

QUANTIFICAÇÃO DE CUSTO E TEMPO NO PROCESSO DE INFORMATIZAÇÃO DAS COLEÇÕES BIOLÓGICAS BRASILEIRAS: A EXPERIÊNCIA DO HERBÁRIO DO INSTITUTO DE PESQUISAS JARDIM BOTÂNICO DO RIO DE JANEIRO

Marcos Gonzalez¹

RESUMO

(Quantificação de custo e tempo no processo de informatização das coleções biológicas brasileiras: a experiência do herbário do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro) Este estudo resume o processo de digitação de cerca de 300.000 espécimes do herbário RB, apresentando os principais resultados em termos de tempo e custo por espécime informatizado (digitado, validado e/ou digitalizado). Tais resultados permitiram apontar o melhor modelo de gestão entre os quatro experimentados no decorrer do período estudado (29 de agosto de 2005 a 12 de julho de 2007). O projeto foi tomado como exemplo de viabilidade para iniciativas similares, e seus resultados indicados como valores de referência. Como tal, foram comparados às estimativas de custo e tempo utilizadas no documento “Diretrizes e estratégias para a modernização de coleções botânicas brasileiras tendo como foco a formação de taxonomistas e a consolidação de sistemas integrados de informação sobre biodiversidade” (Peixoto *et al.* 2006) e as principais conclusões foram que a estimativa de custo por espécime digitado está de acordo com sugerido pela literatura, porém o tempo foi superestimado.

Palavras-chave: informatização de herbários, coleções biológicas, informática para biodiversidade, gestão da informação ambiental.

ABSTRACT

(Quantification of cost and time in the process of computerization of the Brazilian biological collections: the experience of the herbarium of the Instituto de Pesquisas Jardim Botânico do Rio de Janeiro) This study summarizes the process of digitizing about 300,000 specimens of the herbarium RB, presenting the main results in terms of time and cost per computerized specimen (typed, validated and/or scanned), values that allowed to point the best management model among the four experienced in elapsing of the studied period (August 29, 2005 to July 12, 2007). The project was taken as an example of feasibility for similar initiatives, and their suitable results as reference values. As such, they were compared to estimates of cost and time used in the document “Guidelines and strategies for the modernization of brazilian botanical collections with a focus on the training of taxonomists and consolidation of integrated systems of information on biodiversity” (Peixoto *et al.* 2006) and the main conclusions were that the estimated cost per specimen typed is in agreement with this work, but the time was overestimated.

Key words: computerization of herbaria, botanical collections, biodiversity informatics, environmental information management.

INTRODUÇÃO

Conforme a Instrução Normativa IBAMA nº 160, de 27 de abril de 2007, o herbário do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, registrado no *Index Herbariorum* em 1938 sob a sigla RB, é uma coleção biológica científica, isto é, uma “coleção de material biológico devidamente tratado, conservado e documentado de acordo com normas e padrões que garantam a segurança, acessibilidade, qualidade, longevidade, integridade e interoperabilidade dos dados da coleção,

pertencente à instituição científica com objetivo de subsidiar pesquisa científica ou tecnológica e a conservação *ex situ*” (IBAMA 2007). Tratando-se de uma coleção botânica, seu material biológico são amostras de plantas — excisatas, fragmentos de madeira, frutos, lâminas com cortes anatômicos, pólen ou DNA, ilustrações, fotos, textos e objetos confeccionados com plantas. A coleção do RB, cerca de 450.000 espécimes desidratados, representa mais de 8% de todo o acervo depositado nos herbários brasileiros, estimado em 5,3 milhões de

Artigo recebido em 02/2009. Aceito para publicação em 04/2009.

¹Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, R. Pacheco Leão 915, 22460-030, Rio de Janeiro, RJ, Brasil. gonzalez@jbrj.gov.br

exsicatas (Peixoto & Morim 2003; Forzza *et al.* 2008).

