumentando a importância do tema “biodiversidade” nas agendas científicas e políticas de um país. Essa organização também promove o desenvolvimento científico com a possibilidade do uso e reuso dos dados e da integração de dados de diferentes fontes e áreas de conhecimento. Contribui também para estudos e para a resolução de problemas globais, ao disponibilizar e integrar dados, informações e conhecimentos locais nas infraestruturas globais. E acredita-se também que para setores que eventualmente necessitam de um controle ou sigilo de determinados dados, é necessário desenvolver uma estratégia específica.

Na tese, portanto, foi exposta a importância da política de acesso livre e aberto, colocando foco nos dados sobre espécies e espécimes mantidos em coleções biológicas, que representam um componente importante nos processos de tomada de decisão. O tratamento especial a esses dados deve-se a vários fatores como: a quantidade significativa de amostras existentes em coleções no mundo, calculada como sendo mais que 2 bilhões (ARIÑO, 2010); a existência de material-testemunho (amostra ou *voucher*) que pode ter a sua nomenclatura e identidade revisadas; a possível associação com dados das anotações do campo; e a disponibilidade de dados históricos, com o registro da ocorrência de espécies desde o período paleontológico e em áreas hoje perturbadas (p.ex. urbanizadas ou áreas agrícolas). Portanto, percebe-se que muitos desses dados são únicos e insubstituíveis, o que aumenta a importância de conhecê-los sem reserva.

Problemas associados a esses dados incluem seja o viés geográfico (a maioria das coletas ocorre ao longo das estradas e rios e próximos às cidades e estações biológicas) seja a imprecisão taxonômica, com erros na identificação ou falta de atualização nomenclatural, além de material

não identificado. Como a utilidade dos dados depende da acurácia na identificação da amostra (GRAHAM, FERRIER, *et al.*, 2004), a tese procurou destacar as atividades de curadoria das coleções biológicas e o trabalho do taxonomista na identificação das amostras como importantes serviços científicos que devem ser valorizados e reconhecidos, inclusive para estimular a atividade, aumentando a disponibilidade e qualidade dos dados.

Fundamental também reconhecer a importância da taxonomia como atividade de pesquisa, que necessita avançar no uso de tecnologias emergentes principalmente em relação ao desenvolvimento de redes e da tecnologia da informação. Wheeler *et al.* (2004) indicam a necessidade de tratar a taxonomia como uma ciência internacional de larga escala. A meta de descobrir, descrever e classificar as espécies do planeta, qualificam-na como *big science*. Afirmam que uma ciência em escala planetária necessita de uma ferramenta também em escala global, e citam como exemplo o futuro uso de instrumentação remota (p.ex, uso de microscópio remoto para identificação de material biológico via rede). Os autores ainda indicam como provável a revisão, pelos pares, de descrições de espécies publicadas eletronicamente e disponibilizadas instantaneamente, além do acesso livre e aberto às publicações de interesse, a qualquer pessoa interessada (WHEELER, RAVEN e WILSON, 2004).

Essa visão de 2004 já está se concretizando. A Lista de Espécies da Flora do Brasil é uma experiência bem sucedida de formação de uma rede de especialistas de diferentes locais do mundo, responsáveis pela classificação e publicidade instantânea e *on-line* de dados sobre espécies de plantas do Brasil. A disponibilização de dados e imagens de amostras de herbários *on-line* (muitas em alta resolução) através do INCT Herbário Virtual da Flora e dos Fungos e dos aplicativos que permitem analisar cada imagem em detalhe, e compará-las com outras, abre a possibilidade de realizar um serviço de identificação de material *on-line*.

Um outro conceito importante, que também foi apresentado na tese, é que além do princípio do acesso livre e aberto aos dados, informações e conhecimento, é importante que os desenvolvimentos das e-infraestruturas ocorram em ambientes abertos. A utilização de software, ferramentas e padrões proprietários, mesmo quando os dados são disponibilizados sem restrição, pode dificultar ou até mesmo impedir o seu acesso e uso.

Dependendo da análise desejada, aos dados sobre espécies e espécimes é necessário associar dados climáticos, geológicos, mapas, imagens de satélite, dados econômicos, sociais, entre muitos outros. Assim, para o país é importante desenvolver e manter diferentes e-

infraestruturas, de diferentes áreas de conhecimento, em ambientes não proprietários, capazes de receber, preservar, recuperar e servir dados e informações em formatos úteis e utilizáveis para pessoas e para outros sistemas.

Essa atividade exige planejamento de longo prazo e a competência de equipes especializadas.

Assad (p. 170, 2000) afirma:

*“Definir políticas para biodiversidade requer muito mais do que ouvir e contar com a colaboração de competentes especialistas. Requer identificar, colher, interpretar e organizar de forma inteligível todas as complexas variáveis que são fundamentais para promover a conservação e uso sustentável da diversidade biológica, além de envolver e comprometer os diversos atores que participam dos diferentes níveis na execução das atividades.”*

As e-infraestruturas representam uma necessidade e uma oportunidade para que os diferentes atores possam participar do processo, não só de definição de políticas para biodiversidade, mas também de monitoramento das ações e dos resultados alcançados.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ACKOFF, R. L. From Data to Wisdom. **Journal of Applied Systems Analysis**, 16, 1989. 3-9.
- ARIÑO, A. H. Approaches to estimating the universe of natural history collections data. **Biodiversity Informatics**, 7, n. 2, 2010. <https://journals.ku.edu/index.php/jbi/article/view/3991>.
- ASSAD, A. L. D. **Biodiversidade: institucionalização e programas governamentais no Brasil**. Tese (doutorado) Universidade Estadual de Campinas, Instituto de Geociências. Campinas, p. 174. 2000. (As71e).
- AZEVEDO, C. M. D. A.; SILVA, F. Á. D. **Regras para o Acesso Legal ao Patrimônio Genético e Conhecimento Tradicional Associado**. Ministério do Meio Ambiente. Brasília. 2005.
- BARBOSA, M. R. D. V.; PEIXOTO, A. L. Coleções Botânicas Brasileiras: situação atual e perspectivas. In: PEIXOTO, A. L. ( ). **Coleções Biológicas de Apoio ao Inventário, Uso Sustentável e Conservação da Biodiversidade**. Rio de Janeiro: Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, 2003. Cap. 6, p. 113-139. ISBN 85-7224-012-8.
- BATISTA, G. A. Francis Bacon: para uma educação científica. **Revista Teias**, 11, n. 23, 2010. [http://www.periodicos.proped.pro.br/index.php?journal=revistateias&page=article&op=viewFile&path\[\]=670&path\[\]=561](http://www.periodicos.proped.pro.br/index.php?journal=revistateias&page=article&op=viewFile&path[]=670&path[]=561).
- BELUZZO, A. M. D. M. **O Brasil dos Viajantes**. São Paulo: Fundação Odebrecht, 1994.
- BRAVERMAN, H. **Trabalho e Capital Monopolista: A degradação do trabalho no século XX**. 3a ed. 1981. ed. Rio de Janeiro: [s.n.], 1974. 137-147 p.
- BRIDGEWATER, P. et al. **GBIF Review 2009 - From Prototype to Full Operation: Managing Expectations**. Copenhagen, p. 39. 2010. [http://imgbif.gbif.org/CMS\\_ORC/?doc\\_id=2984&download=1](http://imgbif.gbif.org/CMS_ORC/?doc_id=2984&download=1).
- BROCKWAY, L. H. Science and Colonial Expansion: The Role of British Royal Botanic Gardens. **American Ethnologist**, v. 6, n. 3, p. 449-465, Aug. 1979. Disponível em: <[www.jstor.org/stable/643776](http://www.jstor.org/stable/643776)>. Acesso em: 19 jul. 2012.
- BROWNE, J. **Darwin's Origin of Species - A Biography**. London: Atlantic Books, 2006. 174 p. ISBN 978-1-84354-394-7.
- BUSH, V. **Science the Endless Frontier**. Washington. 1945. (<http://www.nsf.gov/od/lpa/nsf50/vbush1945.htm>).

