



Instituto de Pesquisas Jardim Botânico do Rio de Janeiro
Escola Nacional de Botânica Tropical

O Processo de Informatização de herbários: Estudo de caso

Flávio Luna Peixoto

2005



Instituto de Pesquisas Jardim Botânico do Rio de Janeiro
Escola Nacional de Botânica Tropical

O Processo de Informatização de herbários: Estudo de caso

Flávio Luna Peixoto

Orientador: Haroldo Cavalcante de Lima

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Botânica, Escola Nacional de Botânica Tropical do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, como parte dos requisitos necessários a obtenção do título de Mestre em Botânica.

Rio de Janeiro, fevereiro de 2005

O Processo de Informatização de herbários:
Estudo de caso

Flávio Luna Peixoto

Dissertação submetida ao corpo docente da Escola Nacional de Botânica Tropical, Instituto de Pesquisas Jardim Botânico do Rio de Janeiro - JBRJ, como parte dos requisitos necessários a obtenção do grau de Mestre.

Aprovada por:

Prof. _____ - Orientador
Dr. Haroldo Cavalcante de Lima

Prof. _____
Dra. Marli Pires Morim

Prof. _____
Dr. Mike Hopkins

Rio de Janeiro, 22 de fevereiro de 2005.

Ficha Catalográfica

Peixoto, Flávio Luna
P379p O processo de informatização de herbários : estudo de caso /
Flávio Luna Peixoto. – Rio de Janeiro, 2005.
ix, 69 f. : il. ; 28 cm.

Dissertação (mestrado) – Instituto de Pesquisas Jardim Botânico
do Rio de Janeiro/Escola Nacional de Botânica Tropical, 2005.
Orientador: Haroldo Cavalcante de Lima.
Banca examinadora: Marli Pires Morim; Mike Hopkins.
Bibliografia.

1. Herbários 2. Coleções científicas 3. Informatização. I. Título.
II. Escola Nacional de Botânica Tropical.

CDD 580.74

Agradecimentos

A Jorge Fontella Pereira, que me apresentou formalmente um herbário e me inseriu no contexto da informatização, no Herbarium Bradeanum.

A Haroldo Cavalcante de Lima, pela amizade e orientação nessa área de trabalho, nova para mim.

A Denis Filer pela instrução, atenção e pela atitude sempre positiva.

À Diretoria da Escola Nacional de Botânica Tropical, à Coordenação do Programa de Pós-Graduação e ao secretariado da ENBT pela oportunidade de realizar minha pesquisa e pelo apoio no decorrer do curso.

Aos membros da banca examinadora pelo tempo e atenção dedicados à minha pesquisa.

À FAPERJ pelo apoio através da concessão de bolsa.

À Márcia e Abílio por estarem sempre disponíveis, também via e-mail, telefone ou fax.

Ao corpo de pesquisadores e funcionários do Jardim Botânico por ter me acolhido tão calorosamente.

À Regina, Alda, Jose, César e todos do IAN pela atenção, apoio, e por tornar minhas visitas à Belém tão agradáveis.

À Elsie e toda nova diretoria do Herbarium Bradeanum, por ter permitido a continuidade de minha pesquisa.

A todos que se envolveram voluntariamente na informatização do Herbarium Bradeanum.

Aos curadores de herbário pelo tempo consumido com meu questionário.

A Eduardo Dalcin pelas referências bibliográficas.

À Marli pelos inúmeros toques durante minha pesquisa e pela revisão da dissertação.

A Marcelinho e Juan que tantas vezes me mostraram o “caminho das pedras” da Botânica.

A minha mãe pelo carinho e pela orientação informal.

A meu pai pelas palavras de apoio nos momentos de dificuldade.

A Claudia pelo amor e apoio irrestrito.

Lista de Siglas e Abreviaturas	viii
Lista de Tabelas e Gráficos	ix
Resumo	1
Abstract	3
Introdução	5
Material e Métodos	8
Revisão Bibliográfica	11
Resultados	17
A informatização do herbário IAN	17
A informatização do herbário HB	22
A informatização do herbário RB	27
Análise dos questionários	32
Discussão	37
Conceito de banco de dados	37
Conceito de informatização	38
Etapas da informatização	40
Estimativa de tempo de execução e custo da entrada e revisão de dados	43
Situação atual e perspectivas da informatização dos herbários brasileiros	45
Próximos passos	47
As condições ideais para herbários brasileiros realizarem a informatização	48
Conclusões	50
Referências Bibliográficas	53
Bibliografia Complementar	57
Lista de sites utilizados	58
Anexo 1: modelo de questionário	59

Anexo 2: tabulação das respostas ao questionário	62
Apêndice: artigo enviado para a revista Ciência Hoje das Crianças	67

Lista de Siglas e abreviaturas utilizadas

- BGCI – Botanic Gardens Conservation International
- BRAHMS – Botanic Research and Herbarium Management System
- CDB – Convention on Biological Diversity
- CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico
- CRIA – Centro Referência em Informação Ambiental
- DFID – Department for International Development
- EMBRAPA – Empresa Brasileira de Pesquisa Agropecuária
- GBIF – Global Biodiversity Information Facility
- GTI – Global Taxonomy Initiative
- HB – Herbarium Bradeanum, Rio de Janeiro
- HISPID – Herbarium Information Standards and Protocols for Interchange of Data
- HUEFS – Herbário da Universidade Estadual de Feira de Santana
- IAN – Herbário do Centro de Pesquisa Agroflorestal da Amazônia Oriental da EMBRAPA, Belém
- IBGE – Herbário do Instituto Brasileiro de Geografia e Estatística, Brasília
- INPA – Instituto Nacional de Pesquisa da Amazônia, Manaus
- INPI – International Plant Name Index
- JBRJ – Instituto de Pesquisas Jardim Botânico do Rio de Janeiro
- LNCC – Laboratório Nacional de Computação Científica
- MG – Herbário do Museu Paraense Emílio Goeldi, Belém
- ORSTOM - Institut de Recherche pour le Développement, atualmente IRD
- PUC-RJ – Pontifícia Universidade Católica do Rio de Janeiro
- R – Herbário do Museu Nacional – UFRJ, Rio de Janeiro
- RB – Herbário do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro

SERPRO – Serviço Federal de Processamento de Dados

SGC – Sistema para Gerenciamento de Coleções

SUDAM – Superintendência de Desenvolvimento da Amazônia

UERJ – Universidade do Estado do Rio de Janeiro

UFP – Herbário da Universidade Federal de Pernambuco, Recife

Lista de Tabelas e Gráficos

Tabela 1: Custo da entrada de dados das etiquetas do herbário da EMBRAPA – Amazônia Oriental (IAN) em banco de dados no período de 1995 a 2002.	20
Tabela 2: Lista de campos usados no herbário da EMBRAPA – Amazônia Oriental (IAN) para entrada de dados e informação a ser incluída	20
Tabela 3: Custo da entrada de dados das etiquetas do Herbarium Bradeanum em banco de dados no período de 2001 a 2003.	26
Tabela 4: Lista de campos usados no Herbarium Bradeanum (HB) para entrada de dados e informação a ser incluída	26
Tabela 5: Estimativa do custo da entrada de dados das etiquetas do herbário do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro (RB) em banco de dados.	30
Tabela 6: Lista de campos usados no herbário do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro (RB) para entrada de dados e informação a ser incluída	30
Figura 1: Ritmo de entrada de dados no Herbarium Bradeanum (HB) de outubro de 2001 a junho de 2003	25
Figura 2: Porcentagem de herbários em Informatização no Brasil	33
Figura 3: Programas mais utilizados para a informatização de herbários no Brasil	33
Figura 4: Porcentagem de herbários brasileiros que contam com pessoal especializado em informática para apoio à informatização	34
Figura 5: Principais dificuldades apontadas pelos curadores de herbários brasileiros para realizar a informatização	35

As coleções botânicas têm papel fundamental no inventário da diversidade biológica. Além de material essencial e indispensável para estudos taxonômicos elas guardam dados que são utilizados em muitos outros ramos da ciência e úteis em políticas de conservação. A implantação de bancos de dados para herbários, comumente chamada informatização, é necessária para que essas coleções possam atender à crescente demanda pelas informações nelas contidas. Uma das importantes demandas encontra-se exposta nas Estratégias para Conservação de Plantas, da Convenção da Diversidade Biológica, que prevê o levantamento de dados sobre as espécies vegetais e seus diferentes habitats, especialmente aqueles mais ameaçados, em todo o mundo. Esse trabalho buscou conhecer melhor o planejamento e o desenvolvimento da informatização através da análise desse processo em três herbários brasileiros: O herbário do Centro de Pesquisa Agroflorestal da Amazônia Oriental da EMBRAPA (IAN), o Herbarium Bradeanum (HB) e o herbário do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro (RB), cada um deles em uma distinta fase de informatização. A análise foi feita através de visitas a estes herbários, participação em discussões sobre o processo de informatização nestas instituições e a própria informatização de parte do acervo de uma delas. Para melhor entendimento, procurou-se saber como a informatização estava se processando em outros herbários brasileiros utilizando-se questionário enviado a seus curadores e entrevistas não formais com curadores e pesquisadores durante os dois últimos congressos nacionais de botânica, realizados em Belém e Viçosa. Utilizou-se de apoio bibliográfico, especialmente de artigos oriundos do exterior, já que o tema ainda é pouco explorado no Brasil. Desta forma foi possível identificar as diferentes etapas do processo e as maiores dificuldades encontradas

nas três instituições estudadas, além de uma forma de estimar o custo e o tempo necessário para informatizar um herbário. Sugere-se que todo processo de informatização seja iniciado com a elaboração de um projeto detalhado e que nele sejam previstas diferentes etapas, entre as quais se destacam: 1) escolha do software a ser utilizado, 2) tratamento das informações, 3) treinamento de pessoal, 4) digitação dos dados, 5) correção de erros e 6) disponibilização dos dados. Também sugere-se a definição de um responsável pelo trabalho que venha a se tornar o administrador do banco de dados, quando a informatização estiver concluída, preferentemente pessoal da instituição, ficando assegurada a continuidade do trabalho de informatização. Os resultados apresentados podem ser úteis para o planejamento e execução de novos projetos de informatização e para orientar políticas regionais ou nacionais para a recuperação de informações sobre diversidade vegetal.

Botanic collections play an important role on the inventory of biological diversity. They are an essential part of taxonomic studies and contain data that can be used in many other scientific fields. The use of data-banking in herbaria, usually known as computerization, is necessary so that these collections can meet the growing demand for access to the information they hold. One of these demands is expressed in the Global Strategy for Plant Conservation, of the Convention on Biological Diversity, which aims to gather data about plants and their habitats, especially the most threatened ones. This work researched the planning and execution of the computerization through analysis of the process in three Brazilian herbaria: the herbarium of Centro de Pesquisa Agroflorestal da Amazônia Oriental da EMBRAPA (IAN), the Herbarium Bradeanum (HB) and the herbarium of Instituto de Pesquisas Jardim Botânico do Rio de Janeiro (RB), each in a different stage of computerization. Data was collected by visiting these herbaria, by attending meetings about computerization and by computerizing part of one collection. To understand the way other Brazilian herbaria are using data bases, a questionnaire was sent to curators. Interviews were used during the 53rd and 54th Congresso Nacional de Botânica, held in Belém and Viçosa. Bibliografic research was used to support the results, especially articles from Europe and North America, since this is a relatively new subject in Brazil. Thus, it was possible to identify the different stages of the computerization process and the most common problems that emerged in the three institutions analyzed. A way to estimate cost and time of computerizing a collection is presented. It is suggested that such a process should begin with a detailed project, comprising various different stages: 1) choosing the software; 2) treatment of the data; 3) training; 4) data entering; 5) data correction; and 6) making data available after computerizing. As a means to assure the continuity of the

process, the designation of a person responsible for the process, who should become the data-base administrator when the computerization is complete, is also suggested. The results presented here should be useful for planning and executing new computerization projects and for establishing regional or national policies on biodiversity data retrieval.

Um dos maiores problemas para o homem na atualidade é o fato de que a destruição do meio ambiente acontece mais rápido que o inventário sobre as espécies de animais e plantas e os ecossistemas onde vivem (Prance, 2001). Ecossistemas inteiros podem desaparecer sem que se tenha conhecimento dos seres ali existentes e das relações que ocorrem entre eles. Além do óbvio empobrecimento da natureza pela extinção de espécies, a perda da biodiversidade traz um prejuízo incalculável para a humanidade, pois ela guarda muitas soluções naturais para problemas econômicos e de saúde, a maioria ainda desconhecida (BGCS, IUCN & WWF, 1989).

Como forma de diminuir as perdas de espécies vegetais, foi criada a Estratégia Global para Conservação de Plantas (Global Strategy for Plant Conservation). Para atingir seu objetivo, esse plano pretende levantar dados sobre a diversidade vegetal – estimada atualmente em cerca de 300.000 espécies – e os diferentes habitat, especialmente aqueles mais ameaçados e criar uma rede de trabalho que interligue as iniciativas conservacionistas em todo o mundo. Para isso, prevê que sejam elaboradas até 2010 listas de espécies conhecidas, incluindo seu uso e sua distribuição. A estratégia reconhece que o uso de sistemas de informação, como bancos de dados, é uma ferramenta essencial para o levantamento e a conseqüente conservação da flora mundial (CBD & BGCI, 2004).

Novas tecnologias têm sido aplicadas para aumentar a capacidade de inventariar a biodiversidade e lidar com as informações provenientes desse inventário, bem como para tentar usar a natureza de maneira sustentável (Zorzetto, 2002), mas, por motivos muito variados, nem sempre o resultado pode ser considerado satisfatório.

As coleções científicas têm papel fundamental para a realização do inventário da biodiversidade. As coleções botânicas, além de material essencial e indispensável para

estudos taxonômicos, guardam os passos da evolução da ciência botânica, um histórico dos pesquisadores que a ela se dedicaram, um retrato de potencialidades etnobotânicas (veja-se as “simples” anotações em etiquetas de herbário) e um quadro de modificações ambientais geradas por ações humanas ao longo de amplos intervalos de tempo.

A informatização dos herbários, agilizando a busca de dados para as pesquisas em botânica e ciências afins e tornando as informações aí depositadas disponíveis para um público mais diversificado, é fundamental para o avanço da ciência, ao acelerar a troca de informações e facilitar o gerenciamento de coleções (Brito & Joly, 1999). Quando bem aplicadas, tecnologias de informação podem contribuir para melhorar a organização e a recuperação dos dados depositados em coleções biológicas (Wilson, 2001 e Bisby, 2000), visando o planejamento de políticas de manutenção e ampliação dos seus acervos bem como a conservação dos espécimes, a unidade funcional das coleções (Stuessy, 1996).

A maior parte dos herbários brasileiros iniciou recentemente o processo de informatização, em geral de maneira bastante independente (Peixoto & Barbosa, 2002). Apesar disso, tais processos podem apresentar – em todas as etapas ou em algumas delas – problemas semelhantes. Dúvidas sobre o software a ser utilizado, tratamento das informações, que tipo de informação divulgar e como divulgá-la, quais recursos materiais e humanos devem ser aplicados se repetem, tornando a informatização mais lenta e, em alguns casos, interrompendo ou inviabilizando o processo. Atualmente no país, mesmo em herbários com os dados dos espécimes totalmente digitalizados, as informações resultantes do processo nem sempre estão disponíveis para a comunidade científica ou para a sociedade por motivos muito variados.

Partindo da constatação de que os herbários brasileiros são depositários de informações fundamentais para o inventário da biodiversidade e de que conhecer a biodiversidade é fundamental para a conservação, a pesquisa aqui apresentada pretende contribuir para que projetos de informatização de coleções científicas em botânica e disponibilização das informações para a sociedade possam ser planejados e executados com mais precisão, otimizando assim esforços e insumos.

O trabalho foi realizado como estudo de caso em três herbários brasileiros que encontram-se em etapas diferentes do processo de informatização: O Herbário da EMBRAPA/CPATU (Belém, PA, informatizado), O Herbarium Bradeanum (Rio de Janeiro, RJ, em informatização) e o Herbário do Jardim Botânico do Rio de Janeiro (Rio de Janeiro, RJ, concluindo o projeto de informatização).

A pesquisa teve como objetivo a busca de ferramentas auxiliares para a execução de trabalhos de informatização de coleções por instituições botânicas, através de uma análise comparativa em herbários que já realizaram o processo ou o estão implantando, buscando responder a duas questões principais: 1. quais são as dúvidas/problemas mais comuns no processo de informatização de uma coleção botânica? 2. quais soluções foram encontradas pelos herbários brasileiros que estão mais adiantados no processo?

Para responder tais questões, a pesquisa procurou apontar: 1) as diferentes etapas identificáveis no processo; 2) os problemas recorrentes na informatização de um herbário; 3) as soluções encontradas por curadores e suas equipes para dificuldades comuns; 4) os meios empregados por cada herbário para realizar o trabalho; 5) uma maneira de estimar o tempo e o custo de realização das tarefas.

Os herbários brasileiros são muito distintos entre si no tamanho e abrangência de suas coleções e na disponibilidade de recursos humanos e financeiros. Para atingir os objetivos da pesquisa, foi analisada a informatização em três herbários brasileiros: IAN, HB e RB. Cada um deles encontra-se em etapa distinta do processo de informatização. Os dados apresentados a seguir foram retirados do site Index Herbariorum Online (uma lista de sites utilizados está disponível na página 57).