Cada amostra de uma coleção biológica é um testemunho de um espécime vivo em seu habitat natural, portanto dado primário que embasa o conhecimento sobre diversidade da flora, fauna e microbiologia. As coleções são, por conta disso, a mais poderosa ferramenta para o conhecimento sistemático e o entendimento das relações evolutivas e biogeográficas da biota de uma determinada área, região ou continente. São ainda a melhor documentação permanente da composição da vegetação e de ecossistemas em áreas que se modificam ao longo do tempo, seja pela ação antrópica ou por efeito de eventos e perturbações naturais, que alteram irremediavelmente a cobertura viva (Suarez & Tsutsui 2004).

A partir da Convenção da Diversidade Biológica — um dos acordos da Conferência das Nações Unidas sobre Meio Ambiente e Desenvolvimento, realizada em 1992 no Rio de Janeiro — instituições depositárias de coleções foram encorajadas a disseminar pública e gratuitamente os dados relacionados à biodiversidade que detinham sob sua curadoria. No entardecer do século XX, no Brasil e no mundo, diversas coleções iniciaram projetos independentes de “informatização” de acervos — termo aqui entendido tanto como o processo de digitação das etiquetas dos *vouchers* quanto como digitalização, sendo este último considerado como o processo de captura digital das imagens dos espécimes (Lane *et al.* 2000; Magalhães *et al.* 2001; Canhos *et al.* 2004b; Canhos *et al.* 2006; Gaikwad & Chavan 2006). No Instituto de Pesquisas Jardim Botânico do Rio de Janeiro (JBRJ), embora ações de informatização estivessem sendo testadas desde a década de 1980, resgatar e tornar disponível todos os dados contidos em seu centenário acervo era considerada, no início deste século, uma meta longe de ser cumprida, em curto ou mesmo em médio prazo (Baumgratz *et al.* 2001). Em 2007, no entanto, a instituição disporia de um

dos maiores bancos de dados do país sobre diversidade florística nacional, o Jabot, um resultado direto do projeto institucional “Informatização do Acervo do Herbário do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro”.

Este estudo apresenta os resultados alcançados por esse projeto, em termos de custos, tempo de informatização de um espécime e modelos de gestão, com o intuito de fornecer referências para o desenvolvimento de experiências similares e de políticas públicas para o setor de coleções biológicas no Brasil. De modo geral, são diligências que dependem de indicadores para corroborarem metas. Entretanto, a literatura sobre o assunto, além de escassa, baseia-se em casos que seguem metodologias diferentes para o cálculo de custos e tempos (Canhos *et al.* 2004a). Assim, para tornar os resultados mais comparáveis — e portanto mais úteis — buscou-se apresentar no decorrer do texto não só os dados quantitativos primários que diretamente os compuseram, mas também um resumo dos processos operacionais que indiretamente os moldaram.

Com base nos valores obtidos, foram avaliados, por fim, custo e tempo utilizados no documento “Diretrizes e estratégias para a modernização de coleções botânicas brasileiras tendo como foco a formação de taxonomistas e a consolidação de sistemas integrados de informação sobre biodiversidade” (Peixoto *et al.* 2006), que estima serem necessários 5,08 milhões de reais e 10 anos para digitar 3,32 milhões de espécimes botânicos, um custo unitário, portanto, de R\$ 1,52 (um real e cinquenta e dois centavos), e um volume de 332 mil espécimes por ano.

MATERIAIS E MÉTODOS

Orçado em R\$ 2.823.329,60 (dois milhões, oitocentos e vinte e três mil, trezentos e vinte e nove reais e sessenta centavos), o projeto “Informatização do Acervo do Herbário do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro.” foi dividido em duas etapas.

A primeira, executada entre abril de 2005 e julho de 2007, foi patrocinada pela Petrobras — Petróleo Brasileiro S.A., através da Lei Federal de Incentivo à Cultura (Lei Rouanet, nº 8.313/91), sob o número Pronac 03-5411 (para maiores detalhes sobre o projeto, consultar JBRJ 2003). A segunda etapa permanece em captação de recursos.