- CANHOS, D. A. L. et al. **Rede Brasileira de Informação em Biodiversidade**. Brasília, p. 60. 2000. ([http://sistemas.mma.gov.br/sigepro/arquivos/\\_6/relatorio2000%20FAT.pdf](http://sistemas.mma.gov.br/sigepro/arquivos/_6/relatorio2000%20FAT.pdf)).
- CANHOS, D. A. L.; CANHOS, V. P. Disseminação de Informação: O uso da Internet. In: GARAY, I. E. G.; DIAS, B. F. S. **Conservação da Biodiversidade em Ecossistemas Tropicais**. Petrópolis: Editora Vozes, 2001. ISBN 85.326.2529-0.
- CASH, D. W. et al. Knowledge systems for sustainable development. **PNAS**, v. 100, n. 14, p. 8086-8091, July 2003.
- CASTELLS, M. **The rise of the network society (A sociedade em rede)**. Tradução de Roneide Venâncio Majer. São Paulo: Editora Paz e Terra S.A., v. 1, 1999. 617 p. ISBN 8521903294.
- CASTRO, P. F. D. D. **Avaliação de impacto de programas de pesquisa em biodiversidade**. Tese (doutorado) Universidade Estadual de Campinas, Instituto de Geociências. Campinas, p. 222. 2011. (C279a). Orientador: Maria Beatriz Machado Bonacelli.
- CHAPMAN, A. D. **Uses of Primary Species-Occurrence Data**. Copenhagen. 2005. <http://www2.gbif.org/UsesPrimaryData.pdf>.
- CHAPMAN, A. D. **Numbers of Living Species in Australia and the World**. Commonwealth of Australia. Canberra. 2009. (978 0 642 56861 8). <http://www.environment.gov.au/biodiversity/abrs/publications/other/species-numbers/2009/index.html>.
- CHAPMAN, A. D.; GRAFTON, O. **Guide to best practices for generalising primary species-occurrence data**. Copenhagen, p. 27. 2008. (ISBN: 87-92020-06-2). <http://www2.gbif.org/BPsensitivedata.pdf>.
- COZZENS, S. E. Frontiers of evaluation: United States. **Revista Brasileira de Inovação**, Campinas (SP), v. 11, n. n.esp., p. 47-66, julho 2012. ISSN 2178-2822.
- CRIA. **Relatório de Gestão 2011**. Centro de Referência em Informação Ambiental. Campinas, p. 44. 2011.
- DALY, D. C. Trilhas Botânicas do Rio Negro. In: (EDS), A. A. D. O. E. D. C. D. **Florestas do Rio Negro**. São Paulo: Companhia das Letras, UNIP, 2001. Cap. 1. ISBN ISBN 85-359-0142-6. <http://ecologia.ib.usp.br/guiaigapo/images/livro/RioNegro01.pdf>.

- DAVID, P. **The Economic Logic of "Open Science" and the Balance between Private Property Rights and the Public Domain in Scientific Data and Information: a Primer.** Washington, p. 238. 2003. (ISBN: 0-309-52545-4). <http://www.nap.edu/catalog/10785.html>.
- DAVID, P. A. The Historical Origins of 'Open Science': An Essay on Patronage, Reputation and Common Agency Contracting in the Scientific Revolution. **Capitalism and Society**, 3, n. 2, 2008. <http://www.bepress.com/cas/vol3/iss2/art5>.
- DIAS, B. F. S. et al. **Roles of the Clearing-House Mechanism in Promoting and Facilitating the Implementation of the Convention on Biological Diversity.** Brasília. 1997. <http://www.cbd.int/doc/ref/chm-roles-br-en.pdf>.
- DIETRICH, S. M. C. Ciência e Tecnologia no Brasil: Uma Nova Política para um Mundo Global. In: SCHWARTZMAN, S. **Ciência e Tecnologia no Brasil: A Capacitação Brasileira para a Pesquisa Científica e Tecnológica.** Rio de Janeiro: Fundação Getúlio Vargas, 1996. p. 73-92. <http://www.schwartzman.org/simon/scipol/pdf/botanica.pdf>.
- EGLER, I.; SANTOS, M. D. M. ( . ). **Diretrizes e estratégias para a modernização de coleções biológicas brasileiras e a consolidação de sistemas integrados de informação sobre biodiversidade.** Brasília: Centro de Gestão e Estudos Estratégicos, 2006. ISBN CDU 574:32:5/6(81).
- ELZINGA, A. Features of the current science policy regime: Viewed in historical perspective. **Science and Public Policy**, v. 39, n. 4, p. 416-428, 2012. ISSN 1471-5430.
- FORZZA, R. C. et al. New Brazilian Floristic List Highlights Conservation Challenges. **BioScience**, v. 62, n. 1, p. 39-45, January 2012. ISSN 1525-3244. doi:10.1525/bio.2012.62.1.8.
- GADOTTI, M. **História das ideias pedagógicas.** São Paulo: Ática, 2005.
- GARVEY, W. D.; GRIFITH, B. G. Communication and information processing within scientific disciplines: empirical findings. In: W.D., G. **Communication: the essence of science.** [S.l.]: Osford: Pergamon Press, 1979.
- GBIF. **GBIF Annual Report 2010.** [S.l.]. 2011. (87-92020-17-8).
- GBIF. **GBIF Strategic Plan 2012-2016 - Seizing the Future.** [S.l.]. 2011. (87-92020-18-6).

- GIBBONS, M. et al. **The New Production of Knowledge: Dynamics of Science and Research in Contemporary Societies**. London: SAGE Publications Ltd, 1994. 179 p. ISBN 978-0-8039-7794-5.
- GIOVANNI, R. et al. The Real Task of Selecting Records for Ecological Niche Modelling. **Natureza & Conservação**, v. 10, n. 2, p. 139-144, 2012.
- GODDARD, A. et al. **GBIF Position Paper on Data Hosting Infrastructure for Primary Biodiversity Data**. GBIF. Copenhagen, p. 34. 2011. (87-92020-38-0). Acessível em [http://links.gbif.org/gbif\\_position\\_paper\\_data\\_hosting\\_infrastructure\\_primary\\_biodiversity\\_data\\_en\\_v1](http://links.gbif.org/gbif_position_paper_data_hosting_infrastructure_primary_biodiversity_data_en_v1).
- GRAHAM, C. H. et al. New developments in museum-based informatics and applications in biodiversity analysis. **Trends in Ecology and Evolution**, v. 19, n. 9, p. 497-503, September 2004. ISSN 0169-534.
- GRAY, J. et al. Online Scientific Data Curation, Publication, and Archiving. **SPIE Astronomy Telescopes and Instruments**, Waikoloa, August 2002. 6. Disponível em: <http://research.microsoft.com/apps/pubs/default.aspx?id=64568>.
- GRDI2020. **Preliminary Roadmap Report: Global Scientific Data Infrastructures: The Big Data Challenges**. [S.l.]. 2010. (<http://www.grdi2020.eu/StaticPage/About.aspx>).
- GREEN, D. G.; CROFT, J. F. **Proposal for Implementing a Biodiversity Information Network**. Campinas, p. 5-17. 1994.
- HOBBSAWM, E. J. **Era dos extremos: o breve século XX; 1914-1991**. Tradução de Marcos Santarrita. 2. ed. [S.l.]: Companhia das Letras, 2008. 598 p. ISBN 9788571644687.
- INPI. **Guia de Depósitos de Patentes**. Instituto Nacional da Propriedade Industrial. Rio de Janeiro, p. 62. 2008.
- ITL EDUCATION SOLUTIONS LIMITED. **Introduction to Information Technology**. [S.l.]: Dorling Kindersley, 2006. ISBN 9788177581188.
- IUCN. **IUCN Red List Categories and Criteria: Version 3.1**. Gland: IUCN, 2001. 30 p. ISBN 2-8317-0633-5. Disponível em: [http://www.iucnredlist.org/documents/redlist\\_cats\\_crit\\_en.pdf](http://www.iucnredlist.org/documents/redlist_cats_crit_en.pdf).

- JOLY, C. A.; SPEGLICH, É. Programa Biota/Fapesp: um novo paradigma no estudo da conservação e do uso sustentável da biodiversidade. **Ciência e Cultura**, São Paulo, v. 55, n. 3, July/Sept. 2003. ISSN 0009-6725. Disponível em: <[http://cienciaecultura.bvs.br/scielo.php?pid=S0009-67252003000300024&script=sci\\_arttext](http://cienciaecultura.bvs.br/scielo.php?pid=S0009-67252003000300024&script=sci_arttext)>.
- KIM, K. C.; BYRNE, L. B. Biodiversity loss and the taxonomic bottleneck: emerging biodiversity science. **Ecological Research**, v. 21, n. Special Issue, p. 794-810, October 2006.
- LASTRES, H. M. M.; ALBAGLI, S. Introdução: Chaves para ao Terceiro Milênio na Era do Conhecimento. In: LASTRES, H. M. M.; ALBAGLI, S. **Informação e globalização na era do conhecimento**. Rio de Janeiro: Editora Campus Ltda, 1999. p. 7-26. ISBN 85-352-0489-X.
- LATHAM, R.; SASSEN, S. Digital Formations: Constructing an Object of Study. In: SASSEN, R. L. A. S. **Digital Formations: IT and new architectures in the global realm**. Princeton: Princeton University Press, 2005. Cap. 1, p. 1-34. ISBN 0-691-11987-2.
- LEINER, B. M. et al. **Histories of the Internet**. [S.l.]. 1997.
- LEITE, F. C. L.; COSTA, S. Repositórios institucionais como ferramentas de gestão do conhecimento científico no ambiente acadêmico. **Perspectivas em Ciência da Informação**., Belo Horizonte, v. 11, n. 2, p. 206-219, mai./ago. 2006. Disponível em: <<http://www.scielo.br/pdf/%0D/pci/v11n2/v11n2a05.pdf>>.
- LUCAS, C. R. Bases de dados bibliográficas como mapas virtuais de conhecimento. In: GUIMARÃES, E. (. ). **Produção e Circulação do Conhecimento**. Campinas: Pontes Editores, v. II, 2003. Cap. Parte III - 2, p. 9. ISBN CDD 070.4495.
- LUNA, F. J. Frei José Mariano da Conceição Veloso e a divulgação de técnicas industriais no Brasil colonial: discussão de alguns conceitos das ciências químicas, Rio de Janeiro, 16, n. 1, 2009. [http://www.scielo.br/scielo.php?pid=s0104-59702009000100009&script=sci\\_arttext](http://www.scielo.br/scielo.php?pid=s0104-59702009000100009&script=sci_arttext).
- MAMEDE, A. C. H. A experiência do projeto flora fanerogâmica do estado de São Paulo. In: (ORG), A. L. P. **Coleções Biológicas de apoio ao Inventário, Uso Sustentável e Conservação da Biodiversidade**. Rio de Janeiro: Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, v. 1, 2003. Cap. 5, p. 238. ISBN 85-7224-012-8.