O Herbário do Centro de Pesquisa Agroflorestal da Amazônia Oriental da EMBRAPA (IAN), em Belém, Pará, foi fundado em 1945. Conta com uma coleção de cerca de 180.000 exsicatas, a maioria de espécies arbóreas da Amazônia.

O Herbarium Bradeanum (HB), no Rio de Janeiro, RJ, foi fundado em 1958. Guarda aproximadamente 90.000 exsicatas, com destaque para as coleções de Bromeliaceae, Orchidaceae e Pteridófitas e para os cerca de 1300 tipos nomenclaturais. Seu acervo conta com coletas de Guido Pabst, Alexander Kurt Brade, Edmundo Pereira, Alberto Castellanos, Dimitri Sucre, Roberto Kautsky, R. Reitz e Gert Hatschbach, entre outros.

O Herbário do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro (RB) foi fundado em 1890. Sua coleção de cerca de 400.000 exsicatas contém coletas de vários pesquisadores importantes para a Botânica no Brasil, como Barbosa Rodrigues, A.C. Brade, A.P. Duarte, A.C. Ducke, A.F.M. Glaziou, F.C.P. St. Hilaire, F.C.Hoehne, J.G. Kuhlmann, G.F.J. Pabst, E. Pereira, R. Reitz, C.T. Rizzini, C.A.W. Schwacke, R. Spruce, E. Ule, D. Sucre, P. Campos Porto e P.K.H. Dusen, G. Martinelli, entre outros. Sua coleção de tipos nomenclaturais tem cerca de 5000 exemplares.

Também foram analisados os dados relativos às coleções correlatas guardadas nestes herbários (carpotecas, xilotecas e outras). A informatização destas coleções, quando ocorreu, também foi considerada. Portanto, o termo herbário é aplicado aqui em *lato sensu*.

Os processos nestas instituições foram descritos e os dados obtidos estão relatados em Resultados. A metodologia aplicada foi adaptada de Sarasan (1981), quando foi aplicada para a análise do uso de recursos de informática em coleções científicas de diversas instituições dos Estados Unidos.

No herbário da Embrapa / Amazônia Oriental, em Belém, foram realizadas duas visitas técnicas, em agosto de 2002 e julho de 2003. Nestas visitas, além de assistir a um curso sobre a utilização do BRAHMS (Botanical Research and Herbarium Management System), programa desenvolvido pela Universidade de Oxford para a informatização de coleções botânicas, foi possível obter informações sobre uma tentativa anterior de informatização, que não foi concluída. Também foram recebidas informações e opiniões valiosas sobre o planejamento de um projeto de informatização.

No Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, durante os anos de 2003 e 2004, foi possível acompanhar algumas discussões e os esforços na direção da informatização de acervos. O acesso à documentação também foi muito facilitado. Houve a participação, como convidado em uma reunião da Comissão de Curadores do Jardim Botânico.

No Herbarium Bradeanum, foi possível adquirir experiência prática de gerenciamento do processo de informatização de parte de uma coleção, num período de dois anos. Essa experiência foi fundamental para formar lastro para o entendimento dos avanços e problemas relatados pelo curador e técnicos responsáveis pela informatização do herbário da Embrapa, em Belém e também possibilitou acompanhar melhor as discussões para a informatização da coleção do herbário do JBRJ. Outro facilitador foi o domínio do software BRAHMS.

Depois de coletados, os dados sobre a informatização nos três herbários supra citados foram analisados comparativamente. Para essa parte do estudo foi importante a experiência do autor da pesquisa no trabalho de informatização de coleções científicas. O principal objetivo desta análise foi dissecar o processo, identificando e detalhando as diferentes etapas que devem ser cumpridas para o bom andamento da informatização. Desta forma buscou-se ainda criar uma fórmula para a elaboração de cronogramas para a informatização de coleções que se adapte às diferentes realidades dos herbários brasileiros.

Para auxiliar essa análise foram aplicados conceitos obtidos através da revisão bibliográfica. A bibliografia foi particularmente importante para esclarecer termos empregados usualmente, como campos mínimos de informação, dados confiáveis e funcionamento de um banco de dados para curadoria e pesquisa em botânica. Além das referências bibliográficas citadas no texto, em seu final são apresentadas uma lista complementar de bibliografia e uma lista de sites na Internet relevantes para a pesquisa, que não estão citados ao longo do trabalho. Eles contêm informações úteis para pesquisas taxonômicas e em coleções científicas.

A metodologia utilizada incluiu também a aplicação de um questionário, distribuído a curadores nos meses de junho a agosto de 2004, para revelar suas necessidades e os problemas enfrentados para a execução de suas tarefas (O modelo de questionário está no anexo 1 e a tabulação das respostas está no anexo 2). Para melhorar a representatividade das respostas ao questionário, ele foi aplicado pessoalmente durante o 55º Congresso Nacional de Botânica, realizado em Viçosa, Minas Gerais, de 18 a 23 de julho de 2004. O questionário visou também conhecer melhor o estado atual da informatização dos herbários brasileiros e prever quando estes dados podem estar disponíveis digitalmente.

Com a mesma finalidade do questionário, foram realizadas entrevistas abertas (semi-estruturadas) com pesquisadores e curadores. O contato obtido em cursos ministrados pelo autor (no Herbarium Bradeanum, em novembro de 2002 e durante o 55º Congresso Nacional de Botânica, em julho de 2004) também foi útil ao revelar as pretensões dos usuários de bancos de dados e as dificuldades para sua implementação em diferentes instituições.

Convém destacar que não fez parte do presente estudo a análise em profundidade sobre software e ferramentas de programação para criação de banco de dados. O foco foi voltado para o usuário dos programas (pesquisadores, curadores e instituições) e para a recuperação de dados e circulação de informação científica a partir do uso da informática. Ou seja, como o uso acertado dos recursos de programas de computador pode auxiliar no manejo de coleções e na transformação de dados em informação.

As siglas citadas ao longo do texto estão apresentadas, junto com seu significado, na página viii.

Os primeiros textos a tratar de uso de informática em coleções científicas datam da década de 1960. Curadores e pesquisadores dos Estados Unidos e Europa manifestavam sua preocupação em lidar com a quantidade crescente de material armazenado e as informações relacionadas, apontavam caminhos para o uso de ferramentas eletrônicas e descreviam suas tentativas de aplicação de novas tecnologias.

Squirres (1966) sugeriu quatro principais problemas para a falta de esforço na direção da informatização: antipatia de curadores quanto à adoção de técnicas modernas; ignorância de especialistas, especialmente em sistemática, das várias tecnologias disponíveis para seu trabalho; falta de recursos para a modernização; e medo de desperdiçar tempo de pesquisa em tarefas como a criação de um sistema ou das informações nele depositadas.

Como solução para os dois primeiros problemas ele sugeriu, baseado na sua experiência no Museu de História Natural do Smithsonian Institution, a criação de um comitê para pesquisar as possibilidades de aplicação da informática e a contratação de pessoal especializado na área. Os dois problemas finais parecem permanentes, segundo o autor. Ele considerou que o processo de digitalização de dados seria dispendioso e que a tarefa de verificação dos dados de fato demandaria tempo que poderia ser dedicado à pesquisa.

Soper & Perring (1967) discutem a necessidade de criação de bases eletrônicas e o tempo que os curadores gastariam com sua implementação. Descrevem os equipamentos que estavam disponíveis.

Crovello (1967) analisa as possibilidades de uso de processamento de dados em coleções biológicas e apresenta as dificuldades que podem ser encontradas.

Hall (1972) explica os tipos de banco de dados que podem ser usados para taxonomia e mostra como deve ser construída sua estrutura.

Brenan (1974) resume os trabalhos apresentados e as discussões levantadas durante uma conferência sobre o uso de processamento de dados em coleções taxonômicas da Europa. Até então a maioria dos herbários europeus não havia implantado nenhum sistema de informação eletrônico.

Morse (1974) identifica tipos diferentes de dados contidos em coleções biológicas (sobre os espécimes, ecológicas e relacionadas à taxonomia). Explica as vantagens do uso de bancos de dados e avalia as opções então disponíveis.

Hall (1974) descreve as possibilidades de uso de um banco de dados e as etapas necessárias para sua implementação.

Sarasan (1981) descreve experiências de diversas instituições dos Estados Unidos no uso de recursos de informática. Relata também resultados de uma pesquisa realizada pela Association of Systematics Collections (ASC) com coleções de todo o país. E enumera alguns motivos que levam ao fracasso de tentativas de informatizar coleções, entre eles: falta de recursos financeiros; falta de agilidade na tomada de decisões; falta de um projeto para a informatização; baixos salários; falta de continuidade administrativa; ausência de padrões nas informações a serem digitalizadas. A autora esclarece ainda termos que são usados em projetos e processos de informatização como:

- Registro: uma série de campos relacionados, todos com dados de um único espécime;
- Elemento de dados: a menor unidade de informação a que se faz referência. Ano de coleta é um exemplo de elemento de dados.
- Campo de dados: é a área de um registro onde um tipo de dado é inserido. Podem conter mais de um elemento de dados ou não (data de coleta é um exemplo de campo de dados); e
- Controle de sintaxe: uma forma de controlar a entrada de dados, de modo a uniformizá-los. Por exemplo, anos só podem ser incluídos com 4 dígitos.

Vogellehner & Speck (1988) apresentam o programa DIDEA-FR, para uso em jardins botânicos.

Russel & Arnold (1989) avaliam suas experiências em 15 anos de uso de computadores em herbários, através do sistema PRECIS para informações taxonômicas.

Macrander & Haynes (1990) apresentam os primeiros resultados do sistema SERFIS (Southeastern Regional Flora Information System) para a criação de bancos de dados de coleções de vários herbários nos Estados Unidos.

Morin & Gomon (1993) descrevem vários processos para a recuperação e disponibilização de dados em diferentes países.

Conn (1996, ed.) apresenta uma lista de campos que devem ser respeitados na criação de banco de dados de informações sobre biodiversidade. Esses campos foram criados para tornar possível a troca de informações entre os bancos de dados de todas as coleções.

Davenport & Prusak (1998 apud Dias & Belluzzo, 2003) conceituam dado, informação e conhecimento.

Em 2001, a revista *Taxon* publicou um número dedicado exclusivamente ao processamento de dados para o levantamento e conservação da biodiversidade. Alguns artigos se destacam:

Wilson (2001), que descreve as vantagens de novos meios de transmissão de informação como a Internet e o que deve ser feito para otimizar seu uso. Dá exemplos de sites que apresentam floras digitais de alguns estados norte-americanos como resultados de uma nova forma de organização e colaboração no trabalho de produção e distribuição de informação.

Heywood (2001), que aponta a necessidade da aplicação de novas tecnologias e conceitos na taxonomia e dá um novo papel aos herbários, às coletas neles depositadas e à maneira como os dados são guardados e recuperados. Sobre banco de dados para herbários, afirma que sempre que possível devem ser usados programas e plataformas já existentes para evitar o risco de erros e perda de dados e, principalmente, orientar recursos para onde sejam realmente necessários.

Prance (2001), que apresenta um quadro geral das coletas de plantas e sua conservação. Afirma que o uso de novas tecnologias, especialmente banco de dados e de imagens pode acelerar o trabalho do taxonomista e ajudar o levantamento da biodiversidade ao mostrar mais clara e rapidamente os pontos de maior diversidade (*hotspots*) e os que são, por outro lado, pouco pesquisados.

Smith *et al.* (2003), descrevem o processo de informatização realizado em herbários da África e uma maneira de estimar os custos que pode ser aplicada a outras instituições.

Uma das dificuldades encontradas por quem planeja informatizar uma coleção científica no Brasil é exatamente a pequena quantidade de material de referência em Português sobre o processo de informatização e sobre as experiências já realizadas.

Durante a década de 1970, o CNPq coordenou o Programa Flora, que tinha entre seus objetivos inventariar a flora brasileira e criar métodos eficientes de consulta a esse inventário. O desenvolvimento do trabalho dos cientistas envolvidos está resumido em alguns relatórios e publicação:

Teixeira (1979) mostra primeiros resultados do Programa Flora e explica a construção de um banco de dados de espécimes de herbários e de bancos satélite com informações sobre espécies vegetais.

Teixeira (1980) atualiza os resultados do Programa nos núcleos regionais e descreve o tratamento das informações feito pelo SERPRO (Serviço Federal de Processamento de Dados) para a criação do Banco de Dados do Programa Flora, sob cuidado da EMBRAPA.

Teixeira (1984) descreve a origem do Programa Flora, seus objetivos e sua estruturação e métodos de trabalho, especialmente em alguns núcleos regionais.

Outros trabalhos levantam a situação da pesquisa em Botânica no Brasil e apontam caminhos a serem seguidos, sempre incluindo o uso de ferramentas de processamento de dados.

Nogueira (1987) retrata a realidade da Botânica brasileira na primeira metade dos anos 80 e apresenta estratégias para formar recursos humanos e inventariar mais rapidamente a diversidade vegetal brasileira.

Peixoto (1999) avalia a pesquisa Botânica brasileira e seus desafios, limites e objetivos para o século XXI através do estado das coleções científicas.

Muitos herbários apresentaram o desenvolvimento de seu banco de dados em Congressos de Botânica. Os livros de Resumos dos Congressos Nacionais de Botânica da Sociedade Brasileira de Botânica publicaram vários deles:

Martins-da-Silva e Ferreira (1996) descrevem a informatização da família Lecythidaceae no início do banco de dados do herbário IAN.

Ferreira & Martins-da-Silva (1997) descrevem a informatização da família Vochysiaceae.

Martins-da-Silva *et al.* (1998) apresentam a formação dos bancos de dados dos herbários IAN e MG.

Sonehara *et al.* (2000) descrevem o uso e as funções do banco de dados do herbário da Universidade do Tuiuti, no Paraná.

Daly & Silveira (2000) descrevem a criação do banco de dados da Flora do Acre.

Nunes *et al.* (2002) divulgam o sistema de gerenciamento de dados da coleção do herbário HUEFS, o programa Herbário 2.0, desenvolvido em Access.

Borges *et al.* (2002) descrevem e avaliam o banco de dados do herbário UFP, usando o programa BRAHMS.

Oliveira & Senna-Valle (2002) apresentam os passos para a informatização da família Convolvulaceae do herbário R.

Alguns textos se referem a sistemas específicos:

Campos dos Santos *et al.* (2000) apresentam um estudo sobre a base de dados biológicos do INPA e avaliam a possibilidade de criar uma ferramenta para análise ambiental mais ampla.

Magalhães *et al.* (2001) apresentam a proposta de um sistema de informações para organizar, integrar e gerenciar os dados sobre biodiversidade presentes em coleções biológicas, a partir da experiência do INPA. Fazem então uma análise mais ampla dos problemas, prioridades e possíveis parcerias envolvidas na criação e implantação de um sistema deste tipo, incluindo ações de curto, médio e longo prazo.

Em 2002, foi realizado na cidade de Indaiatuba, estado de São Paulo, um simpósio para a apresentação e discussão de inovações tecnológicas para pesquisa científica. Os resumos das palestras apresentadas estão disponíveis, em meio eletrônico (<http://www.cria.org.br/eventos/tdbi/>). Entre eles é importante destacar:

Peixoto & Barbosa (2002), que mostram a situação atual das coleções botânicas no Brasil. Indicam que apenas 29% dos herbários brasileiros têm mais da metade de seu acervo informatizado, ou concluíram a informatização;

Thiers (2002), que indica como se apresentou a necessidade de melhorar o acesso a informações do Jardim Botânico de Nova Iorque, como o processo foi dividido em etapas e como as informações se tornam disponíveis para pesquisadores em todo mundo;

Chapman (2002), que apresenta novas tecnologias para acelerar o processo de produção das floras;

RCS Arte Digital (2002), que se apresenta como empresa responsável pela criação do Herbário Virtual do Instituto de Botânica de São Paulo e mostra os objetivos do projeto, bem como os recursos nele utilizados até aquele momento.

Alguns outros textos, igualmente recentes, apresentados durante o 53º Congresso Nacional de Botânica e publicados posteriormente em formato de livro, devem ser considerados:

Canhos *et al.* (2002), que apresentam programas e sistemas para o tratamento e distribuição de informações de interesse científico, criados em diversas parcerias com o CRIA (Centro de Referência em Informações Ambientais);

Thomas & Magill (2002), que discutem a importância do uso de bancos de dados para a sistemática, ecologia e conservação, especialmente através da geração de checklists;

Monteiro Filho & Costa (2002), que apresentam as etapas do processo de informatização dos herbários IBGE e HRB. Na verdade, discutem as etapas para a criação de um banco de dados próprio que será utilizado na informatização, embora nem todas tenham sido cumpridas;

Gamarra-Rojas (2002), que mostra a elaboração do checklist de plantas do nordeste como um produto da informatização das coleções da região e relata as fases de execução desse processo, discutindo ainda as vantagens e desvantagens da metodologia aplicada.

Peixoto (2003a) organizou a publicação de um livro sobre coleções biológicas. Em vários capítulos, os diferentes autores apontam a importância da informatização e apresentam o estado da arte das coleções brasileiras. Além da informatização, o livro também destaca o valor da visita de taxonomistas a diferentes coleções.