- MCDADE, L. A. et al. Biology Needs a Modern Assessment System for Professional Productivity. **BioScience**, v. 61, n. 8, p. 619-625, August 2011. ISSN 1525-3244. doi:10.1525/bio.2011.61.8.8.
- MEADOWS, A. J. **A comunicação científica**. Tradução de Antonio Agenor Briquet de Lemos. Brasília DF: Lemos Informação e Comunicação Ltda., 1999. ISBN 85-85637-15-3.
- MEADOWS, D. H. et al. **The Limits to Growth: a Report for the Club of Rome's Project on the Predicament of Mankind**. New York. 1972.
- MONTEIRO, J. G. B. **Desenhos botânicos da Viagem Filosófica x Flora brasiliensis**. [S.l.]. 2010.
- NIELSEN, A. M. **GBIF Strategic Plan 2012-2016**. Copenhagen, p. 12. 2011. (ISBN: 87-92020-18-6). [http://links.gbif.org/sp2012\\_2016.pdf](http://links.gbif.org/sp2012_2016.pdf).
- NOGUEIRA, E. **Uma história brasileira da Botânica**. Brasília: Paralelo 15 - Marco Zero, 2000. ISBN 85-279-0211-7.
- NOVECK, B. S. **Wiki government: how technology can make government better, democracy stronger, and citizens more powerful**. Washington, DC: Brookings Institution Press, 2009. ISBN 978-0-8157-0275-7.
- NOWOTNY, H.; SCOTT, P.; GIBBONS, M. **Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty**. Cambridge: Polity Press, 2001. 278 p. ISBN 978-0-7456-2608-6.
- NRC, N. R. C. **A New Biology for the 21st Century: Ensuring the United States Leads the Coming Biology Revolution**. Washington DC: The National Academies Press, 2009. ISBN 978-0-309-14486-5.
- NSB. **Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century**. National Science Foundation. Washington DC, p. 9-87. 2005.
- OECD. **Frascati Manual: Proposed Standard Practice for Surveys on Research and Experimental Development**. 6th. ed. Paris: OECD Publication Services, 2002. 266 p. ISBN 978-92-64-19903-9.
- OECD. Science, Technology and Innovation for the 21st Century. Meeting of the OECD Committee for Scientific and Technological Policy at Ministerial Level, 29-30 January 2004 - Final Communique, January 2004. Disponível em: <[http://www.oecd.org/document/15/0,2340,en\\_21571361\\_21590465\\_25998799\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/15/0,2340,en_21571361_21590465_25998799_1_1_1_1,00.html)>. Acesso em: 05 jul. 2011.

- OECD. **OECD Principles and Guidelines for Access to Research Data from Public Funding**. Paris. 2007. <http://www.oecd.org/dataoecd/9/61/38500813.pdf>.
- OECD MEGASCIENCE FORUM. **Final Report of the OECD Megascience Forum Working Group on Biological Informatics**. OECD. Paris, p. 74. 1999. [http://www.gbif.org/fileadmin/Temp\\_for\\_New\\_Web\\_Site/BIrepfin.pdf](http://www.gbif.org/fileadmin/Temp_for_New_Web_Site/BIrepfin.pdf).
- OKF, O. K. F. **The Open Data Manual**. [S.l.]. 2010.
- OLIVO, L. C. C. D. **Reglobalização do Estado e da Sociedade em Rede na Era do Acesso**. Florianópolis: Fundação Boiteux, 2004. 224 p. ISBN 85-87995-25-1.
- PCAST. **Sustaining Environmental Capital: Protecting Society and the Economy**. White House. Washington DC, p. 1-126. 2011.
- PEIXOTO, A. L. **Coleções Biológicas de Apoio ao Inventário, Uso Sustentável e Conservação da Biodiversidade**. Rio de Janeiro: Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, 2003. 238 p. p. ISBN 85-7224-012-8.
- PEIXOTO, F. L.; LIMA, H. C. D. L. **A Informatização de Herbários Brasileiros: estudo de caso**. Jardim Botânico do Rio de Janeiro. Rio de Janeiro, p. 20. 2005.
- PIRANI, J. R. **Sistemática: Tendências e Desenvolvimento, Incluindo Impedimentos para o Avanço do Conhecimento na Área**. [S.l.]. 2005. <http://www.cria.org.br/cgee/documentos/PiraniTextoSistemática.doc>.
- POMBO, O. Interdisciplinaridade e integração dos saberes. **Liinc em Revista**, Rio de Janeiro, v. 1, n. 1, 2005. ISSN 18083536. Disponível em: <<http://revista.ibict.br/liinc/index.php/liinc/article/viewFile/186/103>>.
- PRANCE, G. T. Floristic Inventory of the Topics: Where Do We Stand? **Annals of the Missouri Botanical Garden**, v. 64, n. 4, p. 659-684, 1977. Disponível em: <<http://www.jstor.org/stabel/2395293>>.
- PRIORE, M. D.; VENANCIO, R. **Uma Breve História do Brasil**. São Paulo: Editora Planeta do Brasil Ltda., 2010. ISBN 978-85-7665-529-9.
- REICHMAN, O. J.; JONES, M. B.; SCHILDHAUER, M. P. Challenges and Opportunities of Open Data in Ecology. **Science**, v. 331, p. 703-705, February 2011. ISSN 1095-9203. Disponível em: <[www.sciencemag.org/content/331/6018/703.full.html](http://www.sciencemag.org/content/331/6018/703.full.html)>.

- ROWLEY, J. The wisdom hierarchy: representations of the DIKW hierarchy. **Journal of Information Science**, Chicago, 33, n. 2, 15 February 2007. 163-180. <http://jis.sagepub.com/cgi/content/abstract/33/2/163>.
- RUMBLE, J. **Scientific Discovery and Large-Scale Databases**. [S.l.]. 2005.
- SÁ, M. R. O botânico e o mecenas: João Barbosa Rodrigues e a ciência no Brasil na segunda metade do século XIX, Rio de Janeiro, VIII, n. Suplemento, 2001. <http://www.scielo.br/pdf/hcsm/v8s0/a06v08s0.pdf>.
- SACHS, J. D. From Millennium Development Goals to Sustainable Development Goals. **The Lancet**, v. 379, n. 9832, p. 2206-2211, June 2012. Disponível em: <<http://www.thelancet.com/journals/lancet/article/PIIS0140-6736%2812%2960685-0/fulltext>>.
- SALLES, L. et al. Era uma vez uma Instrução Normativa. **Jornal da Ciência**, p. 3, Março 2007.
- SALLES-FILHO, S.; BIN, A.; FERRO, A. F. P. Abordagens abertas e as implicações para a gestão de C,T&I. **Conhecimento & Inovação**, Outubro/Novembro/Dezembro 2008. ISSN 1984-4395. Disponível em: <<http://www.conhecimentoeinovacao.com.br/materia.php?id=169>>.
- SANJAD, N. **Emílio Goeldi (1859-1917): a ventura de um naturalista entre a Europa e o Brasil**. Rio de Janeiro: Empresa Marketing Cultural - EMC, 2009. ISBN 978-85-87933-22-5.
- SANTOS, A. A. S.; GÓES, J. A. Considerações sobre a preservação da biodiversidade e o papel da Lei nº 11.284/06. **Revista Jus Navigandi**, Teresina, 10, n. 1036, Abril 2006. <http://jus.uol.com.br/revista/texto/8341/consideracoes-sobre-a-preservacao-da-biodiversidade-e-o-papel-da-lei-no-11-284-06>.
- SARUKHÁN K., J.; SECO M., R. M. (. **Conabio - Dos Décadas de História 1992 - 2012**. Conabio. D.F, p. 103. 2012. (978-607-7607-59-5).
- SCAFF, L. M. A lição de coisas do Museu Paraense Emílio Goeldi. In: ARTE, F. N. D. **Museu Paraense Emílio Goeldi**. Rio de Janeiro: FUNARTE, 1981. Cap. 1, p. 12. ISBN CDU 7.031:571/572(81).
- SCHOFIELD, P. N. et al. Sustaining the Data and Bioresource Commons. **Science**, v. 330, p. 592-593, October 2010. Disponível em: <[www.sciencemag.org/cgi/content/full/330/6004/592/DC1](http://www.sciencemag.org/cgi/content/full/330/6004/592/DC1)>.

- SCHROEDER, R. E-Research Infrastructures and Open Science: Towards a New System of Knowledge Production. **Prometheus**, v. 25, n. 1, p. 1-17, March 2007. ISSN 1470-1030.
- SETTE, L. D.; VAZOLLER, R. F. **Visão da Sociedade Brasileira de Microbiologia Frente ao Novo Projeto de Lei sobre Coleta e Uso de Recurso Genético da Biodiversidade Nacional**. [S.l.]. 2010.
- SHAPIN, S. **The Scientific Revolution**. 1998. ed. Chicago: The University of Chicago Press, 1996. ISBN 0226750213.
- SHEPHERD, G. J. **Conhecimento de diversidade de plantas**. Ministério do Meio Ambiente (MMA). Brasília. 2003.
- TAPSCOTT, D.; WILLIAMS, A. D. **Macrowikinomics: rebooting business and the world**. [S.l.]: Penguin Group, 2010. 1-424 p. ISBN 978-1-59184-356-6.
- TAPSCOTT, D.; WILLIAMS, A. D. **Wikinomics - How Mass Collaboration Changes Everything**. New York: Penguin Group, 2010. 351 p. ISBN 978-1-59184-367-2.
- TAYLOR, J. **A viagem do Beagle: A Extraordinária Aventura de Darwin a Bordo do Famoso Navio de Pesquisa do Capitão Fitzroy**. Tradução de Gilson César Cardoso de Sousa. São Paulo: Editora da USP, 2009. 192 p. ISBN 978-85-314-1163-2.
- THE ROYAL SOCIETY SCIENCE POLICY CENTER. **Science as an open enterprise: open data for open science**. The Royal Society. London, p. 1-104. 2012. (978-0-85403-962-3).
- VICKERY, B. C. **Scientific Communication in History**. Lanham: Scarecrow Press, INC, 2000. 255 p. ISBN 0-8108-3598-3.
- VIEIRA, E. **Os Bastidores da Internet no Brasil**. 1a. ed. Barueri, SP: [s.n.], 2003. ISBN 85-204-1708-6.
- WCED. **Our Common Future: Report of the World Commission on Environment and Development**. Oxford. 1987.
- WHEELER, Q. D.; RAVEN, P. H.; WILSON, E. O. Taxonomy: Impediment or Expedient? **Science**, v. 303, n. 5656, p. 285, 16 January 2004. ISSN 1095-9203. Disponível em: <<http://www.sciencemag.org/content/303/5656/285.full>>.
- WISZ, M. S. et al. Effects of sample size on the performance of species distribution models. **Diversity and Distributions A Journal of Conservation Biogeography**, v. 14, n. 5, p. 763-773, March 2008. ISSN 10.1111/j.1472-4642.2008.00482.x.

WOOD, J. et al. **Riding the Wave - How Europe can gain from the rising tide of scientific data**. European Union. [S.l.], p. 36. 2010.

## **ANEXO 1. PESQUISA DE AVALIAÇÃO E PLANEJAMENTO**

Essa pesquisa, realizada no âmbito de tese de doutorado em Política Científica e Tecnológica, do DPCT, Instituto de Geociências da Unicamp, tem por objetivo avaliar um método de planejamento e definição de ações prioritárias para subsidiar a definição de políticas públicas de apoio a coleções biológicas, visando a sua participação em redes de dados de acesso público e aberto. Os dados e informações obtidas serão de uso exclusivo para esse fim.

A pesquisa parte da análise dos provedores de dados da rede *speciesLink* sobre suas próprias coleções. O questionário apresentado foi formulado a partir de dados secundários sobre o processo de planejamento e avaliação interna realizada no Museu Nacional de História Natural Smithsonian (trabalho apresentado pela Dra. Carol Roetzel Butler, Chief of Collections, no Simpósio Latino-Americano de Coleções Biológicas e Biodiversidade: Conhecimento e Gestão organizado pelo Instituto Oswaldo Cruz/Fiocruz); da publicação Diretrizes e estratégias para modernização de coleções biológicas brasileiras e a consolidação de sistemas integrados de informação sobre biodiversidade (CGEE, 20006); e da experiência do CRIA com a rede *speciesLink* que há mais de dez anos está trabalhando com coleções botânicas, zoológicas e microbianas no desenvolvimento de redes de informação de acesso público através da Internet.

No resultado da pesquisa, as coleções não serão identificadas, apenas solicita-se que sejam caracterizadas para avaliar possíveis pontos comuns, de acordo com o tamanho do acervo e grupo taxonômico trabalhado.

### **Questionário**

Todas as respostas deverão partir da avaliação da situação atual da coleção. Ao longo do questionário são apresentados campos livres para comentários. A inclusão desses campos não tem como objetivo conhecer especificidades das coleções, mas receber comentários para incluir eventuais omissões ou corrigir falhas do questionário.

### **IDENTIFICAÇÃO DO ENTREVISTADO**

Nota: nenhuma coleção, pessoa, ou instituição será identificada na análise a ser apresentada. Solicito a identificação caso seja interessante comparar as avaliações de acordo com o perfil do entrevistado.

Nome:

Cargo:

Acrônimo da Coleção:

Caracterização da Instituição:

- Universidade Pública
- Universidade Privada
- Instituição de Pesquisa Pública
- Associação de Interesse Público de Direito Privado
- Empresa de Direito Público
- Empresa de Direito Privado

Outra (especificar):

## CARACTERIZAÇÃO DA COLEÇÃO

### Tipo do Acervo

- Coleção Botânica
- Herbário
- Carpoteca
- Xiloteca
- Banco de DNA
- Coleção Zoológica
- Vertebrados
- Invertebrados Terrestres
- Invertebrados Aquáticos

- Coleção Microbiana
- Arqueias
- Bactérias
- Fungos Filamentosos
- Leveduras
- Protozoários
- Coleção de Fósseis

### Tamanho do Acervo (classificar por número de registros)

- até um mil registros
- entre um mil a 5 mil registros
- entre 5 mil a 50 mil registros
- entre 50 mil a 100 mil registros
- entre 100 mil a 200 mil registros
- entre 200 mil a 500 mil registros
- entre 500 mil a um milhão de registros
- acima de um milhão de registros

### Avaliação da Coleção

As questões a seguir irão contribuir para a avaliação das necessidades e definição das ações prioritárias de apoio às coleções biológicas participantes da rede *speciesLink*.

### **Recursos Humanos**

Não se trata da avaliação das pessoas, mas da disponibilidade de pessoal qualificado para o cargo em número adequado.

#### **Curadoria:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

#### **Taxonomia/Sistemática:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

#### **Apoio Técnico:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

#### **Pesquisa:**

- Muito fraco ou ausente
- Fraco ou insuficiente

- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

Comentários:

### **Infraestrutura Física**

#### **Recebimento e Preparo de Material:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

#### **Armazenamento:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

#### **Acesso e Manipulação de Material:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

**Segurança:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

Comentários:

**Informatização****Processo de Documentação:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

**Equipe****Digitação de dados:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

**Digitalização das Amostras:**

- Muito fraco ou ausente
- Fraco ou insuficiente

- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

**Equipe de Apoio em Informática:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

**Equipamento****Digitação e Armazenamento de Dados:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

**Digitalização das Imagens:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

**Serviços Internet:**

- Muito fraco ou ausente
- Fraco ou insuficiente

- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

Comentários:

### **Atividades**

#### **Gestão da Coleção:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

#### **Intercâmbio:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

#### **Pesquisa:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

#### **Ensino:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

#### **Serviços:**

##### **Demanda Interna:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

##### **Demanda Externa:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

Comentários:

### **Governança**

#### **Reconhecimento Institucional:**

- Muito fraco ou ausente
- Fraco ou insuficiente

- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

**Fixação de Recursos Humanos:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

**Orçamento Anual:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

**Planejamento Estratégico:**

- Muito fraco ou ausente
- Fraco ou insuficiente
- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

**Apoio Externo a Projetos:**

- Muito fraco ou ausente
- Fraco ou insuficiente

- Médio ou suficiente
- Forte ou bom
- Muito forte ou muito bom

Comentários:

**Ações Prioritárias**

De acordo com a sua experiência em relação à sua coleção, favor classificar as ações a seguir como sendo desde de prioridade máxima a não prioritária.

**Ampliação do Acervo**

- Prioridade máxima
- Prioritário
- Medianamente prioritário
- Baixa prioridade
- Não é prioritário

**Fixação de Recursos Humanos**

- Prioridade máxima
- Prioritário
- Medianamente prioritário
- Baixa prioridade
- Não é prioritário

**Formação de recursos humanos**

- Prioridade máxima
- Prioritário
- Medianamente prioritário

Baixa prioridade

Não é prioritário

#### **Informatização e limpeza de dados**

Prioridade máxima

Prioritário

Medianamente prioritário

Baixa prioridade

Não é prioritário

#### **Interação com outras coleções**

Prioridade máxima

Prioritário

Medianamente prioritário

Baixa prioridade

Não é prioritário

#### **Investimento na capacitação em gestão da coleção**

Prioridade máxima

Prioritário

Medianamente prioritário

Baixa prioridade

Não é prioritário

#### **Melhoria da infraestrutura**

Prioridade máxima

Prioritário

Medianamente prioritário

Baixa prioridade

Não é prioritário

#### **Planejamento**

Prioridade máxima

Prioritário

Medianamente prioritário

Baixa prioridade

Não é prioritário

#### **Reconhecimento institucional**

Prioridade máxima

Prioritário

Medianamente prioritário

Baixa prioridade

Não é prioritário

**Comentários:**



## ANEXO 2. E-MAILS ENVIADOS

### ***E-mail* enviado no dia 04 de junho de 2012 aos curadores, diretores e chefes de departamento das coleções biológicas da rede *speciesLink*:**

Prezados curadores, diretores e chefes de departamento das coleções biológicas da rede *speciesLink*,

Estou desenvolvendo a minha tese de doutorado em política científica e tecnológica sobre sistemas de informação em biodiversidade, de acesso livre e aberto, e a formulação de políticas públicas na era digital. Gostaria de testar uma metodologia que parte da análise dos próprios provedores de dados, utilizando como exemplo as coleções biológicas da rede *speciesLink*. Para tanto preparei um questionário cujo preenchimento está sendo solicitado somente para os curadores e diretores ou chefes de departamento das coleções participantes da rede *speciesLink*.