Por fim, alguns manuais de bancos de dados para herbários também contêm conceitos relevantes para o entendimento do processo de informatização:

Breyer & Tsuboi (1996) que apresentam seu programa para informatização de herbários (Sistema Herbário) e discutem aspectos da curadoria de coleções e de sua informatização.

Filer & Bennet (2002), que citam conceitos como campos mínimos de informação, revisão de informações e apresentam uma série de produtos (publicações, websites, etiquetas, listas...) que podem ser produzidas com o software BRAHMS.

Peixoto (2002), que indica os passos para a informatização de uma coleção científica a partir de um software de banco de dados para herbários já existentes (BRAHMS).

O Processo de informatização dos herbários

A Informatização do herbário IAN

A primeira experiência com informatização da coleção do herbário IAN se deu dentro do Programa Flora do CNPq. Junto com o Museu Paraense Emílio Goeldi, a EMBRAPA Amazônia Oriental fez parte do núcleo de Belém do Projeto Flora Norte, iniciado em 1976. Seguindo a metodologia do Programa Flora, as informações das coleções eram levantadas e transcritas para fichas. Eram digitadas no SERPRO, RJ, trabalho que mais tarde foi transferido para o LNCC.

O volume de dados coletados foi grande. Segundo Nogueira (1987), até 1981 haviam sido preenchidos 146.868 formulários no Núcleo de Belém. Isso pode ser verificado com uma consulta à coleção do IAN. Grande parte das exsicatas está marcada com o carimbo do Programa Flora.

O Banco de Dados da Flora Brasileira ficou a cargo do LNCC, que verificava e acrescentava as informações enviadas pelos herbários. Desta forma os botânicos ligados às coleções tinham pouco contato com a tecnologia envolvida no processo. Apenas em 1984 o Museu Paraense Emílio Goeldi instalou um terminal conectado ao LNCC para ter acesso às informações.

Já a partir da primeira metade da década de 80 o apoio do CNPq ao Programa Flora diminuiu. A falta de recursos e de uma política clara, inclusive sobre os objetivos do Programa, levou ao término de diversos núcleos, especialmente os criados depois de 1981. O Programa foi Abandonado e os dados deixados sob a guarda do LNCC não se tornaram amplamente disponíveis.

No início da década de 90, o IAN passou por um novo processo de informatização usando o Software SISFITO. O Programa era desenvolvido em parceria da EMBRAPA e o ORSTOM – Institut Français de Recherche pour lê Développement en Cooperation. Gerard Cochneau e Patrick Séche eram os responsáveis pelo seu aperfeiçoamento, e contavam com o apoio dos Botânicos Georges Fotius e Irenice Alves Rodrigues.

Entretanto, o projeto não obteve sucesso dos pontos de vista da inclusão e disponibilização dos dados. A distância e a dificuldade de comunicação, especialmente a barreira dos idiomas, levaram ao abandono do trabalho, com menos de 500 etiquetas digitadas. O final da parceria trouxe um problema: o programa ainda não estava completo, e o herbário IAN não contava com nenhum programador capaz de terminá-lo. Outro problema do programa era a interface com o usuário, pouco amigável. No entanto, a experiência foi importante para mostrar aos pesquisadores da instituição as facilidades que poderiam advir com a informatização de suas coleções.

Em 1996, o laboratório de botânica da EMBRAPA – Amazônia Oriental ganhou novo prédio e iniciou um terceiro processo de informatização do herbário, utilizando o sistema BRAHMS, desenvolvido pela Universidade de Oxford. O Chefe da EMBRAPA, Dr. Dílson Frazão conheceu o programa numa visita à Inglaterra. De volta ao Brasil, fez uma proposta a Denis Filer, um dos responsáveis pelo desenvolvimento, de consultoria para a criação do banco de dados do IAN.

No início do trabalho, foram incluídos os dados da coleção de Tipos Nomenclaturais e das famílias Lecytidaceae, Malvaceae e Vochysiaceae. A consultoria inicial foi atrelada a um projeto de manejo florestal. A informatização dessas famílias, relevantes para o manejo, foi a forma encontrada para apresentar resultados mais rapidamente. Entre 1996 e 1997 foram digitados os dados de 4421 exsicatas.

Em 1998, o projeto de informatização passou a contar com apoio da SUDAM, através do projeto “Levantamento da diversidade vegetal da Amazônia, através da informatização dos herbário MG e IAN”. O apoio permitiu a contratação de dois digitadores, o que acelerou bastante a entrada de dados. Entre 1998 e 1999 foram digitados os dados de 65.066 exsicatas da coleção do IAN.

Os digitadores eram profissionais e ganhavam por produção, trabalhando 8 horas por dia, em dois computadores que ficavam na sala do herbário. Para a instalação do banco de dados principal foi adquirido um terceiro computador, que tem o acesso restrito para a segurança das informações. Atualmente o responsável pelo banco de dados continua fazendo cópias de segurança semanalmente em CD.

Para a realização do projeto, foi estabelecida uma parceria com o Museu Paraense Emílio Goeldi, localizado na mesma cidade de Belém, para informatizar a coleção do herbário daquela instituição. Essa iniciativa foi importante porque permitiu usar a experiência do pessoal do IAN no treinamento e na manutenção do banco de dados do Museu Paraense Emílio Goeldi, aumentando o volume de dados digitalizados para a formação de uma rede amazônica de coleções.

A partir de 2000, a informatização passou a contar com apoio formal do projeto DENDROGENE (convênio DFID / EMBRAPA - Amazônia Oriental). O objetivo do projeto é criar uma base de dados com informações sobre espécies vegetais da Amazônia. O IAN passou então a atuar como colaborador direto no desenvolvimento do software BRAHMS.

A participação no desenvolvimento do BRAHMS fez com que o Dr. Denis Filer visitasse o IAN ao menos duas vezes por ano. Nessas visitas, ele apresentava novas ferramentas do programa e organizava treinamentos. Problemas surgidos durante a informatização eram discutidos para que fossem criadas soluções.

Apesar disso, em 2000 houve uma interrupção no financiamento, o que levou a um intervalo no trabalho de entrada de dados.

Com a volta do apoio financeiro, de 2001 a 2002 foram digitados os dados de 112.132 exsicatas. A partir de 2002 iniciou-se um grande esforço de correção de dados, feita preferencialmente a partir de fotografias digitais das exsicatas e suas etiquetas.

O IAN terminou a digitação de dados da coleção do herbário, carpoteca e xiloteca em 2002. Também neste ano completou a correção e atualização dos dados da coleção de Tipos, com apoio de pesquisadores do Museu Nacional – UFRJ.

Os relatórios disponíveis sobre a informatização do herbário IAN tratam especialmente da prestação de contas às instituições financiadoras. Enquanto mostram claramente o custo do trabalho e a importância do financiamento, eles não apresentam nenhuma informação sobre o dia-a-dia da informatização e dos problemas encontrados.

Os custos estão resumidos na tabela abaixo (tabela 1). Os valores dos equipamentos são estimados de acordo com valores atuais. Não estão incluídos valores gastos com material de consumo. O Projeto de informatização com o apoio da SUDAM começou pagando R\$0,20 por etiqueta digitada (da carpoteca, xiloteca e do herbário) e depois passou a pagar R\$0,25. A correção dos dados digitados custa R\$0,15 por etiqueta revista e, como a correção ainda não acabou, o valor apresentado foi projetado a partir do número de etiquetas que foram digitadas.

Embora o projeto de informatização não tenha criado a função de curador do banco de dados, um funcionário do herbário tornou-se responsável por esse trabalho. Como ele não se dedica exclusivamente, apenas parte de seu salário foi incluída no cálculo de custo da informatização. Os encargos sociais não estão incluídos neste cálculo.

A tabela 2 apresenta os campos usados para a entrada de dados e a informação que deve ser colocada em cada campo.

Tabela 1: Custo da entrada de dados das etiquetas do herbário da EMBRAPA – Amazônia Oriental (IAN) em banco de dados no período de 1995 a 2002.

Digitadores			
Período	Etiquetas	Valor	Fonte Financiadora
1995-1997	4421		EMBRAPA
1998	25000	R\$ 5.000,00	SUDAM
1999	40066	R\$ 8.013,20	SUDAM
2001	45030	R\$ 9.006,00	SUDAM
2002	67102	R\$ 16.165,50	SUDAM
Responsável pelo banco de dados			
1995-99, 2001-02		R\$ 50.400,00	EMBRAPA
material			
Equipamento	Quantidade	Valor unitário	Valor total
PC para digitação	2	R\$ 1.500,00	R\$ 3.000,00
PC para banco	1	R\$ 3.000,00	R\$ 3.000,00
Máquina digital	3	R\$ 1.600,00	R\$ 4.800,00
Valor total:			
	R\$ 99.384,70	Etiquetas:	181619
Custo por etiqueta digitada:		R\$ 0,54	
Custo por etiqueta digitada e corrigida:		R\$ 0,69	

Tabela 2: Lista de campos usados no Herbário da EMBRAPA – Amazônia Oriental (IAN) para entrada de dados e informação a ser incluída

Campo	Informação
Registro	Número de tombo do material

Coletor	Nome do coletor
Número	Número da coleta
Coladic	Coletores adicionais
DiaCol	Dia da coleta
MesCol	Mês da coleta
AnoCol	Ano da coleta
Família	Família botânica
Gênero	Nome Genérico
Sp1	Epíteto Específico
Autor 1	Autor do nome da espécie
Infra 1	Categoria Infra-específica
Sp2	Epíteto Infra-específico
Autor 2	Autor do epíteto infra-específico
Detby	Determinador do exemplar
MesDet	Mês da determinação
AnoDet	Ano da determinação
País	País da coleta
Estado	Estado da Coleta
Município	Município da coleta
Gazetteer	Localidade da coleta
NotasLocalidade	Nota sobre a localidade
Obs	Observação de campo
Foto	Ligação para imagens
Uso	Uso
Lat	Latitude da coleta
N/S	Norte / Sul
Long	Longitude da coleta
E/W	Leste / oeste
Alt	Altitude da coleta
Dups	Em quais herbários há duplicatas
Flcode	Código para flor
Frcode	Código para fruto
Carp	Material na carpoteca

Xiloteca	Material na xiloteca
Vernacular	Nome vulgar
Idioma	Idioma do nome vulgar
Comment	Comentários
TreeNumber	Número da árvore
Typestat	Categoria de tipo

Informatização do Herbário HB

O Herbarium Bradeanum (HB) iniciou a informatização de sua coleção no segundo semestre de 2001. A partir de uma visita de Denis Filer, da Universidade de Oxford, responsável pelo desenvolvimento do programa BRAHMS, foi realizado um treinamento e iniciada a digitação das informações no banco de dados.

No HB, decidiu-se pelo software BRAHMS por vários motivos, entre os quais se destacam:

- baixo custo: fornecido sem custo algum, mediante *download* a partir do site www.brahms.co.uk e posterior registro no mesmo;
- compatibilidade: o programa importa dados já guardados em outros programas como Microsoft Excel e Access de modo simples e seguro;
- adequação: o BRAHMS é desenvolvido especificamente para o gerenciamento de herbários e informações botânicas, e é bastante completo;
- suporte: a equipe de desenvolvimento presta suporte permanente a instituições registradas através de e-mail, atualizações do manual e visitas (consultorias);
- experiências anteriores: o BRAHMS já está em uso em outros herbários brasileiros, o que permite troca de informações sobre seu uso dentro do Brasil;

O HB não contava então com nenhum apoio financeiro para o projeto. No início do trabalho, a entrada de dados ficou a cargo de apenas uma pessoa, que recebia uma ajuda de custo para o deslocamento até a instituição e alimentação. Desta forma, apenas um dia por semana era dedicado a essa tarefa.

Na verdade, a informatização da coleção do HB era vista por seu diretor, Dr. Jorge Fontella Pereira, como uma necessidade, não apenas para facilitar as tarefas de curadoria, mas como uma maneira de conseguir (abrir portas para) outros financiamentos. Acreditava

que a partir da publicação das primeiras informações via Internet ou através de catálogos impressos, poderiam ser obtidos parceiros para dar continuidade do processo.

A responsabilidade pela informatização do HB, levou o autor de presente dissertação a aprender o uso básico do BRAHMS no início do primeiro semestre de 2002 com Denis Filer, em duas visitas de um dia feitas por ele ao Museu Nacional- UFRJ e ao HB e com Gracialda Ferreira, do Herbário IAN, em visita de uma semana ao Herbarium Bradeanum. Mais tarde, em um curso realizado no Herbário IAN, organizado pela sua curadora, professora Regina Célia Martins, adquiriu conhecimentos para administrar melhor as informações do banco de dados.

As informações adquiridas foram repassadas ao pessoal do HB, constituído principalmente por estagiários (alunos de graduação) e pesquisadores voluntários. Foram feitos vários treinamentos individuais, à medida que novos estagiários se apresentavam para executar a entrada de dados.

Para iniciar a entrada de dados optou-se pela inclusão de todas as informações das etiquetas, além de (mais) informações bibliográficas especialmente da coleção tipo. Os dados sobre habitat, hábito e a descrição dos espécimens foram agrupadas em um único campo (numa única observação), para facilitar a digitação. Foram criados padrões para os nomes de coletores e cientistas autores de nomes de espécies a partir de consulta a publicações como o Authors of plant names (Brummitt & Powell,1992) e o site IPNI (www.ipni.org).

Com o andamento da informatização, alguns estagiários foram designados para ajudar de forma mais dedicada a digitação dos dados. Apesar disso, devido ao tamanho da coleção, via-se distante a conclusão da digitação dos dados de todos os exemplares do HB. Por isso, decidiu-se que a entrada de dados deveria ser iniciada pela coleção de tipos nomenclaturais e depois pelas coleções mais importantes do herbário (Bromeliaceae, Orchidaceae e Pteridófitas). Essa ordem permitiria que alguns objetivos fossem alcançados mais rapidamente, como a criação do site sobre a coleção de tipos. Outra decisão importante foi que novos exemplares provenientes de coletas, permutas ou doações teriam seus dados digitados antes de serem incluídos na coleção. No HB, os exemplares novos oriundos de coletas eram predominantemente provenientes de expedições à Ilha Grande, em Angra dos Reis, RJ, no projeto Flórula da Ilha Grade, desenvolvido em parceria com a Universidade do Estado do Rio de Janeiro (UERJ). A importância dessa atitude está no fato de que o volume de trabalho de entrada de dados não aumenta com a inclusão de novos materiais a coleção.

A digitação dos dados logo que o material chega à instituição mostrou-se útil também como incentivo ao uso do BRAHMS. Não apenas pela obrigação de fazê-lo: ao usar o programa, o pesquisador percebe que é muito mais fácil entrar os dados e gerar etiquetas e listas através dele, do que digitar as informações várias vezes, para gerar cada produto isoladamente.

Os estagiários envolvidos na informatização do HB eram voluntários e não geraram nenhum custo à instituição. Além disso, eram estudantes oriundos de cursos de graduação da área biológica, já tinham algum conhecimento de botânica, o que facilitava bastante o entendimento das informações constantes nas etiquetas. Entretanto, como a sua frequência era em dias variáveis e, em geral, o tempo dedicado ao HB muito pequeno (4 a 16 horas semanais), a entrada de dados foi feita lentamente. Em semanas de provas nas universidades, por exemplo, os estagiários se concentravam nos estudos e deixavam de lado a digitação.

Para avaliar o andamento da informatização, a velocidade e a qualidade (ausência de erros) da entrada de dados a digitação foi acompanhada cuidadosamente, através de anotações sistemáticas do tempo dedicado à entrada de dados e sua relação com o número de etiquetas digitadas. Os arquivos de dados produzidos foram comparados com o material original para quantificar os erros. No HB o volume de erros era pequeno.

Foi possível observar uma média de 50 registros digitados por dia de trabalho no HB. Foi notado também que o tempo ideal de trabalho na digitação é de cerca de 4 horas, ou meio período. Depois disso, há uma considerável diminuição do ritmo, devido à fadiga e falta de concentração.

A coleção de tipos recebeu atenção especial em relação à correção. Os dados foram comparados com os contidos no catálogo de tipos publicado pelo HB (Fontella-Pereira 1996), com os apresentados pelo Index Kewensis e por sites como o IPNI e o W3TROPICOS do Missouri Botanical Garden (<http://mobot.mobot.org/W3T/Search/vast.html>) Além disso, as listas de nomes científicos foram enviadas para alguns especialistas para revisão.

No ano de 2002, o Projeto “Coleções Biológicas de Apoio ao Inventário, Uso Sustentável e Conservação da Biodiversidade”, dentro do Programa de Ciência e Tecnologia para Gestão de Ecossistemas, do Ministério de Ciência e Tecnologia, decidiu apoiar a informatização do HB. O apoio visava levantar dados sobre o processo de informatização de coleções biológicas com a finalidade de subsidiar programas específicos para este setor da ciência. Esse apoio financeiro (R\$ 1200,00 por mês durante dez meses)

foi usado para o pagamento de digitadores, o que acelerou a entrada de dados neste período.

A partir do início de 2003, o ritmo de digitação foi diminuindo gradualmente, até se reduzir aos exemplares novos que dão entrada na coleção. Isso foi causado pelo final do apoio financeiro para informatização e pelo fato de o responsável pelo processo estar se dedicando ao curso de mestrado.