O questionário apresentado foi formulado a partir de dados secundários sobre o processo de planejamento e avaliação interna realizada no Museu Nacional de História Natural Smithsonian (trabalho apresentado pela Dra. Carol Roetzel Butler, Chief of Collections, no Simpósio Latino-Americano de Coleções Biológicas e Biodiversidade: Conhecimento e Gestão organizado pelo Instituto Oswaldo Cruz/Fiocruz); da publicação *Diretrizes e estratégias para modernização de coleções biológicas brasileiras e a consolidação de sistemas integrados de informação sobre biodiversidade* (CGEE, 2006); e da experiência do CRIA com a rede *speciesLink*, que há mais de dez anos está trabalhando com coleções botânicas, zoológicas e microbianas no desenvolvimento de redes de informação de acesso público através da Internet.

Os itens a serem avaliados incluem:

- Recursos humanos (curadoria, taxonomia/sistemática, apoio técnico, pesquisa)
- Infraestrutura física (recebimento e preparo de material, armazenamento, acesso e manipulação de material, segurança)
- Informatização (processo e documentação, equipe, equipamento, serviços Internet)
- Atividades (gestão da coleção, intercâmbio, pesquisa, ensino, serviços)
- Governança (reconhecimento institucional, fixação de recursos humanos, orçamento anual, planejamento estratégico, apoio externo a projetos)
- Ações prioritárias

Todos os itens serão avaliados de acordo com uma escala que varia entre "muito fraco ou ausente" a "muito forte ou muito bom". O preenchimento do questionário deve ser muito simples e espero poder contar com a experiência e reflexão de cada um de vocês para que a análise resultante possa mostrar o valor da avaliação participativa.

Encaminho anexo o ofício da Dra. Maria Beatriz Machado Bonacelli, professora do Departamento de Política Científica e Tecnológica do Instituto de Geociências da Unicamp que é quem está orientando esse trabalho.

O questionário está disponível *on-line* no endereço: <http://www.cria.org.br/~dora> e peço que ele seja encaminhado até o dia 15 de junho pf.

Muito obrigada,

Dora Ann Lange Canhos

Aluna de doutorado do

Programa de Pós-graduação em Política Científica e Tecnológica

DPCT/IG/Unicamp

***E-mail* enviado no dia 13 de junho de 2012:**

No dia 04 de junho pp enviei um *e-mail* aos curadores, diretores e chefes de departamento responsáveis pelas coleções biológicas participantes da rede *speciesLink* solicitando o preenchimento de um questionário *on-line* para a minha tese de doutorado em política científica e tecnológica. Agradeço a todos que responderam e solicito àqueles que têm interesse em responder e ainda não o fizeram, que o façam até o final desta semana. O referido questionário está disponível na página <http://www.cria.org.br/~dora>.

Novamente muito obrigada,

Dora Ann Lange Canhos

Aluna de doutorado do

Programa de Pós-graduação em Política Científica e Tecnológica

DPCT/IG/Unicamp

### ANEXO 3. CARTA DA ORIENTADORA



**Universidade Estadual de Campinas  
Instituto de Geociências  
Departamento de Política Científica e Tecnológica**



Caixa Postal 6152 – CEP 13083-970 Campinas – SP  
Tel/Fax: +55 19 3521 4555 Tel: +55 3521 4597  
e-mail: dpct@ige.unicamp.br

---

Campinas, maio de 2012

Aos  
Curadores, Responsáveis e Técnicos  
das Coleções Biológicas da rede speciesLink  
N E S T A

Prezados Senhores e Senhoras

A aluna de doutorado do Programa de Pós-graduação em Política Científica e Tecnológica do Departamento de mesmo nome (DPCT), do Instituto de Geociências da Unicamp, Dora Ann Lange Canhos, está desenvolvendo sua pesquisa acadêmica de dissertação sobre o tema sistemas de informação de acesso livre e aberto e a formulação de políticas públicas na era digital.

Para que o trabalho ganhe maior confiabilidade e densidade, solicitamos sua valiosa contribuição a partir da resposta a um questionário simples de avaliação das coleções biológicas brasileiras participantes da rede speciesLink.

As informações apresentadas não serão de forma alguma utilizadas individualmente, nem o nome da instituição aparecerá no trabalho relacionada a algum número ou dado.

O trabalho final poderá ser enviado à instituição caso haja interesse.

Reforçamos assim nossa solicitação de ajuda.

Na esperança de contar com a sua atenção, nos colocamos à disposição para qualquer esclarecimento que se faça necessário.

Atenciosamente,

Maria Beatriz Machado Bonacelli  
Profa. DPCT/IG/Unicamp  
Orientadora



#### ANEXO 4. TABELAS ORIGINAIS COM AS RESPOSTAS DAS COLEÇÕES BIOLÓGICAS

As colunas de identificação da instituição e da pessoa responsável pelo preenchimento do formulário foram omitidas, assim como os comentários e observações, uma vez que esses, em sua grande maioria também citavam a instituição.

Anexo 4. 1. Caracterização da Instituição

Nº.	Data - hora	Universidade		Instituto de Pesquisa	Associação	Empresa		Outra
		Pública	Particular			Pública	Privada	
1	15-06-2012-14-34-34						1	
2	13-06-2012-15-15-05					1		
3	04-06-2012-10-56-08	1						
4	04-06-2012-11-46-23							1
5	04-06-2012-12-04-12	1						
6	04-06-2012-12-09-25	1						
7	04-06-2012-13-38-07						1	
8	04-06-2012-14-46-52	1						
9	04-06-2012-16-49-05	1						
10	04-06-2012-17-06-27			1				
11	04-06-2012-21-41-08	1						
12	05-06-2012-10-45-29		1					
13	05-06-2012-10-48-53	1						
14	05-06-2012-12-08-08	1						
15	05-06-2012-12-55-04	1						
16	05-06-2012-16-13-56							

Nº.	Data - hora	Universidade		Instituto de Pesquisa	Associação	Empresa		Outra
		Pública	Particular			Pública	Privada	
17	05-06-2012-17-16-43			1				
18	06-06-2012-11-57-24							1
19	06-06-2012-15-45-33	1						
20	07-06-2012-14-45-47	1						
21	09-06-2012-18-28-26	1						
22	11-06-2012-17-56-07	1						
23	13-06-2012-14-31-42	1						
24	13-06-2012-14-52-27	1						
25	13-06-2012-18-55-16	1						
26	13-06-2012-19-57-34	1						
27	14-06-2012-10-42-17			1				
28	14-06-2012-14-35-47	1						
29	14-06-2012-15-45-07	1						
30	14-06-2012-16-13-56		1					
31	14-06-2012-16-14-40	1						
32	14-06-2012-16-37-01	1						
33	14-06-2012-23-19-46	1						
34	15-06-2012-07-23-50							1
35	15-06-2012-12-08-49		1					1
36	15-06-2012-13-51-53	1						
37	15-06-2012-18-12-07	1						

Nº.	Data - hora	Universidade		Instituto de Pesquisa	Associação	Empresa		Outra
		Pública	Particular			Pública	Privada	
38	18-06-2012-10-32-05						1	
39	19-06-2012-16-39-11	1						
40	20-06-2012-07-45-10	1						
41	20-06-2012-12-38-07			1				
42	05-06-2012-22-55-24							1
43	06-06-2012-11-06-25							1
44	04-06-2012-11-04-30			1				
45	04-06-2012-11-15-28	1						
46	04-06-2012-11-23-37	1						
47	04-06-2012-11-58-36	1						
48	04-06-2012-12-47-14	1						
49	04-06-2012-14-09-45			1				
50	04-06-2012-14-31-45	1						
51	05-06-2012-10-33-03			1				
52	05-06-2012-15-05-46			1				
53	05-06-2012-16-35-33			1				
54	05-06-2012-18-00-48		1					
55	06-06-2012-12-12-21			1				
56	07-06-2012-18-33-04	1						
57	09-06-2012-13-49-06	1						
58	10-06-2012-20-46-40		1					

Nº.	Data - hora	Universidade		Instituto de Pesquisa	Associação	Empresa		Outra
		Pública	Particular			Pública	Privada	
59	11-06-2012-10-17-49		1					
60	11-06-2012-14-42-00			1				
61	11-06-2012-14-44-48			1				
62	13-06-2012-13-50-13		1					
63	13-06-2012-14-02-25	1						
64	13-06-2012-14-22-05		1					
65	13-06-2012-14-59-53			1				
66	13-06-2012-15-10-51	1						
67	13-06-2012-16-07-25	1						
68	13-06-2012-16-17-14	1						
69	13-06-2012-16-36-04					1		
70	13-06-2012-16-40-21	1						
71	13-06-2012-17-49-27	1						
72	14-06-2012-00-17-06	1						
73	14-06-2012-09-26-19	1						
74	14-06-2012-14-28-16			1				
75	14-06-2012-16-00-11	1						
76	14-06-2012-16-53-08		1					
77	14-06-2012-18-32-09	1						
78	15-06-2012-11-41-29	1						
79	15-06-2012-12-56-22			1				