A informatização dos tipos está completa, assim como a família Bromeliaceae. Orchidaceae tem cerca de 35% do material digitado, e as famílias de pteridófitas, 10%.

Para disponibilizar os dados visando atender demandas de pesquisadores e da sociedade em geral, no HB decidiu-se pela criação de uma página na Internet para apresentar as informações sobre a coleção de tipos nomenclaturais. A página, feita em linguagem HTML está pronta, mas por falta de recursos, ainda não está decidido como ela será colocada no ar.

Apesar das oscilações no ritmo da entrada de dados, no HB, foram consumidos 20 meses para a informatização de cerca de 12.000 exsicatas. O gráfico (figura 1) a seguir apresenta o ritmo de entrada de dados por trimestres:

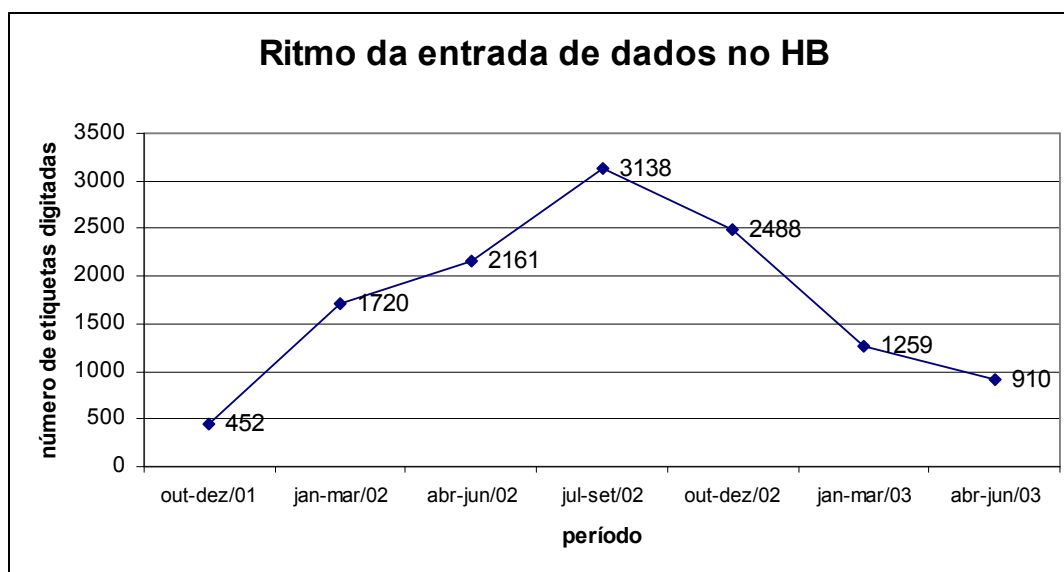


Figura 1: Ritmo de entrada de dados no Herbarium Bradeanum (HB) de outubro de 2001 a junho de 2003

Os valores empregados na informatização do HB estão representados na tabela abaixo (tabela 3). Os computadores usados na entrada de dados e para guardar o banco de dados são antigos e não foram adquiridos para esse fim específico. Por isso seu custo não está relacionado na tabela.

A tabela 4 apresenta os campos usados para a entrada de dados e a informação que deve ser colocada em cada campo.

Tabela 3: Custo da entrada de dados das etiquetas do Herbarium Bradeanum em banco de dados no período de 2001 a 2003.

Digitadores			
Período	Etiquetas	Valor	Fonte Financiadora
2001	452	R\$ 360,00	HB
2002	9419	R\$ 12000,00	CNPq
2003	2169		
Valor total: R\$ 12.360,00			
Etiquetas:		12040	
Custo por etiqueta digitada:		R\$ 0,97	

Tabela 4: Lista de campos usados no Herbarium Bradeanum (HB) para entrada de dados e informação a ser incluída

Campo	Informação
Registro	Número de tombo do material
Família	Família botânica
Gênero	Nome genérico
CF	Detalhe da identificação
Sp1	Epíteto específico
Autor 1	Autor do nome da espécie
Infra 1	Categoria infra-específica
Sp2	Epíteto infra-específico
Autor 2	Autor do epíteto infra-específico
Infra 2	Segunda categoria infra-específica
Sp3	Segundo epíteto infra-específico
Autor 3	Autor do segundo epíteto infra-específico
País	País da coleta
Estado	Estado da coleta
Município	Município da coleta
Gazetteer	Localidade da coleta

NotasLocal	Notas sobre a localidade
Alt	Altitude
AltMax	Altitude máxima (quando há faixa de alt.)
Obs	Observações de campo
Coletor	Coletor
Prefixo	Prefixo do número de coleta
Número	Número de coleta
Sufixo	Sufixo do número de coleta
Coladic	Coletores adicionais
CollDD	Dia da coleta
CollMM	Mês da coleta
CollYY	Ano da coleta
Detby	Determinador
DiaDet	Dia da determinação
MesDet	Mês da determinação
AnoDet	Ano da determinação
Lat	Latitude
N/S	Norte / Sul
Long	Longitude
E/W	Leste / Oeste
Dups	Em quais herbários há duplicatas
Inicial	Quantas duplicatas foram feitas
Avaiable	Quantas duplicatas estão disponíveis
Vernacular	Nome vulgar
Idioma	Idioma do nome vulgar
Typestat	Categoria de tipo
Bibliografia	Bibliografia sobre o tipo
Comment	Comentários

A Informatização do Herbário RB

Os pesquisadores do Jardim Botânico do Rio de Janeiro há muito tempo perceberam a necessidade da utilização de bancos de dados para facilitar e acelerar a

pesquisa botânica. Por isso, vários projetos criaram suas bases, usando programas diversos como Access e Alice, entre outros. Essas bases não eram interligadas e foram feitas para atender as necessidades de seus organizadores, sem pensar em uma conexão entre elas para que atendessem à instituição como um todo.

Na década de 1980, foi realizado um projeto piloto de informatização do Herbário RB, em parceria com o LNCC (Laboratório Nacional de Computação Científica), coordenado pela então curadora do herbário. Foram digitados os dados de cerca de 8000 exsicatas. Entretanto este projeto piloto não teve continuidade. Os dados digitados só foram encontrados em um relatório impresso, não existindo mais em meio digital.

Em 2001, o JBRJ realizou um concurso para a contratação de pessoal, entre os quais 3 analistas de sistemas que têm, entre outras atribuições, a tarefa de criar uma base de dados unificada para o Herbário RB e o corpo de pesquisadores.

Após a contratação, os analistas passaram um período de aproximadamente 18 meses conhecendo a realidade bastante particular de uma instituição de pesquisa e especialmente de suas coleções científicas.

Ao mesmo tempo era realizada uma análise dos programas disponíveis no mercado, para saber qual deles melhor atenderia as demandas do JBRJ. Neste período, o programa LUCID, para criação de chaves dicotômicas para identificação, foi traduzido para o português por um dos analistas. Essa realização deu à instituição a licença de uso do programa sem custo adicional.

Os analistas chegaram à conclusão de que nenhum dos programas disponíveis atenderia plenamente às necessidades do RB e dos pesquisadores. Decidiram então criar um novo banco de dados, usando ferramentas mais modernas.

Uma série de reuniões com pesquisadores e curadores foi realizada para apresentar e discutir pontos do banco de dados. Um assunto recorrente nessas reuniões eram os campos a serem utilizados para a entrada de informações.

Em 2003 foi firmado um convênio com a PUC-RJ para o desenvolvimento do banco de dados. Atendendo a uma resolução do Governo Federal que recomenda o uso de software livre, o banco de dados do JBRJ utilizará a ferramenta Postgree. Entretanto, ainda não foi decidido se o resultado do desenvolvimento será também livre. Ou seja, se o programa final para a informatização do RB será disponibilizado sem custo para outras instituições. Como o JBRJ dispõe de uma considerável infra-estrutura, o programa deverá utilizá-la, funcionando através da rede para permitir múltiplos acessos.

Apesar da referida infra-estrutura, o projeto de informatização prevê a aquisição de um computador para funcionar como um servidor dedicado exclusivamente para o banco de dados (para garantir a segurança das informações e a velocidade do processo) e de 16 computadores para a entrada de dados. Um projeto específico para informatização do herbário foi elaborado visando à busca de recursos para tal. O projeto foi encaminhado a empresas na busca de apoio através da lei Ruanet (nº 8.313/91) de apoio à cultura.

O projeto prevê que as exsicatas da coleção do RB sejam digitalmente fotografadas e depois tenham os dados de suas etiquetas incluídos no banco de dados e revisados através das fotografias. A previsão é de digitação de 20.000 exsicatas por mês. Todo o processo deverá durar três anos e está dividido em duas etapas de 18 meses. A primeira está orçada em R\$ 1.414.494,00 (tabela 5). Este valor inclui a compra de equipamentos, o desenvolvimento do software e a contratação de pessoal.

Mas as bases de dados já existentes não foram deixadas de lado. Elas foram analisadas, tiveram os campos unificados sempre que possível e então importadas para um banco único, em BRAHMS. A base resultante é usada para testar o funcionamento do banco de dados em desenvolvimento e vai diminuir bastante o trabalho de digitação dos dados. Em outubro de 2004, a base em BRAHMS já contava com dados de mais de 70.000 exsicatas. A tabela 6 apresenta os campos usados para a entrada de dados e a informação que deve ser colocada em cada campo.

Da mesma forma, a base de dados existente em Access sobre a coleção de tipos nomenclaturais do RB foi importada para o BRAHMS, corrigida e atualizada. Todos os tipos foram fotografados e suas imagens ligadas ao banco de dados. Após a verificação de erros e a inclusão de dados bibliográficos, a lista de tipos foi colocada na rede interna do Jardim Botânico para revisão pelos taxonomistas. As informações sobre os tipos do RB estarão disponíveis na Internet via rede SpeciesLink a partir de dezembro. A atualização da informatização dos tipos foi realizada com a atuação de 3 pessoas: 2 funcionários do JBRJ com o apoio de um analista de sistemas. Eles utilizaram um computador já disponível dentro do herbário. Os cerca de 5000 tipos do RB foram revisados em cerca de 8 meses. Durante a informatização, eram feitas cópias de segurança semanais.

Antes do trabalho, o pessoal envolvido foi treinado no uso do BRAHMS pelos analistas de sistema do JBRJ, que já haviam estudado o programa. Eles também foram os responsáveis por sanar dúvidas e resolver problemas durante a informatização.

Informatização da Xiloteca RBw

Em 1997, os pesquisadores ligados à Xiloteca do Jardim Botânico (RBw) decidiram informatizar sua coleção. Não havia apoio financeiro para a compra de equipamento ou contratação de digitadores. Os próprios pesquisadores digitaram as informações nos computadores de suas salas. A ordem de entrada foi alfabética, com as famílias a serem digitadas sendo sorteadas entre os envolvidos. A digitação foi feita em Excel, depois os dados foram importados para o Access e atualmente encontram-se BRAHMS.

No início os dados se limitavam a gênero, espécie, registro e lâmina, que acreditavam ser suficientes para auxiliar nas tarefas de curadoria. Logo foi percebido que seria mais interessante para a pesquisa inserir todos os dados, o que foi feito a partir de então. Para completar os dados, outras coleções (especialmente o herbário e outras xilotecas) onde se encontram material relacionado aos espécimes representados na xiloteca foram consultadas.

A informatização de cerca de 7000 exemplares consumiu cinco anos. Durante a realização dessa pesquisa, nenhum relatório sobre o processo foi encontrado.

Tabela 5: Estimativa do custo da entrada de dados das etiquetas do herbário do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro (RB) em banco de dados.

Etiquetas	Valor
400.000	R\$ 1.414.494,00
Valor por etiqueta digitada e corrigida:	R\$ 3,53

Tabela 6: Lista de campos usados no herbário do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro (RB) para entrada de dados e informação a ser incluída

Campo	Significado
ACCESSION	Número RB da exsicata, se houver
SUFFIX	Campo para determinar o número de duplicatas
FAMILY	Família botânica
SUBFAM*	Subfamília
GENUS	Nome genérico
CF	Detalhe da identificação.
SP1	Epíteto específico
AUTHOR1	Autor do nome da espécie

RANK1	Categoria infra-específica
SP2	Epíteto infra-específico
AUTHOR2	Autor do epíteto infra-específico
RANK2	Segunda categoria infra-específica
SP3	Segundo epíteto Infra-específico
AUTHOR3	Autor do segundo epíteto infra-específico
VERNACULAR	Nome vulgar
TYPESTAT	Categoria de tipos
COUNTRY	País
MAJORAREA	Estado
MINORAREA	Município
GAZETTEER	Localidade
LAT	Latitude
NS	Norte ou sul
LONG	Longitude
EW	Leste ou oeste
ALT	Altitude
LOCNOTES	Notas de localidade
FLOR*	Verifica a existência de flor. Marcar com um x
FRUTO*	Verifica a existência de fruto. Marcar com um x
FUSTE*	Fuste do exemplar
ALTURA*	Altura do exemplar
COLLECTOR	Coletor principal
NUMBER	Número da coleta do principal coletor
ADDCOLL	Nome dos demais coletores
COLLDD	Dia da coleta
COLLMM	Mês da coleta
COLLYY	Ano da coleta
DETBY	Determinador do exemplar
DETDD	Dia de determinação
DETMM	Mês de determinação
DETY Y	Ano de determinação

DUPS	Herbários onde há duplicatas do material
NOTES	Observações de campo
USOS	Estado de caráter etonobotânico. Lista de usos a que se destina
PROJETO*	Nome do projeto
HABITAT*	Descrição do local onde o espécime habita
HABITO*	Habito do espécimem

* campos sugeridos pelos pesquisadores do JBRJ

Análise dos questionários

Entre junho e agosto de 2004 foram enviados questionários a todos os herbários brasileiros. Esse questionário (reproduzido no anexo 1) buscava coletar informações sobre a informatização nessas coleções. Foram obtidas 34 respostas, o que corresponde a 30% dos 114 herbários ativos no Brasil (Barbosa e Peixoto, 2003). Embora seja um número aparentemente pequeno, ele está compatível com o resultado encontrado por pesquisas semelhantes realizadas por outros cientistas (Peixoto, 2003b).

A análise das respostas obtidas permitiu traçar um quadro da informatização de herbários no Brasil. Também foi conhecer as principais preocupações e dificuldades enfrentadas pelos curadores para comparar com os três herbários estudados na presente pesquisa.

Dos 34 herbários que responderam, 80% (27 herbários) estão realizando a informatização. A figura 2 apresenta graficamente esses dados. Destes, 37% (10) estão usando o BRAHMS. 11% (3) usam o Excel, outros 11% (3) usam o programa desenvolvido pela Rede Mineira de Herbários. O Access é utilizado por 7% (2) dos herbários, o mesmo número de usuários do ELCEN, desenvolvido pelo CENARGEN. Dois herbários usam bancos de dados desenvolvidos pela instituição. Também houve respostas sobre a utilização de outras ferramentas: Specify, Herbário 2.0 e SGC, com um herbário usando cada programa. O gráfico da figura 3 mostra como está distribuído o uso do programas de banco de dados nos herbários brasileiros.



Figura 2: Porcentagem de herbários em Informatização no Brasil



Figura 3: Programas mais utilizados para a informatização de herbários no Brasil

Quanto à possibilidade de importação ou exportação de dados para outros programas, 81% (22) dos herbários afirmaram que seu banco de dados tem essa capacidade. Apenas 4 instituições não podem trocar dados com outras coleções.

55% (15) dos herbários receberam o programa que utilizam sem custo. Os outros compraram pacotes prontos, desenvolveram seu próprio banco de dados ou não responderam. E 63% (17) dos herbários contam com suporte para usar o programa escolhido, seja ajuda presencial, através da Internet ou por telefone. Apenas 6 herbários (22%) afirmaram contar com apoio especializado (analistas de sistemas) para solucionar problemas com o banco de dados. Por outro lado, em 63% (17) dos herbários não há qualquer pessoa responsável pela manutenção do banco de dados, e o curador, um

pesquisador ou um aluno acumula essa função. A figura 4 apresenta o problema da falta de pessoal especializado em banco de dados nos herbários brasileiros.

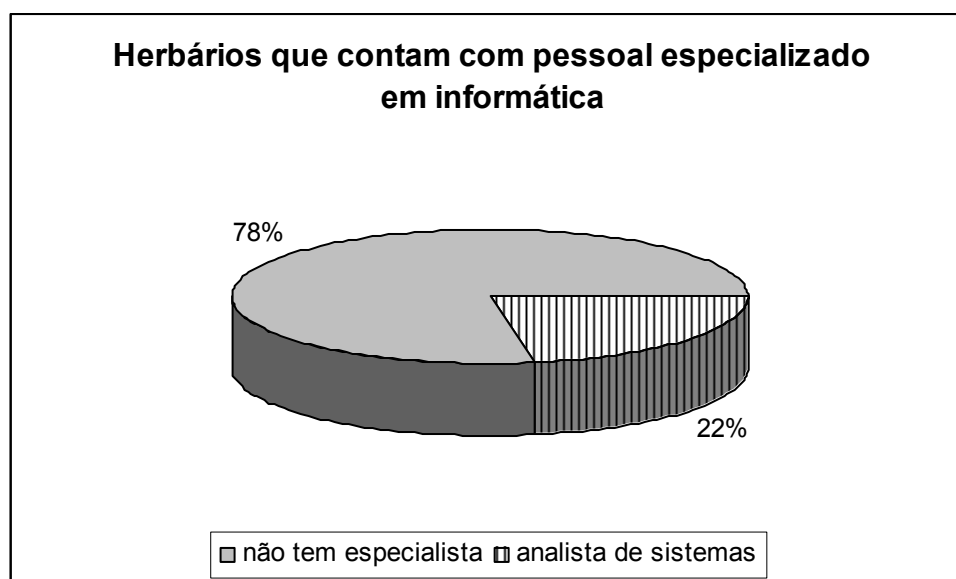


Figura 4: Porcentagem de herbários brasileiros que contam com pessoal especializado em informática para apoio à informatização

A maior parte dos herbários (52%, 14) usa um ou dois digitadores, mas o tempo dedicado à entrada de dados varia muito entre todas as instituições. Na maioria das coleções não há dedicação exclusiva à digitação, e essa tarefa consome 10 horas por semana ou menos de todos os envolvidos em 48% (13) delas. Apenas 5 herbários (18%) usam mais de 30 horas por semana em digitação.

Dos herbários que estão em informatização ou já têm sua coleção em um banco de dados, um não digitou todas as informações das etiquetas, deixando de fora os campos com as observações do hábito e detalhes da localidade. Em relação à coleção de tipos, 30% (8) dos herbários estão digitalizando os dados das etiquetas e também imagens das exsiccatas. Os outros 55% (15) se limitam às etiquetas.

Na questão sobre a prioridade da informatização, 40% (11) dos herbários dão prioridade à digitação dos dados de toda a coleção, 22% (6) priorizam a informatização dos tipos e 11% (3) priorizam os grupos de maior interesse dentro de sua coleção.

Facilitar o acesso aos dados de sua coleção não é prioridade para todos os herbários. Enquanto 59% (16) pretendem disponibilizar informações via internet, 18% (5) só permitirão consultas ao banco de dados a partir de visitas ao próprio herbário.

Os problemas mais lembrados pelos curadores foram a falta de recursos humanos especializados (48%, 13 vezes), a falta de equipamentos (18%, 5 vezes), dificuldades para

revisar e uniformizar os dados (15%, 4 vezes) e a falta de pessoal para trabalhar nas coleções (11%, 3 vezes). Também foram citados falta de planejamento prévio, falta de uma política nacional para coleções, e falta de contato entre as instituições em informatização.

Apesar das dificuldades relativas à falta de pessoal especializado, de equipamentos e de pessoas para trabalhar nas coleções, a falta de apoio financeiro – que poderia ser usado para contratar pessoal ou pagar por seus serviços – foi citada por apenas dois curadores (7%). 25% (7) deles afirmaram realizar a informatização sem nenhum recurso financeiro e 22% (5) disseram usar apenas recursos próprios. Três herbários (11%) afirmaram estar sendo informatizados apenas com recursos externos, de agências de fomento à pesquisa. O gráfico da figura 5 apresenta as maiores dificuldades apontadas pelos curadores para realizar informatização de suas coleções.

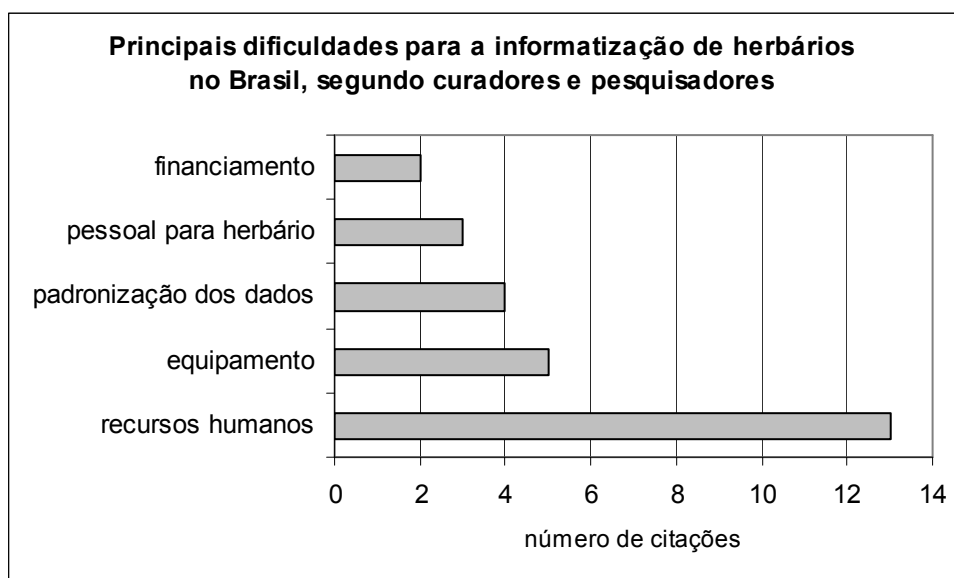


Figura 5: Principais dificuldades apontadas pelos curadores de herbários brasileiros para realizar a informatização

Os herbários que não estão sendo informatizados afirmaram que os maiores problemas são a falta de pessoal e a falta de equipamento.

Devido em grande parte aos problemas citados, mais da metade (55%, 15) dos herbários já passou por tentativas de informatização que não foram concluídas.

A previsão de todos os curadores que responderam estar informatizando suas coleções é que até o final desta década elas estejam informatizadas.

A questão 20 não está representada porque a quantidade e a variabilidade dos dados obtidos na resposta tornam difícil sua tabulação. A questão 23 teve respostas textuais, algumas bastante extensas, e também não foi possível tabular para apresentação.

Conceito de Banco de dados

Um banco de dados é um conjunto de dados com significado e que têm relação entre si. Pode ser pequeno e simples, como uma agenda telefônica, ou grande e complexo, como o arquivo de fixas de um herbário.

Elmasri & Navathe (2000) acrescentam que um banco de dados deve ter as seguintes características:

- ser uma representação do mundo real;
- ser uma coleção lógica e coerente de dados que contém algum significado. Uma página de texto é uma coleção de dados – palavras – relacionadas, mas não é um banco de dados. Da mesma forma, uma série de dados, mesmo que representativos, que não têm relação entre si não compõe um banco de dados;
- ser projetado, construído e povoado com dados que possuem um objetivo específico. Um banco de dados tem um grupo provável de usuários e algumas finalidades que interessam a esses usuários.

Os citados autores resumem: “Em outras palavras, um banco de dados possui alguma fonte, da qual os dados são derivados, algum grau de interação com eventos do mundo real e um público que está ativamente interessado no conteúdo do banco de dados”.

Elmasri & Navathe (2000) afirmam que para criar um banco de dados num computador, é necessário usar um sistema de gerência de banco de dados, ou seja, um programa que permita definir, construir e manipular o banco de dados. Esse programa pode estar pronto e ser adaptado ou pode ser criado para uso específico com algum tipo de dados.

Definir um banco de dados significa especificar os tipos de dados que serão armazenados e as estruturas e restrições para a entrada desses dados. Construir o banco é guardar os dados em um meio de armazenamento que seja controlado pelo programa. Manipular os dados significa consulta-los, atualiza-los e extrair relatórios a partir do banco.

Os autores chamam de sistema de banco de dados o conjunto de banco de dados e do programa de gerência.

Embora o software Excel não esteja incluído na definição de sistemas de gerência de banco de dados dos autores, para esta pesquisa o uso de planilhas eletrônicas para armazenar digitalmente as informações foi considerado como processo de informatização. Isso se deve ao fato de que o Excel guarda os dados com segurança e a maioria dos programas de banco de dados é capaz de importar esses dados de maneira simples. Por ser um programa amplamente conhecido e disponível, pode ser uma boa opção para as instituições que não podem implementar um sistema de banco de dados imediatamente.

Conceito de informatização

A utilização correta de bancos de dados em coleções, entre outras ferramentas produzidas pela informática, é fundamental para que os herbários possam continuar a atender a crescente demanda de informações sobre biodiversidade (Walters, 1993 e Heywood, 2001).

A introdução e recuperação dos dados das etiquetas num programa de banco de dados é denominada pela comunidade científica de informatização (Peixoto, 2003b). Mas a palavra informatizar tem um significado mais amplo. Segundo o dicionário Houaiss (Houaiss & Villar, 2001):

“informatizar: aplicar os métodos ou recursos da informática a (um fato, um problema, atividades industriais, de administração ou de documentação etc.).”

Além de banco de dados, informatizar pode se referir ao uso de várias outras ferramentas da informática, dos simples computadores à Internet. Para esse estudo, no entanto, o termo informatizar terá o significado menos amplo, referindo-se à aplicação de tecnologias de banco de dados às coleções botânicas. Neste contexto, informatizar tem o mesmo significado do termo usado em toda a bibliografia obtida em Inglês, *computerize*; ou *computerization* para informatização.

Para explicar melhor o valor dos dados e o processo intelectual pelo qual podem se transformar em conhecimento, Davenport e Prusak (1998 apud Dias & Belluzzo, 2003) fazem uma distinção. Segundo os autores “dados são um conjunto de fatos distintos e objetivos, relativos a eventos que estão ligados aos estímulos que as pessoas recebem; têm pouca relevância ou propósito, descrevem apenas aquilo que aconteceu. Informação são dados organizados de modo significativo. Já conhecimento é uma mistura fluida de experiência condensada – valores, informação contextualizada e insight experimentado – a qual proporciona uma estrutura para a avaliação e incorporação de novas experiências e informações.”

Como exemplo, os dados em uma etiqueta de herbário registram apenas o fato da coleta, onde ela ocorreu, em que data etc... Para que esses dados se transformem em informação eles devem ser interpretados ou organizados por alguém que detenha algum conhecimento prévio (o botânico, por exemplo). Uma vez transformados em informação, e transmitidos à sociedade, esses dados passam a fazer parte do conhecimento, que por sua vez pode ser aplicado na interpretação de novos dados.

Portanto, por melhor que funcione um banco de dados, ele contém apenas dados. Sua função é simplesmente permitir acessar esses dados de maneira rápida e confiável. Para que sejam úteis, esses dados devem estar disponíveis e serem trabalhados por pessoal qualificado.

De onde se extrai que: (1) Um banco de dados de uma coleção científica é uma ferramenta para auxiliar o pesquisador; jamais irá substituí-lo; e (2) de pouco serve um banco de dados completo e funcional se ele não pode ser usado pela comunidade científica.

Além de recuperar dados com mais eficiência e auxiliar em tarefas de curadoria como controle de intercâmbios e inclusão de novos exemplares, a informatização preserva a coleção. Para destacar o papel do processamento eletrônico de dados na conservação de um herbário, tendo em vista que o acesso aos dados das etiquetas é feito, de modo geral, diretamente nas exsicatas, Soper & Perring (1967) afirmam que “um herbário é um índice que sofre bastante com o uso e que é em grande parte insubstituível”. Por isso, segundo estes autores, qualquer ferramenta que reduza a necessidade de acesso direto às exsicatas tem papel importante na conservação destas.

Na verdade, a informatização de coleções é um campo de conhecimento em pleno desenvolvimento no panorama da pesquisa botânica no Brasil e se insere em uma

preocupação internacional envolvendo as coleções biológicas e seu estudo “por si” como área de conhecimento.

No cenário internacional, entre muitas atividades e ações vinculadas à informatização de coleções, é indispensável mencionar a Global Taxonomy Initiative (GTI), no âmbito da Convenção sobre Diversidade Biológica (CDB). Os objetivos do programa de trabalho da GTI envolvem diversas ações diretamente vinculadas a coleções, tais como: 1. avaliação das necessidades e capacidades em taxonomia nos planos nacional, regional e global; 2. Estabelecimento e manutenção de sistemas e infra-estruturas necessárias para obtenção, estudo e curadoria de espécimes biológicos; 3. Facilitar o estabelecimento de infra-estrutura e sistema de acesso à informação taxonômica, priorizando países de origem dos espécimes; 4. inclusão de objetivos taxonômicos prioritários nos principais programas de trabalhos temáticos da CDB, para gerar informações necessárias ao processo de decisão sobre conservação e uso sustentável da biodiversidade (Peixoto, 2003b).

Etapas da Informatização

A comparação dos processos nos três herbários analisados, à luz dos conceitos levantados pela pesquisa bibliográfica e das entrevistas realizadas permite entender melhor como se desenvolve a informatização de uma coleção científica.

Em princípio destacam-se algumas etapas, nem sempre cumpridas em todas as instituições analisadas, fato este já assinalado por Crovello (1967) e Sarasan (1981). Cada etapa tem sua importância e deve-se pensar em todas para planejar bem um projeto de informatização.

1. Escolha do software: A escolha do software de banco de dados a ser utilizado, ou ainda a decisão pela criação de um software próprio é o primeiro passo para a informatização. Para isso é necessário levantar a capacidade de processamento de dados instalada na instituição, o que inclui, principalmente, equipamento de informática e pessoal disponível para se dedicar à tarefa. Existem diversos programas prontos para a informatização de herbários. Alguns são gratuitos, outros devem ser adquiridos. A escolha deve se basear em critérios como custo, suporte oferecido, compatibilidade com outros programas (para importar e exportar dados), facilidade de uso, disponibilidade de pessoal para trabalhar ou dar treinamento. A compatibilidade é uma questão fundamental para a escolha de um programa. Com o avanço da informatização, já se projetam redes ligando os dados de diversas coleções, e isso é impossível se os programas utilizados por elas forem

incompatíveis. A decisão pela criação de um novo programa deve levar em conta que essa tarefa demanda tempo e recursos extras. E os recursos humanos necessários para desenvolvimento de software não estão disponíveis na maioria das instituições. Para conseguir êxito nessa empreitada, ela deve ser planejada com muito critério e apenas por instituições capacitadas para tal.

2. Tratamento das informações: A segunda etapa é decidir que dados serão colocados no banco, e de que maneira serão organizados. Podem ser digitados todos os dados da etiqueta ou apenas nome científico, coletor e número, data e local da coleta. Os dados podem ser separados ou agrupados em campos para facilitar a entrada de dados e a consulta à base. É interessante também criar critérios para a entrada de dados em certos campos, para criar padrões que facilitem a pesquisa e reduzam a quantidade de erros. Nomes de autor, coletor devem usar sempre o mesmo formato (por exemplo, sobrenome seguido de vírgula e depois iniciais separadas por pontos: Peixoto, F.L.). Todas as datas devem ser inseridas da mesma forma. Quando possível, o uso de bibliotecas traz muitos benefícios para o processo de informatização. Os herbários já em estado mais adiantado de informatização podem fornecer listas de nomes de coletor, informações geográficas... A internet também pode ser consultada para a obtenção de listas de gêneros e espécies. Essa etapa é importante para garantir que os dados sejam inseridos da melhor maneira possível, reduzindo a necessidade de revisão. A escolha dos campos é a base para montar a estrutura das informações no banco de dados. Não é recomendável que se criem campos muito diferentes dos que são utilizados por outros herbários. Isso dificulta o intercâmbio de informações e a importação ou exportação de dados para outros programas. Como pode ser notado nas tabelas com os campos usados por cada um dos três herbários analisados, eles são bastante uniformes. As diferenças ficam em campos adaptados para necessidades específicas de cada coleção, como o número da árvore no IAN. O banco de dados deve respeitar os campos listados no documento HISPID (Conn, 1996), para permitir a troca de informações internacionalmente. Mais recentemente tem sido usada a lista de campos contida no documento Darwin Core, mais simples que o HISPID. (<http://darwincore.calacademy.org/Documentation/DarwinCore2Draft.html>).

3. Treinamento de pessoal: Ainda antes de entrar os dados, é necessário preparar o pessoal. É interessante que o treinamento no uso do programa vá além do necessário para a digitação, incluindo também a realização de consultas e a impressão de relatórios. Em geral, esse treinamento consome uma semana. Se a instituição não dispuser de pessoal especializado em informática, pelo menos uma pessoa deve ser preparada para resolver os

problemas mais comuns do programa, como sua instalação e configuração, além de outras funções mais avançadas, como edição de dados dentro do banco e execução de cópias de segurança do banco. Essa pessoa com mais treinamento deve preferencialmente ter um vínculo formal com a instituição, para que não se perca tempo preparando alguém que pode deixar as tarefas de lado a qualquer momento.

4. Digitação: A digitação das informações é a parte mais lenta do processo de informatização da coleção de um herbário. Seu tempo de execução depende de muitos fatores, que tornam difícil uma previsão muito precisa. Entre esses fatores estão:

- a condição da parte da coleção que se está trabalhando (organização, modo de armazenamento das exsicatas, estado das etiquetas, uniformização das etiquetas - inclusive em idioma);

- a qualificação dos digitadores (digitadores profissionais como os do IAN são mais rápidos, mas aumentam os erros por não conhecerem botânica; alunos de graduação ou pós-graduação como os que incluíram dados no HB podem reduzir os erros, mas reduzem também a velocidade de entrada de dados);

- as informações que serão digitadas (naturalmente, inserir no banco de dados apenas o nome científico e os dados sobre coletor, número e data da coleta é muito mais rápido que digitar todas as informações, especialmente as observações, que podem conter grandes quantidades de texto).

- outro fator que pode alterar a velocidade da digitação é o uso de redes. Se a rede for lenta, ela afetará negativamente a entrada de dados, tornando-a mais lenta. Isso pode se tornar mais evidente quando os herbários fazem a inclusão de seus dados via internet.