Nº.	Data - hora	Universidade		Instituto de Pesquisa	Associação	Empresa		Outra
		Pública	Particular			Pública	Privada	
80	15-06-2012-14-16-35			1				
81	15-06-2012-16-51-25	1						
82	17-06-2012-17-28-23	1						
83	17-06-2012-17-35-34	1						
84	19-06-2012-17-54-32	1						
85	05-06-2012-12-57-03			1				
86	05-06-2012-15-03-29			1				
87	08-06-2012-10-01-42			1				
88	08-06-2012-11-35-35			1				
89	11-06-2012-15-49-00			1				
90	11-06-2012-22-39-34			1				
91	11-06-2012-23-03-34			1				
92	12-06-2012-12-35-46			1				
93	12-06-2012-13-25-33			1				
94	12-06-2012-16-21-39			1				
95	13-06-2012-15-24-51			1				
96	13-06-2012-15-31-49			1				
97	14-06-2012-09-39-14			1				
98	14-06-2012-10-25-54			1				
99	19-06-2012-14-15-50			1				
	<b>Total</b>	<b>48</b>	<b>9</b>	<b>31</b>	<b>0</b>	<b>2</b>	<b>3</b>	<b>6</b>

Anexo 4. 2. Caracterização da Coleção

	Botânica	sub botânica	Zoológica	sub zoo	Microbiana	sub micro	Fósseis	Tamanho
1	1	herbario;Carpoteca;Xiloteca		InvertTerr				3
2	1	herbario;Carpoteca				fungos		4
3	1	herbario						3
4	1	herbario;Carpoteca;Xiloteca						3
5	1	herbario						3
6	1	Xiloteca						1
7	1	herbario;Carpoteca;Xiloteca						3
8	1	herbario						3
9	1	herbario;Carpoteca						2
10	1	herbario						5
11	1	herbario						3
12	1	herbario						3
13	1	herbario						2
14	1	herbario						3
15	1	herbario						3
16	1	herbário						2
17	1	herbario;Carpoteca;Xiloteca						6
18	1	herbario;Carpoteca						3
19	1	herbario						5
20	1	herbario						7
21	1	herbario						3
22	1	herbario;Carpoteca;Xiloteca						2
23	1	herbario						3
24	1	herbario;DNA						2
25	1	herbario						3
26	1	herbario;Carpoteca;Xiloteca						3
27	1	herbario;Carpoteca;Xiloteca						6

	Botânica	sub botânica	Zoológica	sub zoo	Microbiana	sub micro	Fósseis	Tamanho
28	1	herbario;Carpoteca						4
29	1	herbario;Carpoteca						3
30	1	herbario;Carpoteca						3
31	1	herbario						4
32	1	herbario						3
33	1	herbario						5
34	1	herbario;Carpoteca						4
35	1	herbario						3
36	1	herbario						3
37	1	herbario						5
38	1	herbario;Carpoteca;Xiloteca						3
39	1	herbario						3
40	1	herbario;Carpoteca						3
41	1	herbario						4
42	1	herbario						3
43	1	Xiloteca						3
44			1	Vertebrados				3
45			1	Vertebrados				2
46			1	Invertaqua				6
47			1	Vertebrados				3
48			1	Vertebrados				3
49			1	InvertTerr				3
50			1	InvertTerr				8
51			1	Invertaqua				3
52			1	Invertaqua				3
53			1	Vertebrados				3
54			1	InvertTerr				3
55			1	InvertTerr;Invertaqua				8
56			1	InvertTerr				6

	Botânica	sub botânica	Zoológica	sub zoo	Microbiana	sub micro	Fósseis	Tamanho
57			1	Vertebrados				3
58			1	InvertTerr				3
59			1	Vertebrados				3
60			1	InvertTerr				3
61			1	InvertTerr				2
62			1	InvertTerr				3
63			1	Vertebrados				1
64			1	Vertebrados				3
65			1	Vertebrados				3
66			1	InvertTerr				4
67			1	Vertebrados				3
68			1	Vertebrados				2
69			1	InvertTerr;Invertaqua				3
70			1	Invertaqua				4
71			1	Vertebrados				3
72			1	InvertTerr;Invertaqua				
73			1	Vertebrados			1	2
74			1	Invertaqua				3
75			1	InvertTerr				2
76			1	Vertebrados				3
77			1	InvertTerr				3
78			1	InvertTerr				2
79			1	InvertTerr;Invertaqua				3
80			1	InvertTerr				8
81			1	Vertebrados				3
82			1	Vertebrados				3
83			1	Vertebrados				6
84			1	Vertebrados;InvertTerr;Invertaqua				6
85					1	protozoarios		1

	Botânica	sub botânica	Zoológica	sub zoo	Microbiana	sub micro	Fósseis	Tamanho
86					1	bacteria		1
87					1	bacteria		1
88					1	bacteria		2
89					1	bacteria		2
90					1	bacteria		2
91					1	protozoarios		2
92					1	bacteria		2
93					1	bacteria		2
94					1	fungos		1
95					1	fungos		2
96					1	fungos		1
97					1	bacteria		2
98					1	protozoarios		1
99					1	bacteria		1
<b>Total</b>	<b>43</b>		<b>41</b>		<b>15</b>			

Anexo 4. 3. Avaliação dos Recursos Humanos e Infraestrutura Física

No.	Recursos Humanos				Infraestrutura física			
	Curadoria	Taxonomia	Técnico	Pesquisa	Preparo	Armazenamento	Acesso	Segurança
1	3	3	3	2	3	4	3	5
2	4	3	3	2	3	4	3	3
3	3	1	2	3	4	4	5	4
4		1	3	2	2	4	4	4
5	3	2	2	2	2	3	2	2
6	3	3	1	3	1	3	3	1
7	3	3	3	3	4	3	3	4
8	2	2	1	2	3	3	3	2
9	4	2	1	3	3	4	3	2

No.	Recursos Humanos				Infraestrutura física			
	Curadoria	Taxonomia	Técnico	Pesquisa	Preparo	Armazenamento	Acesso	Segurança
10	3	2	2	2	3	2	3	3
11	3	2	4	3	5	5	5	4
12	3	3	2	3	3	2	3	2
13	3	3	2	3	3	3	3	3
14	1	2	1	2	2	2	2	4
15	3	2	2	3	3	3	3	1
16	4	3	4	4	4	4	4	3
17	5	5	4	5	5	3	5	4
18	4	2	2	2	5	5	5	4
19	3	3	4	3	4	4	4	3
20	3	2	2	4	3	4	3	3
21	4	3	2	3	2	2	3	1
22	2	2	2	3	3	3	3	2
23	4	3	3	3	4	5	4	4
24	3	1	2	2	2	3	2	2
25	3	2	2	3	3	3	3	3
26	4	2	2	3	3	4	4	4
27	3	2	2	2	3	3	5	2
28	3	2	2	2	3	3	3	2
29	2		2	2	2	3	3	3
30	4	3	3	3	5	4	5	4
31	2	3	1	3	2	2	2	3
32	3	2	2	2	1	2	2	3
33	5	3	1	5	2	3	3	1
34	3	2	2	2	3	4	3	3
35	3	3	3	3	3	3	3	3
36	2	2	1	2	3	3	3	2
37	5	5	5	4	5	4	5	4

No.	Recursos Humanos				Infraestrutura física			
	Curadoria	Taxonomia	Técnico	Pesquisa	Preparo	Armazenamento	Acesso	Segurança
38	5	2	3	3	4	4	4	5
39	3	3	2	4	4	5	4	5
40	4	2	2	3	3	4	3	2
41	2	2	1	2	2	2	2	3
42	2	2	1	3	3	3	4	4
43	3	2	3	2	3	3	4	2
44	2	2	1	2	2	2	4	4
45	2	4	2	4	2	3	4	3
46	2	2	1	4	1	2	3	2
47	3	3	1	3	3	3	4	3
48	3	3	2	3	4	3	3	2
49	4	4	3	4	3	1	2	2
50	2	3	1	3	3	4	4	4
51	1	2	1	3	3	4	3	2
52	5	5	4	4	4	4	4	4
53	2	2	2	2	2	2	2	1
54	4	4	3	4	5	5	5	5
55	2	3	2	4	2	4	3	3
56	4	5	5	5	5	5	4	4
57	1	5	1	5	3	3	3	2
58	2	2	2	3	4	4	4	4
59	4	5	2	4	5	4	5	4
60	5	4	3	5	2	3	3	2
61	4	5	5	5	5	5	5	4
62	2	2	2	2	2	4	3	4
63	4	2	1	4	3		3	3
64	3	4	3	4	3	3	5	5
65	3	2	3	3	3	2	1	1

No.	Recursos Humanos				Infraestrutura física			
	Curadoria	Taxonomia	Técnico	Pesquisa	Preparo	Armazenamento	Acesso	Segurança
66	2	3	2	3	4	4	3	4
67	2	3	2	4	3	2	3	3
68	3	3	2	4	3	4	4	3
69	2	2	2	2	3	3	3	2
70	3	4	3	3	4	4	4	4
71	2	1	1	2	2	2	2	1
72	1	3	1	3	1	1	1	1
73	2	2	1	2	2	2	2	2
74	2	3	3	4	4	2	3	2
75	2	2	1	3	2	3	2	3
76	3	4	2	3	4	4	4	3
77	3	2	2	3	3	3	3	3
78	2	1	2	3	3	3	3	2
79	2	2	2	2	4	3	4	2
80	3	2	2	3	4	4	4	4
81	5	5	1	3	2	2	3	2
82	1	3	1	3	4	4	4	2
83	2	4	1	4	4	3	4	2
84	3	3	2	4	3	3	3	3
85	3	3	3	4	2	2	2	2
86	5	3	4	3	4	3	4	4
87	3	3	2	3	3	2	3	4
88	4	3	3	2	2	2	2	2
89	4	4	2	3	4	3	4	3
90	2	4	2	3	3	2	2	2
91	5	5	3	5	3	4	4	4
92	2	4	2	4	3	3	4	3
93	3	3	3	4	3	3	3	2