Os resultados obtidos no HB são compatíveis com a média dos outros herbários que não usaram digitadores profissionais, variando de 40 a 50 exsicatas por período trabalhado. A entrada de dados do HB, com cerca de 10.000 exsicatas digitadas no ano de maior envolvimento com a informatização pode parecer lenta quando comparada àquela realizada pelo IAN e à previsão para o RB. Mas se for levado em consideração que ela foi feita com poucos recursos financeiros e com os executores dedicados apenas em tempo parcial, o valor alcançado é considerável, e torna-se ainda mais positivo porque se sabe que a cerca de metade dos herbários brasileiros têm menos de 20.000 exsicatas (Barbosa & Peixoto, 2003).

5. Correção de erros: Após a digitação de parte dos dados, eles devem ser revistos para correção de inevitáveis erros de digitação. A correção de erros de digitação pode ser feita mais tarde pelos próprios digitadores. A correção de erros nas bibliotecas de nomes de autores e coletores e em outras bibliotecas deve ser feita por pessoas com conhecimento específico, capazes de identificar erros mais complexos como diversas formas de apresentar o nome de um mesmo coletor. A correção pode prosseguir com a revisão dos nomes por taxonomistas. Para isso, listas podem ser geradas e enviadas para outras instituições.

6. Disponibilização dos dados: Quando o banco de dados estiver pronto e com as informações revisadas, ele deve se tornar acessível para o maior número de pessoas possível. Isso acelera a pesquisa botânica e preserva o material das coleções, porque reduz a necessidade de consulta direta às exsicatas. Existem muitas formas para tornar os dados disponíveis, desde a impressão de catálogos até a criação de sites na internet. Bancos de dados *on-line* são mais interessantes para a pesquisa porque podem ser atualizados rapidamente, mas não se pode esquecer que um catálogo ou livro impresso pode ser consultado em qualquer lugar, sem requerer eletricidade ou uma conexão de rede.

As formas de disponibilizar dados se constituem no resultado mais visível da informatização. Por isso devem ser usadas como forma de atrair apoios. É importante pensar a forma de disponibilizar informações ainda na elaboração do projeto também porque os dados que sejam mais úteis podem ser inseridos primeiro no banco de dados, tornando-se acessíveis mais cedo.

Estimativa de tempo de execução e custo da entrada e revisão de dados

Uma forma de calcular o tempo que será consumido no processo de informatização é uma ferramenta valiosa para planejar projetos de maneira objetiva. De acordo com as observações feitas nos herbários analisados e com as respostas obtidas com o questionário, uma hora de trabalho de um digitador significa entre 7 e 10 etiquetas incluídas no banco de dados, em média. Essa aproximação se refere ao período de 4 horas de trabalho. Em períodos mais longos, a velocidade diminui e os erros aumentam, o que faz com que o total digitado não seja proporcional ao número de horas do período mais curto.

Isso significa que um digitador, trabalhando 4 horas por dia, 5 dias por semana digitará entre 640 e 700 etiquetas por mês. O trabalho de uma pessoa digitando os dados das etiquetas durante todas as manhãs ou tardes da semana totalizará mais de 8000

exsicatas informatizadas no período de um ano. É normal que no início do trabalho o volume de dados digitados seja inferior à média, mas a velocidade tende a aumentar e até ultrapassar os valores estimados. A velocidade depende do estado das exsicatas e do treinamento dos digitadores. Isso foi percebido no Herbarium Bradeanum (fig 1, pág. 25) e pode ser notado também na tabela 1 (pág 20), que representa o custo e a velocidade da entrada de dados no herbário IAN. Por isso eles não devem ser substituídos com frequência, pois o hábito aumenta a qualidade e a velocidade de seu trabalho.

A revisão e correção dos dados digitados ocorre em média 50% mais rápido. Portanto, as 8000 etiquetas digitadas durante um ano podem ser revistas em cerca de seis meses. Se são consumidos 18 meses para digitar e revisar os dados de 8000 etiquetas, em um ano devem estar prontos os dados de cerca de 5000 etiquetas.

Em projetos que tenham a fotografia da coleção como passo inicial para a informatização, a maior parte do tempo – senão todo ele – gasto no trabalho de fotografar as exsicatas será recuperado durante a entrada de dados. Como não será necessário recorrer outra vez às exsicatas, buscando-as nos armários ou latas e guardando-as outra vez, a digitação e a revisão dos dados será feita muito mais rapidamente.

Em um relatório enviado ao CNPq a partir do projeto “Coleções Biológicas de Apoio ao Inventário, Uso Sustentável e Conservação da Biodiversidade”, Peixoto (2003c) afirma que uma coleção de até 23.000 exsicatas pode ser informatizada em um ano com custo aproximado de R\$ 28.000,00, mais gastos sociais, se já contar com dois computadores. Nesse modelo, o responsável pela informatização também trabalharia na digitação dos dados junto com outros dois estagiários. O valor sugerido de R\$ 1,20 por etiqueta digitada é um pouco mais alto do que o encontrado no IAN e no HB devido à qualificação e treinamento desejados pela autora para o gerente da informatização.

Para herbários maiores, Peixoto (2003c) chega a valor semelhante, R\$ 1,11 por etiqueta, desde que a instituição tenha estrutura adequada (até 7 computadores para coleções com mais de 50.000 exsicatas). É importante notar que esses valores, assim como os obtidos no IAN e HB, não levam em conta os gastos sociais (INSS, ISS), que podem variar muito com o tipo de pessoal utilizado. O pagamento de bolsas isenta a instituição de contribuições, enquanto que a contratação permanente ou por prestação de serviços eleva o gasto com impostos.

Outro fator que pode levar a variações no custo da entrada de dados é a situação local. Nas entrevistas semi-estruturadas, curadores da região sul e sudeste afirmaram não acreditar ser possível a contratação de pessoal para digitação pelos valores pagos pelo IAN,

enquanto o IAN teve dificuldade para encontrar estagiários de cursos de biologia para fazer o trabalho de entrada de dados.

De qualquer forma os valores para volume de etiquetas digitadas e custo da digitação são próximos aos obtidos junto aos herbários pesquisados e são úteis para auxiliar a estabelecer e avaliar cronogramas e orçamentos para a informatização de uma coleção. Como a digitação e a posterior revisão dos dados são as etapas que consomem mais tempo em todo o processo, esses valores servem para determinar com bastante precisão o tempo e o custo que será necessário para concluir o trabalho, quando o projeto estiver sendo preparado.

Situação atual e perspectivas da informatização dos herbários brasileiros

Grande parte dos herbários brasileiros está aplicando recursos de informática em suas coleções. Mesmo o uso de planilha eletrônicas como o Excel para guardar os dados das etiquetas é vantajoso. Apesar de não gerar um banco de dados para recuperar as informações e auxiliar nas tarefas de curadoria, a digitação em planilhas é um passo na direção da informatização, uma vez que a maioria dos programas usados por herbários é capaz de importar esses dados.

Apenas um dos herbários que respondeu o questionário não tem um computador disponível. Esse é o maior impedimento para a informatização. Como recursos de informática (computadores e programas) são quase universalmente disponíveis, praticamente todas as coleções podem ser informatizadas. Se não houver uma pessoa dedicada exclusivamente à entrada de dados – que é a etapa mais lenta da informatização – o processo pode ser consideravelmente lento, mas ainda assim pode acontecer.

Uma das maiores queixas levantadas pelos questionários foi a falta de pessoal qualificado para trabalhar na informatização. Magalhães *et al.* (2001) destacam a dificuldade de pagar salários de mercado para pessoal capaz de planejar, desenvolver e sobretudo implementar e manter um sistema de informação. De fato, apenas 6 herbários (25% dos que responderam) contam com apoio de pessoas ligadas à informática. O custo de pessoal para inserção dos dados também é lembrado, porque o volume de etiquetas ou fichas pode ser muito grande.

Grande parte dos herbários brasileiros não pode contar com um analista de sistema ou outro profissional qualificado para fazer a informatização. Porém, vale destacar que

apenas o Jardim Botânico do Rio de Janeiro contratou pessoal de fato. As outras instituições pesquisadas, IAN e HB, fizeram convênios ou projetos que incluíam a participação da equipe de Denis Filer e de Flávio Peixoto, respectivamente.

É muito mais simples trabalhar com pessoas que tenham conhecimento de informática do que procurar aprender com profundidade o planejamento e construção de bancos de dados. Mas é importante que qualquer pessoa envolvida no trabalho de informatização se familiarize com as práticas e necessidades de um herbário. Por isso os analistas do JBRJ passaram por um período de mais de um ano conhecendo a instituição antes de se envolver diretamente na elaboração do projeto de informatização.

Do outro lado, botânicos deveriam conhecer as possibilidades de uso de bancos de dados e outras ferramentas que a informática pode oferecer para a pesquisa em coleções biológicas, de modo que a comunicação com os responsáveis pelo banco ocorra sem problemas.

Durante reuniões de curadores realizadas nos 53º e 54º Congressos Nacionais de Botânica, e em simpósios assistidos e palestras ministradas, a comunicação entre os presentes apresentou dois problemas recorrentes: (1) o diferente nível de conhecimento dos pesquisadores em relação aos recursos da informática: enquanto alguns estão habituados à sua aplicação, outros sequer conhecem os programas disponíveis e ainda pretendem discutir questões já resolvidas, como campos mínimos para entrada de dados; e (2) as pessoas de informática e de botânica não compreendem a terminologia mínima da outra ciência.

Dois casos exemplificam esses problemas:

Sarasan (1981) cita uma pergunta feita por uma curadora a um programador: como a entrada de dados estava sendo feita por digitadores que não conseguiam separar as informações geográficas nos devidos campos (país, estado, município...) ela precisava saber se mais tarde um sistema poderia identificar e separar os dados. O programador respondeu que isso poderia ser feito. Entretanto, a autora informa que na época os computadores não eram capazes de processar esses dados da maneira requerida. Na verdade, segundo ela, uma pessoa mais habituada a lidar com programadores entenderia sua resposta como “Sim, com fundos ilimitados e sem pressão de tempo, tudo é possível”. Mais tarde os dados geográficos da coleção tiveram que ser re-incluídos na ordem correta.

Em uma reunião para planejar a informatização da coleção do RB, os analistas de sistema apresentaram a estrutura de seu banco de dados, usando o termo “árvore de hierarquia”. Esse termo criou dificuldade de entendimento entre os botânicos presentes,

que já haviam tido problemas em experiências anteriores com programas que usavam “árvores” para estruturar suas informações. A tecnologia apresentada pelos analistas, entretanto, era recente, não havia sido experimentada ainda pelos botânicos. Se o termo “árvore de hierarquia” tivesse sido substituído na apresentação por outro (“pirâmide de hierarquia”, por exemplo), muita discussão teria sido evitada.

Com o aumento do número de instituições com experiência em informatização de coleções, fica mais fácil para outras iniciarem o processo. Um herbário que tenha sucesso na implementação de um banco de dados deve ser capaz de dar suporte a outras coleções. Isso ocorre no IAN, que organiza cursos para o uso do BRAHMS para pessoal de herbários de todo o Brasil, e em especial da Amazônia e no HB, que já realizou treinamentos para pessoas ligadas a herbários do Rio de Janeiro. Essa troca de experiência acelera a informatização e evita a repetição de erros.

Todos os herbários que responderam estar em informatização têm previsão para o final do processo no máximo em cinco anos. Isso é muito positivo, pois permite prever que logo será possível realizar pesquisas e criar listas mais abrangentes sobre a diversidade vegetal no Brasil.

Próximos passos

Com a possibilidade de contar com vários herbários informatizados nos próximos anos, surgem novos passos a serem tomados para melhorar os dados dessas coleções, dando ainda melhor suporte para a pesquisa botânica. Assim caminha-se para chegar ao sentido mais amplo do termo informatizar, ao aplicar outras ferramentas da informática além dos bancos de dados.

- Interligar coleções: os bancos de dados das coleções devem ser ligados, de preferência em rede (Morse, 1974). Assim, uma atualização taxonômica chegaria mais rapidamente a todos os herbários. Isso é particularmente importante para os herbários menores, que não recebem tantas visitas de especialistas em taxonomia e têm suas identificações desatualizadas. A conexão das coleções também aceleraria o trabalho de pesquisa de dados. Um exemplo é a rede SpeciesLink (splink.cria.org.br), que recupera dados de coleções de 11 herbários, predominantemente do estado de São Paulo.
- Melhorar os dados das exsicatas: além da atualização taxonômica, muitos dados podem ser acrescentados àqueles contidos nas etiquetas. As informações geográficas podem ser conferidas e aperfeiçoadas para a criação de mapas.

Bibliotecas podem ser criadas para uniformizar campos como nomes de coletores. Referências bibliográficas podem ser acrescentadas ao banco de dados, tanto em relação aos tipos nomenclaturais quanto à base de dados de espécies.

- Planejar novos esforços de coleta: através da análise conjunta dos exemplares de várias coleções é possível saber que regiões estão bem amostradas e quais precisam ser mais estudadas (Stuessy, 1996). O cruzamento de dados de coletas com informações sobre o ambiente (solo, clima, pluviosidade...) permite que se realizem projeções sobre onde espécies que acredita-se estarem extintas podem ainda viver, além de estimar que áreas têm maior diversidade de espécies, e muito mais.

As condições ideais para herbários brasileiros realizarem a informatização

Um curador que pretenda iniciar hoje a informatização de sua coleção deve avaliar se o herbário tem as condições mínimas para realizar o trabalho. Caso contrário é possível que ele perca tempo e recursos que poderiam ser mais bem empregados em outras tarefas.

O material necessário é simples: um computador, um programa de banco de dados funcionando no computador e pessoas para digitar (Mori *et al.* 1985). Mesmo que o equipamento não seja novo e que as pessoas não tenham muito tempo para dedicar à digitação, a informatização pode ser realizada.

É importante criar um projeto que estabeleça os objetivos da informatização e, se houver, os recursos que serão investidos. O projeto deve levar em conta que grupos são mais importantes na coleção, para que eles sejam informatizados e tenham seus dados disponibilizados mais cedo, gerando assim produtos parciais da informatização. Entretanto, se o herbário tiver recursos suficientes para fazer toda a informatização em um período relativamente curto (menos de dois anos), pode não ser necessário estabelecer uma ordem especial, fazendo-se a entrada de dados pela ordem de organização da coleção.

Na escolha do programa, uma boa política é consultar outros curadores. Nas reuniões de curadores promovidas pela Sociedade Botânica do Brasil têm sido levantadas questões sobre a informatização, mas parece mais interessante realizar a consulta a herbários próximos. Se eles estiverem em etapas mais avançadas da informatização, sua experiência pode ser muito útil. Usar o mesmo programa que herbários vizinhos pode tornar o treinamento de pessoal muito mais simples, e a troca de informações mais fácil depois de concluída a entrada de dados. Mas é ainda mais importante que o programa

escolhido seja capaz de importar e exportar seus dados para outros, de modo a permitir o intercâmbio de maneira mais ampla, regional ou mesmo nacionalmente.

O pessoal envolvido na entrada de dados deve ter alguma forma de compromisso com a instituição. É muito comum que se gaste tempo treinando um estagiário que poucos meses depois deixa o herbário, e isso deve ser evitado.

É importante que exista uma pessoa responsável pelo banco de dados. Não simplesmente pela entrada de dados, mas pela manutenção de todo o sistema, e pela integridade do banco (Magalhães *et al.* 2001). Mesmo que o herbário não tenha um quadro com essa capacitação, ele pode ser contratado através de parcerias com outras instituições, como universidades. Entretanto, sempre que possível esta posição deve ser ocupada por alguém capacitado, contratado definitivamente. Esse cargo, que pode ser chamado de curador do banco de dados, tem de começar a existir, ao menos nos grandes herbários, para que a informatização não fique apenas na entrada de dados.

Por fim, para garantir que o trabalho não seja perdido em caso de defeitos, devem ser feitas cópias de segurança. Essas cópias são feitas sempre que a entrada de dados atinja um volume significativo, por exemplo, mais de mil etiquetas, e sempre que forem realizadas mudanças na base de dados, como atualizações taxonômicas. Existem muitos meios diferentes de fazer essas cópias, também chamadas de *backup*, em CDs, DVDs, fitas DAT, em servidores de rede, entre outros.

Uma política nacional de apoio às coleções, que tenha como um dos objetivos fomentar a informatização, não é apenas desejável, mas realmente necessária. Entretanto, como desde o Programa Flora nenhum outro projeto de amplitude semelhante foi estabelecido, criando até hoje dificuldades para as coleções brasileiras, não é possível esperar mais. Os herbários devem procurar iniciar e concluir seus bancos de dados, para ajudar a botânica a responder às questões urgentes relativas ao levantamento da diversidade e conservação da flora brasileira.

As tarefas de curadoria de um herbário são muitas, e raramente o curador pode se dedicar exclusivamente a elas. Na maioria dos casos, ele tem de se dividir entre a pesquisa, ensino e orientação de estudantes e a curadoria. Talvez por isso alguns curadores ainda evitem iniciar o uso de bancos de dados para gerenciar suas coleções (fato já percebido por Squirres, em 1966). Embora a informática ofereça soluções para facilitar seu trabalho, a implementação dessas soluções demandaria ainda mais tempo e dedicação, neste caso numa área que grande parte dos botânicos não domina e, às vezes, até teme. Quando a posição de curador é temporária, então, fica quase irresistível deixar a decisão de informatizar a coleção para o próximo a ocupar o cargo.

Não foi por acaso, então, que nos três herbários analisados durante a pesquisa, a informatização tomou impulso quando chegaram a estas instituições pessoas ligadas à informática. No IAN, Denis Filer, no HB, Flávio Peixoto e no RB, os analistas de sistemas contratados por concurso público.

Com a disseminação do uso de bancos de dados para herbários, no entanto, esse contexto pode mudar. Uma instituição que conte com uma estrutura melhor poderá ajudar outras menores, especialmente na definição de um projeto e no início do processo de informatização propriamente dito.

A informatização deve começar pela elaboração de um projeto. Durante a elaboração desse projeto devem ser discutidos pontos importantes que precisam ser definidos antes de começar o trabalho de digitalização dos dados. Questões como que software usar, a estrutura do banco de dados, que dados guardar e a maneira de entrar os dados têm de ser respondidas durante a elaboração do projeto, pois não devem ser alteradas depois de iniciado o processo sob o risco de anular o que já foi feito. O objetivo principal

da informatização, as maneiras de disseminar as informações e os critérios para ordenar a entrada de dados (especialmente em grandes coleções, que grupos colocar primeiro no banco de dados) também devem estar claros. As formas de distribuir as informações são importantes para mostrar a utilidade da informatização e da própria coleção para a sociedade e para atrair apoios para o projeto. E o objetivo principal e a ordem de entrada dos dados são importantes para manter o andamento do trabalho, mesmo quando há mudanças na curadoria ou na coordenação da informatização.

O projeto de informatização deve definir ainda um responsável pelo trabalho. Essa pessoa deve, preferencialmente, se tornar o administrador do banco de dados quando a informatização estiver concluída. Ou, ao menos, deve capacitar pessoal da instituição para fazê-lo. Desta forma fica mais assegurada a continuidade do trabalho de informatização em caso de mudanças na curadoria da coleção.

A discussão para a criação de um projeto é útil também para agregar os membros da instituição ao objetivo da informatização. O trabalho requer envolvimento de todos, no planejamento, na execução e na revisão e no aprimoramento das informações após a digitação no banco de dados. É importante que o projeto seja visto como institucional, para que não haja bloqueios e discussões sobre propriedade desta ou daquela parte da coleção ou de um banco de dados de determinado pesquisador. A institucionalização do projeto é valiosa também para garantir apoio a uma empreitada que costuma ser de médio ou longo prazo.

A proposta da pesquisa ora apresentada se baseou em duas questões principais: quais os maiores problemas encontrados pelos herbários para sua informatização e quais as soluções encontradas por eles. O seu desenvolvimento mostrou que o planejamento da informatização deve seguir etapas claras para que o processo ocorra sem os problemas detectados. Essas etapas devem ser consideradas na elaboração do projeto de informatização. Mostrou também que as principais dificuldades que as coleções encontram para a criação de seu banco de dados eletrônico são comuns e muitas delas podem ser resolvidas com troca de informações e experiência entre as instituições. Especialmente para a falta de pessoal especializado em informática, a solução parece ser mesmo o estabelecimento de parcerias.

Outro problema comum à maioria das instituições consultadas é a falta de apoio financeiro para a informatização. Especialmente no que se refere à contratação de pessoal para a criação do banco de dados e para a digitação das informações das etiquetas. Aparentemente, a solução para esse problema depende menos dos herbários do que de

fontes externas. Mas para que surjam apoios, é necessário que a comunidade científica torne clara para a sociedade e para os tomadores de decisão a importância das coleções e da sua informatização. Sem a criação de bancos de dados com as informações das coleções botânicas, será muito difícil respeitar os prazos estabelecidos na Estratégia Global para Conservação de Plantas de ter uma lista de espécies brasileiras até 2010, e cumprir compromissos da CDB. Também é impossível avaliar de maneira apropriada a situação atual da flora e propor medidas para a sua conservação.

Por fim, deve-se considerar que o que é normalmente chamado de informatização, ou seja, a inserção dos dados das etiquetas em um programa de banco de dados é apenas o começo de um processo muito mais amplo. A partir daí, deve-se pensar em melhorar as informações contidas nesse banco. Com dados geográficos mais apurados, um Sistema de Informações Geográficas, com ligações para dados sobre as espécies (como revisões taxonômicas) e o estabelecimento de redes de bancos de dados de coleções botânicas, para que a pesquisa possa dar um novo salto, adaptando-se às possibilidades tecnológicas do século XXI.

Referências Bibliográficas

- Barbosa, M.R.V. & Peixoto, A. 2003. Coleções Botânicas brasileiras: situação atual e perspectivas. In: Peixoto, A.L. (org.) *Coleções Biológicas de Apoio ao Inventário, Uso Sustentável e Conservação da Biodiversidade*. Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, Rio de Janeiro. 113-125.
- BGCS, IUCN & WWF. 1989. *The Botanic Garden Conservation Strategy*. IUCN-BGCS, Richmond, UK.
- Bisby, F.A. 2000. The quiet revolution: biodiversity informatics and the Internet. *Science* 289: 2309-2312.
- Borges, G.M., Gamarra-Rojas, C.F.L., Dalcin, E. & Almeida-Cortez, J. 2002. Avaliação do sistema BRAHMS (Botanical Research and Herbarium Management System) na informatização do herbário UFP. In: 53º Congresso Nacional de Botânica, Recife, PE. *Resumos*. p.414.
- Brenan, J.P.M. 1974. Internacional Conference on the use of electronic data processing in major European plant taxonomic collections. *Taxon* 23: 101-107.
- Breyer, L.M. & Tsuboi, E. 1996. *Manual de operações do Sistema Herbário*. Paralelo 15 Editores. Brasília, DF. 111p.
- Brito, M.C.W. & Joly, C.A. (Eds.) 1999. Infraestrutura para conservação da biodiversidade. In: Joly, C.A. e Bicudo, C.E.M. (Orgs.) *Biodiversidade do Estado de São Paulo, Brasil*. Vol. 7. São Paulo, S.P. 150p.
- Brummitt, R.K. & Powell, C.E. 1992. *Authors of plant names*. Royal Botanic Gardens, Kew. Richmond, UK.
- Campos dos Santos, J.L., By, R.A. & Magalhães, C. 2000. A case study of INPA's bio-db and an approach to provide an open analytical database environment. *International Archives of Photogrammetry and Remote Sensing*. 33 (b4/1):155-163
- Canhos, D.A.L., Canhos, V.P. & Souza, S. 2002. Sistemas de Informação. In: Araújo, E.L. et al. (ed). *Biodiversidade, Conservação e Uso Sustentável da Flora do Brasil*. UFRPE - Imprensa Universitária. Recife, PE. 247-251.

- CBD & BGCI. 2004. *Global Strategy for Plant Conservation*. Convention on Biological Diversity & Botanic Gardens Conservation International. Richmond, UK. 13p.
- Chapman, A. 2002. 21th century floras - new technologies to speed the process. Simpósio: *Flora Brasiliensis Revisited*. Indaiatuba, São Paulo. http://www.cria.org.br/eventos/tdbi/flora/presentations/flora_chapman.ppt, capturado em 06/01/2003
- Conn, B. J. (ed.) 1996. *HISPID3: Herbarium Information Standards and Protocols for Interchange of Data*. Version 3. Royal Botanic Gardens. Sydney, Australia.
- Crovello, T.J. 1967. Problems in the use of electronic data processing in biological collections. *Taxon* 16: 481-490.
- Daly, D.C. & Silveira, M. 2000. O banco de dados da Flora Acreana. In: 51º Congresso Nacional de Botânica, Brasília, D.F. *Resumos*. p.289.
- Dias, M.M.K. & Belluzzo, R.C.B. 2003. *Gestão da Informação em Ciência e Tecnologia sob a ótica do cliente*. EDUSC. São Carlos, S.P. 186p.
- Elmasri R. & Navthe, S.B. 2000. Sistemas de bancos de dados – fundamentos e aplicações. LTC Editora. Rio de Janeiro, RJ.
- Ferreira, G.C. & Martins-da-Silva, R.C.V. 1997. Informatização da família Vochysiaceae do Herbário IAN (Belém, PA) In: 48º Congresso Nacional de Botânica, 1997, Crato-CE. *Resumos*. p.281
- Filer, D.I. & Bennett, J.R. 2002. *Botanical Research And Herbarium Management System Version 5*. Oxford University Press, Oxford, UK. 482.
- Fontella-Pereira, J. 1996. *Tipos nomenclaturais do Herbarium Bradeanum – Angiospermae*. Herbarium Bradeanum. Rio de Janeiro, RJ. 172p.
- Gamarra-Rojas, C. 2002. A experiência do CNIP na elaboração do checklist de plantas do nordeste do Brasil. In: Araújo, E.L. et al. (ed). *Biodiversidade, Conservação e Uso Sustentável da Flora do Brasil*. UFRPE - Imprensa Universitária. Recife, PE. 256-259.
- Hall, A.V. 1972. Computer-based data banking for taxonomic collections. *Taxon* 21: 13-25.
- _____. 1974. Museum specimen record data storage and retrieval. *Taxon* 23: 23-28.
- Heywood, V.H. 2001. Floristics and monography – an uncertain future? *Taxon* 50: 361-380.
- Heywood, V.H., Moore, D.M. Derrick, L.N., Mitchell, K.A. & van Scheepen, J. 1984. The European taxonomic, floristic and biosystematic documentation system – an introduction. In: Allkin, R. & Bisby, A.A. (eds). *Databases in systematics*. London, UK. 79-89.
- Houaiss A & Villar, M.S. 2001. *Dicionário Houaiss da Língua Portuguesa*. Ed. Objetiva. Rio de Janeiro, RJ. 2925p.
- Macrander, A.M. & Haynes, R.R. 1990. SERFIS: a methodology for making multi-herbaria specimen databases a reality. *Taxon* 39: 433-441
- Magalhães, C., Santos, J.L.C. & Salem, J.I. 2001. Automação de coleções biológicas e informações sobre a biodiversidade da Amazônia. *Parcerias Estratégicas* 12: 294-312.

- Martins-da-Silva, R.C.V. & Ferreira, G.C. 1996. Informatização da família Lecythidaceae do Herbário IAN (Belém, PA) In: XVI Congresso Brasileiro de Iniciação Científica em Ciências Agrárias, 1996, Santa Maria-RS. *Resumos*. p.398
- Martins-da-Silva, R.C.V., Secco, R.S., Ferreira, G.C., Venturieri, G.C. 1998. Informatização dos Herbários MG (Museu Paraense Emílio Goeldi) e IAN (Embrapa Amazônia Oriental) In: 49º Congresso Nacional de Botânica, Salvador-BA. *Resumos*. p.134
- Monteiro Filho, C.J. & Costa, N.V. 2002. Processo de Informatização dos herbários IBGE e HRB. In: Araújo, E.L. et al. (ed). *Biodiversidade, Conservação e Uso Sustentável da Flora do Brasil*. UFRPE - Imprensa Universitária. Recife, PE. p.255.
- Mori, S.A. Silva, L.A.M. Lisboa, G. & Coradin, L. 1985. *Manual de manejo do Herbário Fanerogâmico*. Centro de Pesquisas do Cacau. Ilhéus, BA. 97p.
- Morin, N.R. & Gomon, J. 1993. Data banking and the hole of natural history collections. *Annals of the Missouri Botanic Garden*. 80: 317-322.
- Morse, L.E. 1974. Computer-assisted storage and retrieval of the data of taxonomy and systematics. *Taxon* 23: 29-43.
- Nogueira, E. 1987. *Botânica no Brasil: descrição do quadro atual/linhas de ação*. MCT – CNPq. Brasília, DF. 54p.
- Nunes, T.S., Costa, M.J.S.L., Leite, K.R.B., Oliveira, R.P., Miranda, E.B., Santos, C.C. & Queiroz, L.P. 2002. Informatização do herbário HUEFS – Herbário da Universidade Estadual de Feira de Santana – Bahia. In: 53º Congresso Nacional de Botânica, Recife, PE. *Resumos* p.412.
- Oliveira, M.C. & Senna-Valle, L. 2002. Informatização do herbário do Museu Nacional (R): Convolvulaceae Juss. In: 53º Congresso Nacional de Botânica. Recife, PE. *Resumos*. p.413.
- Peixoto, A.L. & Barbosa, M.R.V. 2002. The current situation in Brazil: General strategies, regional differences, local floras, state-level floras and herbarium databasing. In: *Flora Brasiliensis Revisited*. Indaiatuba, São Paulo. Simpósio. http://www.cria.org.br/eventos/tdbi/flora/presentations/flora_regina.ppt, (capturado em 6/01/2003)
- Peixoto, A.L. 1999. Brazilian botany on the threshold of the 21st century: Looking through the scientific collections. *Ciência e Cultura* 51 (5/6): 349-362.
- _____. (org.) 2003a. *Coleções Biológicas de Apoio ao Inventário, Uso Sustentável e Conservação da Biodiversidade*. Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, Rio de Janeiro, RJ. 237p.
- _____. 2003b. Prefácio. In: Peixoto, A.L. (org.) *Coleções Biológicas de Apoio ao Inventário, Uso Sustentável e Conservação da Biodiversidade*. Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, Rio de Janeiro. RJ. 9-17p.
- _____. 2003c. *Coleções Biológicas de Apoio ao Inventário, Uso Sustentável e Conservação da Biodiversidade*. Relatório enviado ao CNPq. CNPq. Brasília, DF. 238p.
- Peixoto, F.L. 2002. *BRAHMS 5 – Manual prático básico*. Apostila. Rio de Janeiro, RJ. 14p.

- Prance, G.T. 2001. Discovering the plant world. *Taxon* 50: 345-359
- RCS Arte Digital. 2002. Creation of the virtual herbarium of São Paulo Botanical Institute. In: *Key Innovations in Biodiversity Informatics Forum*. Indaiatuba, São Paulo. Simpósio. http://www.cria.org.br/eventos/tdbi/bis/image_ibsp. (capturado em 06/01/2003)
- Russel, G.E.G. & Arnold, T.H. 1989. Fifteen years with the computer: assessment of the "PRECIS" taxonomic system. *Taxon* 38: 178-195.
- Sarasan, L. 1981. Why Computer Projects Fail. *Museum News (American Association of Museums, Washington)*. 59 (4): 40-49.
- Smith, G.F., Steenkamp, Y., Klopper, R.R., Siebert, S.J. & Arnold, T.H. 2003. The price of collecting life: overcoming the challenges involved in computerizing herbarium specimens. *Nature* 422: 375-376.
- Sonehara, J.S., Maranhão, L.T. & Uhlmann, A. 2000. Banco de dados para organização da coleção em implantação do herbário da Universidade Tuiuti do Paraná. In: 51º Congresso Nacional de Botânica, Brasília, D.F. *Resumos*, p.290.
- Soper, H.J. & Perring, F.H. 1967. Data Processing in the Herbarium and Museum. *Taxon* 16: 13-19.
- Squirres, D.F. 1966. Data Processing and Museum Collections: A problem for the present. *Curator* 9: 216-227.
- Stuessy, T.F. 1996. Are present sampling, preservation, and storage of plant materials adequate for the next century and beyond? In: Stuessy, T.F. & Sohier, S.H. (eds.). *Sampling the green world*. 271-284. New York, USA
- Teixeira, A. 1979. *Atuação do Programa Flora até dezembro de 1980*. Relatório Datil. CNPq. Brasília, D.F. 14p.
- _____. 1980. *O Programa Flora do CNPq: histórico (1973 a 1976)*. Relatório Datil. CNPq. Brasília, DF. 10p.
- _____. 1984. O Programa Flora do Brasil – História e situação atual. *Acta Amazonica* 14 (1/2), Suplemento: 31-47.
- Thiers, B. 2002. Computerization of Large Collections: Problems, Lessons Learned, and Benefits. Simpósio: *Key Innovations in Biodiversity Informatics*. Indaiatuba, São Paulo. <http://www.cria.org.br/eventos/tdbi/flora/thiers> (capturado em 06/01/2003)
- Thomas, W. & Magill, R.E. 2002. Uma proposta para um check-list on-line das plantas neotropicais. In: Araújo, E.L. et al. (ed.). *Biodiversidade, Conservação e Uso Sustentável da Flora do Brasil*. UFRPE - Imprensa Universitária. Recife, PE. 252-254.
- Vogellehner, D. & Speck, T. 1988. DIDEA-FR, a computer program system in dialogue mode for botanical gardens. *Taxon* 37: 876-884.
- Walters, S.M. 1993. Herbaria in the 21st century: Why should they survive? *Webbia* 48: 673-387.
- Wilson, H.D. 2001. Informatics: new media and paths of data flow. *Taxon* 50: 381-387.
- Zorzetto, 2002. Rede da Vida: Software que permite integração de bancos de dados com informações sobre plantas, animais e microorganismos. *Pesquisa Fapesp* 82: 24-27.