No.	Recursos Humanos				Infraestrutura física			
	Curadoria	Taxonomia	Técnico	Pesquisa	Preparo	Armazenamento	Acesso	Segurança
94	2	2	1	2	2	3	3	3
95	4	1	3	4	3	4	4	3
96	3	3		1	2			3
97	4	3	2	5	5	5	5	3
98	4	5	4	5	4	4	4	3
99	4	4	3	4	4	4	4	4

Anexo 4. 4. Avaliação da Informatização e outras Atividades

No.	Informatização							Atividades					
	Documen- tação	Equipe			Equipamento		Serviços Internet	Gestão	Intercâmbio	Pesquisa	Ensino	Serviços	
		Digitação	Digitalização	Apoio	Digitação	Digitalizacão						Interna	Externa
1	3	2	3	1	2	2	3	3	2	2	2	3	3
2	3	3	2	2	4	1	3	3	3	3	2	3	3
3	5	5	5	4	5	5	4	4	3	3	2	3	2
4	5	5	1	3	5	1	5	3	2	2	4	3	3
5	3	2	2	2	2	1	3	3		2	2	2	2
6	1	1	1	1	1	1	2	1	1	2	3	4	1
7	3	3	3	3	4	3	4	4	4	4	4	3	3
8	3	3	1	1	2	1	2	2	2	2	3	2	2
9	2	1	1	1	4	1	2	3	2	3	3	3	3
10	3	1	1	3	4	2	4	3	3	2	3	4	5
11	5	5	5	4	4	2	4	4	3	3	4	3	3
12	3	3	2	2	2	2	3	3	2	3	4	2	2
13	2	1	1	1	1	1	1	3	3	3	4	2	
14	1	1	1	1	2	1	3	2		3	3	3	3
15	2	2	3	3	4	3	3	3	3	2	3	4	3
16	4	4	4	3	4	4	3	4	4	4	3	3	4

No.	Informatização							Atividades					
	Documen- tação	Equipe			Equipamento		Serviços Internet	Gestão	Intercâmbio	Pesquisa	Ensino	Serviços	
		Digitação	Digitalização	Apoio	Digitação	Digitalizacão						Interna	Externa
17	3	2	2	3	4	5	3	5	5	5	4	5	5
18	4	3	2	3	3	3	3	3	3	2	5	4	5
19	3	3	3	2	4	3	3	3	4	3	3	4	4
20	2	2	2	1	3	1	2	4	3	4	4	2	2
21	3	3	3	1	3	1	4	3	4	3	3	4	3
22	2	2	2	2	3	1	4	2	2	2	3	3	2
23	4	4	1	5	4	1	4	5	5	2	3	4	3
24	3	2	2	1	2	2	2	2	2	2	2	3	2
25	3	2	1	1	2	1	3	3	3	3	3	4	4
26	3	2	2	3	2	1	4	3	3	3	4	4	4
27	2	2	2	2	2	2	3	4	5	2	1	3	3
28	3	2	2	2	3	1	3	3	3	3	5	5	4
29	2	2	2	1	3	2	3	3	2	3	3	3	3
30	3	2	2	2	3	2	5	4	4	3	3	3	3
31	2	2	1	2	2	1	3	3	3	3	3	1	1
32	3	3	2	2	3	1	3	3	3	2	3	2	3
33	4	2	3	1	2	4	4	3	4	5	4	5	5
34	2	2	2	2	2	1	2	3	2	2	2	3	3
35	3	2	2	3	3	3	3	3	3	3	3	3	3
36	2	1	1	1	2	1	3	2	2	2	3	1	1
37	4	4	4	3	3	4	2	4	5	4	4	4	4
38	4	3	3	3	4	4	4	4	2	3	2	3	3
39	4	4	1	5	4	3	4	5	4	4	4	4	3
40	2	1	1	1	3	3	4	4	3	3	3	3	3
41	5	1	1	1	2	1	3	4	3	3	2	1	3
42	2	1	1	2	3	2	2	3	3	4	3	5	2
43	4	3	2	3	4	2	2	3	2	2	3	2	4

No.	Informatização							Atividades					
	Documen- tação	Equipe			Equipamento		Serviços Internet	Gestão	Intercâmbio	Pesquisa	Ensino	Serviços	
		Digitação	Digitalização	Apoio	Digitação	Digitalizacão						Interna	Externa
44	2	2	3	2	3	1	4	3	3	3	2	3	3
45	4	4	4	1	4	1	4	2	2	5	5	4	3
46	4	2	2	1		2	3	2	1	4	3	2	1
47	4	2	2	2	3	3	3	3	3	4	3	3	3
48	3	2	2	1	3	3	3	3	3	3	3	3	3
49	3	3	2	2	2	1	2	2	2	3	3	3	3
50	2	1	1	1	1	1	3	3	4	4	4	2	4
51	3	2	1	2	4	1	3	3	3	4	3	3	3
52	4	4	4	4	4	3	4	4	3	4	4	4	1
53	3	3	3	2	2	2	2	2	2	2	2	3	2
54	4	4		3	3	1	3	4	4	4	4		
55	2	2	2	2	3	2	3	3	4	4	3	4	4
56	4	4	2	3	5	4	5	5	3	5	5	4	4
57	4	2	2	2	4	2	4	3	2	5	4	3	3
58	4	2	2	3	3	1	4	2	3	4	3	3	3
59	5	5	5	3	3	2	5	5	4	4	4	4	4
60	4	3	3	5	4	4	4	5	1	5	4		3
61	4	3	1	3	4	1	5	5	4	5	4	5	4
62	4	3	1	1	2	1	1	2	2	3	1	2	1
63	3	2	2	1	3	3	1	3	2	3	4	3	3
64	3	3	3	2	3	3	4	5	4	5	4	3	3
65	4	4	4	4	4	4	2	4	3	3	3	5	5
66	3	3	3	2	3	2	3	3	3	4	3	3	3
67	2	2	2	2	3	1	2	2	3	3	3	3	2
68	2	1	1	1	1	2	4	3	3	4	4	4	4
69	2	2	2	1	3	1	2	2	1	2	2	2	2
70	4	3	3	3	4	5	5	4	4	4	4		

No.	Informatização							Atividades					
	Documen- tação	Equipe			Equipamento		Serviços Internet	Gestão	Intercâmbio	Pesquisa	Ensino	Serviços	
		Digitação	Digitalização	Apoio	Digitação	Digitalizacão						Interna	Externa
71	1	1	1	1	1	1	2	2	2	2	2	1	2
72	1	1	1	1	2	1	3	1	2	2	2	2	2
73	3	1	1	1	3	1	2	3	2	2	2	3	3
74	4	4	2	3	4	2	3	4	3	4	4	4	4
75	2	2	2	1	3	1	1	2	2	2	1	1	1
76	5	5	3	3	5	4	4	5	2	3	4	4	2
77	3	2	2	3	3	2	3	3	3	3	3	3	3
78	1	1	1	1	2	1	1	2	2	3	3	2	1
79	2	2	1	1	1	1	2	4	3	5	5	5	5
80	2	2	2	2	4	2	2	4	4	4	4	4	4
81	3	3	1	1	1	1	2	5	3	3	1	3	3
82	4	2	2	1	3	2	3	2	2	2	2	2	3
83	3	2	2	1	3	2	2	3	4	4	4	5	5
84	4	2	2	3	4	1	3	2	3	3	3	3	2
85	3	3	2	2	2	2	2	2	4	4	3	3	3
86		1	2	1	1	1	3	2	3	3	3	3	3
87	2	2		3	4	4	4	3	1	3	1	4	1
88	3	2	2	2	1	1	3	3	1	2	2	3	3
89	3	2	2	2	2	3	4	4	4	3	4	3	4
90	2	2	2	3	3	3	2	3	4	3	3	3	4
91	4	4		2	4		4	5	3	5	5	4	4
92	3	2	1	1	2	1	3	3	3	5	5	5	4
93	2	2	3	1	1	1	3	3	2	3	3	3	3
94	2	1	1	1	3	1	3	2	2	2	3	3	2
95	2	2	2	2	3	2	3	3	3	4	4	3	3
96	3	2	2	2	3	2	3	3		3	3	3	3
97	3	3	3	1	3	3	4	3	4	5	5	4	5

No.	Informatização						Atividades						
	Documen- tação	Equipe			Equipamento		Serviços Internet	Gestão	Intercâmbio	Pesquisa	Ensino	Serviços	
		Digitação	Digitalização	Apoio	Digitação	Digitalizacão						Interna	Externa
98	3	3	3	3	3	2	2	3	3	5	3	3	3
99	2	3	3	2	3	2	3	4	2	3	3	4	2