- Boom, B.M. 1996. Societal and scientific information needs from plant collections. In: Stuessy, T.F. & Sohmer, S.H. (eds). *Sampling the green world*. New York, USA 16-27.
- Canhos, D.A.L. & Canhos, V.P. 2001. Disseminação da informação: o uso da internet. In: Garay, I. & Dias, B.F.S. (orgs.) *Conservação da biodiversidade em ecossistemas tropicais: avanços conceituais e revisão de novas metodologias de avaliação e monitoramento*. Editora Vozes, Petrópolis, RJ. 76-87.
- Dias, B.F.S. 2001. Demandas governamentais para o monitoramento da diversidade biológica brasileira. In: Garay, I. & Dias, B.F.S. (orgs.) *Conservação da biodiversidade em ecossistemas tropicais: avanços conceituais e revisão de novas metodologias de avaliação e monitoramento*. Editora Vozes, Petrópolis, RJ. 17-28.
- Heywood, V.H. 1996. A historical overview of documenting plant diversity: are there lessons for the future?. In: Stuessy, T.F. & Sohmer, S.H. (eds). *Sampling the green world*. New York, USA 3-25.
- Reid, W., Barber, C. & Miller, K. 1992. *A estratégia global para biodiversidade: guia para aqueles que tomam decisão*. Fundação o Boticário de Proteção à Natureza. Edição em Português. 37p.
- Shetler, S.G. 1971. Flora North America as an information system. *BioScience* 21:524-532.
- Shetler, S.G. 1974. Demythologizing biological databanking. *Taxon* 23: 71-100.

Sites Utilizados

BRAHMS: www.brahms.co.uk

CBD: www.biodiv.org

Darwin Core: <http://darwincore.calacademy.org>

GBIF: <http://www.gbif.org/>

GTI on line: www.biodiv.org/programmes/cross-cutting/taxonomy/default.asp

International Plant Names Index on line: <http://www.ipni.org>

SpeciesLink: <http://splink.cria.org.br/>

Taxonomia Vegetal no Brasil: <http://www8.ufrgs.br/taxonomia>

W3Tropicos: <http://mobot.mobot.org/W3T/Search/vast.html>

Anexo 1: Modelo de questionário

Questionário para coleções botânicas em informatização/informatizadas:

1. A coleção está sendo informatizada?
 sim não já está informatizada
2. Qual o programa utilizado para informatização?
3. O programa importa/exporta dados de outros programas (Excel, Access, DBase)
 sim não não sei
4. Como o programa foi adquirido?
 compra distribuição gratuita doação desenvolvimento próprio
5. O desenvolvedor do programa em uso fornece algum tipo de suporte/ajuda?
 não através da Internet ajuda presencial _____
6. Sua instituição tem pessoal dedicado exclusivamente à manutenção do programa?
 sim não
7. Quem usualmente resolve os problemas com o programa? (um professor, pesquisador, técnico, analista de sistema...)
8. Quantas pessoas digitam dados no banco?
9. Estime quantas horas por semana o conjunto de pessoas dedica à digitação?
10. Que equipamentos estão disponíveis para a informatização (indique a quantidade):
 - computadores
 - scanners
 - máquina fotográfica digital
 - máquina fotográfica
11. Quais dados foram ou estão sendo digitalizados?

- Somente os nomes de família e binômio específico []
- Nomes científicos (família e binômio) e localidade []
- Nomes científicos, localidade, coletor e data de coleta []
- Nomes científicos, localidade, coletor, data de coleta, determinador e observações []
- Todos os dados da etiqueta []

12. A informatização da coleção tipo inclui:
- Somente dados das etiquetas []
 - Dados das etiquetas e imagem da exsicata []
13. Como os dados estão sendo armazenados ?
- Floppy disk [] Hard drive [] CD [] ZipDisk [] Website []
14. Indique (numere de 1 a 6) a prioridade dada aos objetivos da informatização
- [] Informatizar toda a coleção
 - [] Informatizar a coleção de tipos
 - [] Informatizar grupos de maior interesse
 - [] Obter imagens de toda a coleção
 - [] Obter imagens da coleção de tipos
 - [] Obter imagens dos grupos de maior interesse
15. Como os dados serão disponibilizados?
- [] Site na Internet
 - [] Consultas via correio eletrônico
 - [] CD-ROM
 - [] Publicação de catálogos
 - [] Consulta no local
 - [] _____
16. Como a informatização é financiada?
- [] Com recursos da própria instituição
 - [] Com recursos de agências governamentais
 - [] Com recursos de empresas ou organizações não governamentais
 - [] Sem recursos
17. Houve tentativas anteriores de informatização? Quantas?
18. Em média, quantas exsicatas são digitadas por hora?
19. Quando a informatização em processo foi iniciada?
20. Estime que porcentagem de cada coleção de herbário está informatizada?
- Fanerógamos [] Pteridófitos [] Briófitos [] Algas [] Líquenes []
 - Fungos [] Tipos []
21. Qual a previsão para a conclusão da informatização?
- dos dados das exsicatas: _____
 - das imagens das exsicatas: _____

22. Além de financiamento, que outros problemas ou necessidades você considera importantes no planejamento da informatização?

23. Vivendo e aprendendo... Se você fosse iniciar o processo de informatização hoje, o que faria diferente?

Anexo 2: Tabulação das respostas

Respostas dadas por 34 herbários às 23 questões apresentadas no questionário do enviado entre junho e agosto de 2004

(na s próximas páginas, impressa a partir de uma tabela em Excel)

Respostas dadas por 34 herbários às 23 questões apresentadas no questionário.
A numeração das perguntas está nas células sombreadas e o questionário completo está no anexo 1.

1	2	3	4	5	6	7
s	BRAHMS	s	gratis	internet	n	aluno de mestrado
s	SGC	s	doação	presencial	n	analista de sistema
s	excel	s				tecnico
n						
s	Rede Mineira	n	desenvolvimento	internet, presencial	n	analista de sistema
n						
n						
n						
s	herbario 2	s	doação	internet, tel	n	
s	BRAHMS	s	gratis	CRIA	n	CRIA
s	BRAHMS	s	gratis	outro usuário	n	
s	BRAHMS	s	gratis	internet, presencial	s	bolsitas, pesquisadores
s	Specify	s	gratis	n	em treinamento	pesquisador
s	Access	s			n	
s	BRAHMS	s	gratis	internet	n	outros usuários
s	BRAHMS	s	gratis	internet, presencial	s	pesquisadores
s	BRAHMS	s	gratis	internet	s	analista de sistema
s	Rede Mineira	n	desenvolvimento	internet	n	técnico do herbario
n						
s		s	doação	telefone	n	professor ou bolsista
j	ELCEN	s	desenvolvimento		s	analista de sistema
s	UNB	n			n	alunos
s	herbario IAC	s	desenvolvimento	presencial	n	pesquisadores
s	ELCEN		doação		s	técnico
s	Excel	s				
s	Access	s	compra	n	n	analista de sistema
s	BRAHMS	s	gratis		n	aluno de mestrado
n						
s	Excel	s	compra	n	n	
s	BRAHMS	s	gratis	internet, presencial	s	tecnico
s	Rede Mineira	n	desenvolvimento	internet, presencial	n	técnico
n						
s	sistema para herbario	s	desenvolvimento	presencial	n	analista de sistema
s	BRAHMS	s	compra	internet, presencial	s	curador

		10							14					
8	9	pcs	scanners	maq digital	maq foto	11	12	13	prioridade					
1	0	2	2	0	0	todos	dados	disk, HD, CD	4	1	3	6	2	5
0	0	1	1	0	0	todos	dados	cd	1					
	20	1	1	1		todos		HD, CD	1	5	2	6	4	3
1	0	1	0	0	0	todos	imagem	web	1	2	5	4	3	6
		0	0	0	0									
5	30	3	1	1	3	todos	dados	HD, CD	2	0	1	3	0	0
3 a 5	5	2	2	2	1	sem obs	imagem	HD,CD,WEB	1	3	2	6	4	5
1	20	1				todos	imagem	HD, CD	1	2	6	5	3	4
1	20	1	1	1	1	todos	dados	HD, CD, Zip	1	1	1	6	1	2
3	40	1	1	1	1	todos	imagem	HD, rede	4	1	3	6	2	5
4	2	1	1	1	1	todos	dados	HD	6	1	3	5	2	4
3	4	1	0	0	0	todos	dados	HD						
2	80	2	0	3	0	todos	imagem	HD, CD						
5	8	2	0	1	0	todos	imagem	HD	4	6	5	3	4	3
8 ou 10	1	1	0	0	0	todos	dados	web	1	2	5	4	3	6
2	8	1	0	1	0	todos	dados		1	3	5	2	4	6
2	40	3	1	0	3	todos	elaboração	servidor	1	2		5	3	4
2	0 a 10	1	0	0	0	4	dados	cd	6	5	3	4	2	1
2	35	1	1	0	1	todos	dados	web	3	2	1	6	5	4
1 a 2		1	0	0	1	todos	dados	cd, ZIP	1			2		
3	10	1	1	1	0	todos	dados	HD	2	1	5	4	3	6
2	10	2	0	0	0	todos	dados	HD, disquete						
1														
4	40	4				todos	imagem	CD						
1	4	1				todos	dados	CD, web	1	2	4	6	3	5
10	20	1	1	1	0	todos	imagem	cd, ZIP						
2	25	1	1	1	0	todos	dados	cd, ZIP	5	1	3	6	2	4

15	16	17	18	19	20	21
internet	sem	0	13	2001		-
cd	proprios	2			20% fanerógamas	
local	sem	0		2001		
	sem					
site	agencias	0	10	2002	5% fanerógamas	2006
local, e-mail	sem			2004		2006
site, local	proprios, agencias	0	20	2002	70% fanerógamas e tipo	2006
site	proprios	1	7	2000	50% fanerógamas e pteridófitos	2006
site, e-mail, local	agencias, ong	0	10 ou 15	2002	100%tipos, 10% fanerogamos e briofitos	2008
e-mail, local, catalogos	proprios, agencias	1		2002	50% pteridofitos	2009
site	sem	0	10	1999	40% fanerogamos	
catalogos	sem	0	40	2003	1% fanerógamas, 10% fungos	2009
local	agencias	1	35	1995	90% da coleção	terminada
site	proprios	1	8	2003	8% da coleção geral, 90% tipos	2007
site	proprios, agencias	0		2002	menos de 1000 fanerógamas na rede mineira	
		1				
local	proprios	1	8 a 10	2003	20% fanerógamas	2006
site, local	proprios, ONG	1		1999		
site	sem	s		1998		2006
site	proprios, agencias, ONG	1	5	2002	90% pteridófitos	2006
local	própios, agencias	2		2003		
local, e-mail	própios	1	25	2003	40% fan, 100% pteri, 80% tipos	-
site, e-mail	própios, agências	s				2005
site	agências	n		2002		nã tem
catálogos, local	agencias	1	30	1997	90% fanerógamas, 100% outras coleções	2005
local	própios	1		2002		2005

22	23
pessoal	
rh	tudo
rh, equipamento	
rh equipamento	
suporte, infra-estrutura	
treinamento	
	campo UC
pessoal dedicado	iniciaria por tipos e coleções importantes
rh	mais treinamento antes de começar
rh, planejamento, armazenagem dos dados, revisão dos dados	
rh, política nacional	
	informatizar material novo primeiro
rh	fotografaria primeiro
escolha do software, bibliotecas	começaria antes
pessoal para digitação	pessoal exclusivo
pessoal de apoio, saber tratar as informações	teria um computador e uma pessoa exclusivos
recursos humanos	não usar todos os dados das etiquetas para uniformizar
pessoal para revisão	dobraria a equipe
pessoal dedicado ao herbário	programa mais simples, digitador exclusivo
recursos humanos	
recursos humanos	maleabilidade na entrada de dados
computador exclusivo	
botânicos	usar programas comuns para não ficar preso a uma pessoa ou sistema
financiamento	
financiamento e pessoal de informática	
contato entre instituições para permitir troca de dados	treinamento antes do início. Biblioteca de dados taxonômicos

**Biblioteca da vida:
amostras de plantas que documentam a paisagem brasileira**

Flávio Luna Peixoto

Copacabana é uma das praias mais famosas do mundo. A linha curva de sua orla e as ondas desenhadas em suas calçadas estampam muitos cartões postais do Rio de Janeiro e do Brasil. Mas quem passa por Copacabana hoje não consegue imaginar como o lugar era antes de sua urbanização no século passado.

Antes das construções atuais, nas areias brancas da região havia uma vegetação que não se encontra mais nas praias da zona sul do Rio de Janeiro, a restinga. Existem desenhos e fotos para mostrar como era a paisagem antes dos edifícios e da abertura dos túneis e estradas.

Mas tais imagens não são capazes de registrar muitas informações sobre os animais e plantas que viviam ali. Para encontrar essas informações, podemos recorrer às coleções biológicas. Elas são como bibliotecas da vida, guardando no lugar de livros, amostras da flora, fauna e microorganismos coletadas por cientistas nos lugares por onde eles passaram.

Na restinga de Copacabana, por exemplo, crescia uma pequena árvore que foi descrita pela primeira vez em 1893 e recebeu o nome da praia, *Eugenia copacabanensis*. A amostra original foi coletada em primeiro de fevereiro de 1870, por um botânico francês chamado Glaziou que trabalhava no Rio de Janeiro. Sabemos disso porque ela está até hoje guardada em herbário. Mas o que é exatamente um herbário?

Um herbário é uma coleção de plantas ou partes de plantas que são prensadas e secas, depois costuradas ou coladas em cartolina contendo etiquetas com seu nome

científico e outras informações, como local e data da coleta e nome de quem as coletou. Esse processo é chamado de herborização.

As primeiras coleções de plantas desidratadas pertenciam aos herbalistas, pessoas que as estudavam com fins medicinais. Essas amostras eram organizadas na forma de cadernos, e receberam o mesmo nome dado aos livros que publicavam informações sobre plantas: herbário.

Mas logo se percebeu que o volume que elas ocupavam e principalmente os cuidados especiais que precisavam (para evitar a destruição por insetos ou pela umidade, por exemplo) requeriam um local especial para mantê-las. Assim surgiram os primeiros herbários, na Europa do século XVI.

As primeiras plantas brasileiras conservadas em herbários foram coletadas no século XVII durante a colonização holandesa no nordeste e enviadas para a Holanda. Hoje estão guardadas no Museu de Copenhague, na Dinamarca.

O primeiro herbário brasileiro foi o do Museu Nacional, no Rio de Janeiro, fundado no começo do século XIX. De lá para cá, o conhecimento sobre as plantas brasileiras cresceu e hoje se sabe muito mais sobre elas. O avanço do conhecimento se reflete no tamanho das coleções dos herbários brasileiros. Hoje são quase seis milhões de amostras, depositadas em cerca de 115 herbários, distribuídos por todos os estados brasileiros.

Estas coleções têm muitas utilidades. Por exemplo, servem para ajudar a identificar plantas recém coletadas: um pesquisador pode comparar uma planta ou um pedaço dela corretamente herborizado com as amostras que já estão guardadas no herbário e assim chegar a seu nome correto.

Outra função importante é guardar testemunhos de plantas que ocorrem em uma região. Ao pesquisar a restinga de Copacabana, por exemplo, Glaziou e outros botânicos prensaram e secaram muitas plantas ou partes de plantas coletadas ali e as enviaram para diversos herbários.

Através de consultas a herbários hoje, outro pesquisador pode saber que plantas viviam naquela região, o que não pode mais ser feito em visitas à praia, já que a vegetação não existe mais. Essa informação pode ser usada para planejar um plantio de espécies nativas em uma praça, ou um projeto maior de reflorestamento.

As informações guardadas em um herbário servem também para auxiliar na determinar o status de conservação de uma espécie de planta. Isso quer dizer que podemos saber através das informações das etiquetas das amostras se ela está sendo encontrada com facilidade na natureza ou não, ou ainda se vive em ambientes que estão sendo destruídos

pelo homem. Assim as amostras ajudam a dizer se são espécies raras, ameaçadas ou até mesmo em extinção.

Como a quantidade de informações guardadas em herbários cresce constantemente, é muito trabalhoso consultar as etiquetas das amostras uma a uma (o herbário do Jardim Botânico do Rio de Janeiro tem cerca de 400.000 amostras!). Por isso, muitos herbários estão passando essas informações para computadores. Esse processo, chamado de informatização, torna mais fácil a manutenção da coleção e mais rápidas as pesquisas em herbários.

Voltando a *Eugenia copacabanensis*, a amostra que deu origem ao nome foi enviada à Europa, mas uma duplicata está guardada no Museu Nacional do Rio de Janeiro. Além disso, outros herbários têm material de coletas mais recentes. Se um botânico coletar uma planta semelhante, pode confirmar o nome comparando sua amostra com a que está guardada em uma coleção.

Aprofundando-se mais na pesquisa, ele pode saber que hoje ela não é encontrada naturalmente no bairro que lhe deu o nome, mas vive em outras áreas de restinga: Marambaia, Maricá, Massambaba e Praia do Sul, todas no Estado do Rio de Janeiro. E juntando essas informações com observações feitas em visitas aos locais, pode confirmar que ela é uma espécie ameaçada de extinção.

Como se sabia que *Eugenia copacabanensis* vivia em Copacabana e hoje é ameaçada de extinção, ela foi escolhida para ser replantada em um parque público nesse bairro, o Parque da Chacrinha. Além disso, para ajudar em sua conservação, ela é mantida em outros lugares como o Jardim Botânico do Rio de Janeiro. Conservar plantas dentro e fora de seu ambiente natural e manter coleções herborizadas são tarefas muito importantes desempenhadas pelos Jardins Botânicos.