Anexo 4. 5. Avaliação da Governança

No.	Reconhecimento Institucional	Fixação de Recursos humanos	Orçamento	Planejamento Estratégico	Apoio Externo a Projetos
1	3	5	4	3	
2	2	1	1	3	2
3	3	2	2	3	
4	2	2	3	3	2
5	2	1	1	2	2
6	2	1	1	1	1
7	5	3	3	3	4
8	1	1	1	2	1
9	2	2	2	3	
10	5	2	3	3	2
11	4	4	3	3	4
12	2	2	1	3	
13	3	3	1	3	
14	1	1	1	3	
15	2	2	1	2	4
16	4	2	4	4	4
17	4	3	3	3	4
18	4	3	2	3	
19	4	4	3	1	1
20	4	1	2	2	
21	2	3	1	3	

No.	Reconhecimento Institucional	Fixação de Recursos humanos	Orçamento	Planejamento Estratégico	Apoio Externo a Projetos
22	3	2	2	2	2
23	4	3	5	4	4
24	2	2	2	2	2
25	3	3	2	3	
26	3	2	2	3	2
27	5	1	2	2	2
28	5	3	1	3	4
29	1	2	1	2	2
30	3	3	2	3	2
31	1	1	1	1	2
32	1	1	1	2	2
33	2	2	1	3	
34	3	1	2	3	2
35	3	2	3	3	2
36	1	1	1	1	1
37	4	4	3	4	4
38	2	2	2	3	2
39	4	3	5	4	4
40	4	3	3	3	4
41	1	1	1	2	2
42	3	1	2	3	
43	1	2	1	2	2
44	3	1	1	2	2
45	1	1	1	3	
46	1	1	1	3	2
47	2	1	1	1	1
48	3	2	1	3	
49	2	2	2	3	2
50	2	1	1	3	

No.	Reconhecimento Institucional	Fixação de Recursos humanos	Orçamento	Planejamento Estratégico	Apoio Externo a Projetos
51	4	2	1	3	2
52	5	4	3	4	4
53	1	1	1	1	1
54	4	3	3		2
55	2	2	2	2	2
56	5	3	2	4	5
57	3	1	1	1	2
58	2	1	2	3	
59	4	3	4	4	4
60	5	2	2	4	5
61	5	4	4	4	4
62	2	1	1	1	2
63	1	1	1	1	1
64	3	3	2	3	
65	3	2	2	4	4
66	4	3	3	2	
67	3	3	2	3	
68	3	2	2	3	
69	2	2	2	2	1
70	5	3	4	4	4
71	2	1	1	1	1
72	2	1	1	1	1
73	1	1	1	2	2
74	4	2	2	3	4
75	1	1	1	1	2
76	3	2	3	4	2
77	3	2	1	2	2
78	1	1	1	1	1
79	5	1	2	3	

No.	Reconhecimento Institucional	Fixação de Recursos humanos	Orçamento	Planejamento Estratégico	Apoio Externo a Projetos
80	5	2	2	3	2
81	1	1	1	1	1
82	1	1	1	1	2
83	1	1	1	2	4
84	4	2	2	2	4
85	4	2	2	3	5
86	5	3	3	3	2
87	4	1	1	2	1
88	4	3	3	3	
89	4	2	2	3	2
90	2	1	1	2	1
91	5	3	3	3	4
92	4	3	2	3	
93	4	2	3	3	2
94	4	2	2	2	2
95	3	2	3	3	2
96	4	2	3	2	2
97	3	1	1	3	4
98	4	1	3	3	2
99	2	2	2	3	

Anexo 4. 6. Avaliação das Ações Prioritárias

No.	Ampliação do Acervo	Fixação de Recursos Humanos	Formação de Recursos Humanos	Informatização e Limpeza de Dados	Interação com outras coleções	Investimento na capacitação em gestão da coleção	Melhoria da Infraestrutura	Planejamento	Reconhecimento Institucional
1	2	5	3	2	2	1	1	1	4
2	1	1	2	1	1	1	1	1	4
3	2	1	3	4	3	3	4	3	2
4	3	4	4	1	3	4	3	3	3
5	2	1	2	3	2	2		2	2
6	2	1		1	1	1	1	1	1
7	1	2	2		1	2	2	2	2
8	3	2	3	1	3	2	2	2	1
9	2	3	3	2	3	3	3	2	2
10	5	1	2	1	3	3	1	2	5
11	1	2	1	1	1	3	2	1	2
12	4	4	4	2	2	3	4	3	5
13	2	2	2	2	2	2	2	2	2
14	1	1	1	1	2	2	1	1	1
15	1	1	1	1	1	1	1	1	1
16	2	1	1	2	1	2	2	2	1
17	2	1	1	2	3	2	1	2	2
18	1	1	1	2	1	1	3	3	3
19	3	3	2	1	4	3	3	2	2
20									
21	1	2	2	2	2	3	1	3	2
22	2	1	1	1	2	2	1	2	2
23	3	3	3	2	2	4	2	3	1
24	2	2	2	2	2	2	3	2	2

No.	Ampliação do Acervo	Fixação de Recursos Humanos	Formação de Recursos Humanos	Informatização e Limpeza de Dados	Interação com outras coleções	Investimento na capacitação em gestão da coleção	Melhoria da Infraestrutura	Planejamento	Reconhecimento Institucional
25	2	2	2	2	2	3	2	2	2
26	1	2	2	1	1	2	1	3	2
27	1	2	2	2	3	1	1	2	2
28	2	2	2	1	2	2	2	2	3
29	3	2	2	2	2	3	4	3	2
30	3	2	2	2	2	3	3	2	3
31	4	4	3	3	3	4	4	4	4
32	3	1	2	1	4	3	1	2	2
33	1	1	1	2	3	3	1	3	2
34	2	1	1	1	2	2	3	3	3
35	1	1	2	2	2	2	2	2	2
36	2	1	1	1	2	1	2	2	2
37	1	2	2	2	2	3	1	3	3
38	3	4	4	2	3	4	5	2	4
39	1	3	2	1	2	4	1	2	1
40	3	1	1	1	3	2	3	2	4
41	3	1	2	2	3	2	1	2	1
42	2	3	3	3	3	3	4	4	3
43	3	2	2	3	3	3	1	2	2
44	1	1	3	3	3	3	1	3	3
45	2	2	1	2	3	3	3	4	2
46	2	2	1	3	2	2	1	1	1
47	2	1	2	3	3	3	3	2	1
48	2	1	1	2	3	1	1	2	2
49	2	1	1	1	2	1	1	2	2
50	2	1	1	1	3	1	1	2	2

No.	Ampliação do Acervo	Fixação de Recursos Humanos	Formação de Recursos Humanos	Informatização e Limpeza de Dados	Interação com outras coleções	Investimento na capacitação em gestão da coleção	Melhoria da Infraestrutura	Planejamento	Reconhecimento Institucional
51	3	1	3	2	3	1	3	2	5
52	3	3	2	1	3	2	2	3	1
53	3	1	2	3	3	2	1	2	1
54	3	4	3	2	2	4	3	2	2
55	1	1	1	3	3	2	1	1	1
56	1	3	1	2	3	2	2	2	2
57	2	1	2	3	3	1	2	2	2
58	3	1	2	3	3	1	3	1	2
59	4	2	2	3	4	4	2	4	1
60	2	1	1	2	2	1	1	3	5
61	2	1	3	1	3	3	4	2	2
62	3	4	3	3	4	4	4	3	4
63	4	1	1	1	1	1	2	2	1
64	2	1	2	2	3	3	2	3	3
65	1	2	3	3	2	2	1	1	1
66	3	2	2	2	2	3	4	3	3
67	3	2	2	2	2	2	2	2	4
68	3	1	1	2	3	2	3	3	3
69	4	2	2	4	3	2	3	2	3
70	2	2	1	2	3	3	3	3	3
71	1	1	1	1	2	1	1	1	1
72	1	1	1	1	2	1	1	1	1
73	2	1	1	1	2	1	1	2	1
74	3	1	1	3	3	2	1	2	2
75	3	2	3	2	3	3	2	2	1
76	3	2	3	1	2	3	2	2	2

No.	Ampliação do Acervo	Fixação de Recursos Humanos	Formação de Recursos Humanos	Informatização e Limpeza de Dados	Interação com outras coleções	Investimento na capacitação em gestão da coleção	Melhoria da Infraestrutura	Planejamento	Reconhecimento Institucional
77	2	2	3	2	3	3	2	2	3
78	1	2	2	1	2	1	1	1	1
79	1	1	1	1	2	2	1	2	1
80	3	1	1	1	3	1	1	1	1
81	1	1	1	1	3	3	1	1	1
82	4	1	3	1	3	2	1	1	1
83	4	1	1	1	1	1	1	1	1
84	1	1	1	1	1	1	1	1	1
85	1	1	1	1	2	3	1	3	5
86	2	2		2		3	2	2	4
87	5	1	1	1	3	1	1	1	1
88	2	1	2	2	2	2	1	2	2
89	4	1	1	1	3	2	1	2	2
90	3	3	3	3	2	2	1	1	1
91	3	1	1	1	3	1	1	2	1
92	3	2	2	2	2	2	1	2	2
93	4	1	1	2	3	2	1	2	2
94	5	1	3	2	2	2	3	3	4
95	3	1	1	1	4	4	2	2	2
96	3	1	1	3	2	2	1	2	2
97	2	1	1	2	2	1	2	2	1
98	2	1	1	1	1	2	2	2	3
99	2	3	3	3	3	3	2	2	3