O objetivo principal é recuperar e disponibilizar, de maneira eficiente e ágil, grande parte do acervo do RB. Além disso, o projeto possui outras metas que também consomem recursos, como melhorar a infra-estrutura do herbário e o estado de conservação da coleção, disponibilizar informações e imagens de parte dos espécimes na Internet, produzir peças de divulgação científica sobre a importância da coleção e formar mão de obra especializada no gerenciamento e manutenção do patrimônio científico.

Informatizar um acervo do tamanho do RB foi um desafio para os técnicos e cientistas do Jardim Botânico. A preparação requereu reuniões envolvendo a coordenação do projeto, a equipe de informática, a curadoria e o conselho de curadores do herbário, a fim de se estabelecerem diretrizes que viessem a nortear o processo nos meses seguintes (Peixoto 2005). Questões importantes decididas durante o planejamento foram, por exemplo: a opção pela terceirização da informatização, por meio da contratação de uma empresa especializada; o estabelecimento dos campos das etiquetas de tombamento a serem transcritos (Fig. 1); definir aqueles campos que deveriam ser duplamente digitados (número de tombo, coletores principais, número da coleta e data, que na Figura 1 reaparecem ao final do formulário); e aqueles que deveriam ser selecionados de uma lista, e não digitados letra a letra (p. ex. determinação e unidade geopolítica).

Figura 1 – Tela de entrada de dados do sistema desenvolvido para o projeto de informatização do acervo do herbário RB.
Figure 1 – Data entry screen of the system developed for the project of digitization of the RB herbarium collection.

Para a entrada dos dados, foi montado um laboratório de informatização com 16 computadores ligados ao servidor central do Jabot, operando no modelo cliente-servidor. Foram contratados e treinados para o projeto trinta digitadores profissionais e dois supervisores, um atuando na chefia dos digitadores e outro, biólogo, atuando como interlocutor do herbário. Cada digitador cumpria turnos de 4 horas e os supervisores, 8 horas diárias, todos trabalhando cinco dias por semana. Serviços como a manutenção de exsicatas e o trânsito de material da coleção foram prestados por 18 jovens de baixa renda contratados pelo Centro Sócio-ambiental do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro. Para a digitalização das imagens, foram alocados dois computadores, cada um ligado a um escâner e operado por um técnico, supervisionados por outro biólogo, durante três meses.

Os requisitos do processo da digitação exigiram a criação de um programa de entrada de dados com foco na usabilidade e ergonomia. A empresa contratada para a digitação recomendou que a interface (Fig. 1) fosse a mais simples possível, ou seja: que dependesse minimamente do *mouse*; que promovesse a navegação pelos campos por meio das teclas TAB e Enter, para o campo seguinte, e Shift TAB, para retornar um campo; que oferecesse um atalho via teclado, de preferência um botão, para salvar o registro (foi programada a tecla F12); que os campos com vocabulário controlado (táxon e unidade geopolítica) fossem auto-completáveis (isto é, à medida que letras fossem digitadas, o sistema mostraria os valores da lista que combinassem com a sequência); que as listas dos campos auto-completáveis pudessem ser abertas pelo teclado (foi programada a sequência [ALT] [Seta para baixo]); que, após salvar um registro, o sistema avisasse que o salvou e retornasse para uma tela em branco, para o próximo registro.

Alguns controles foram introduzidos no banco de dados para que o processo pudesse ser monitorado. Foram fornecidas a cada digitador e supervisor uma conta e uma senha,

de tal forma que a criação de lotes e a digitação de cada espécime ficassem associadas a um usuário, permitindo a geração de relatórios, em tempo real, para gerenciamento da produtividade de todos em conjunto e de cada digitador individualmente.

As estatísticas que serão apresentadas adiante foram obtidas com base na entidade ‘Testemunho’ do banco de dados Jabot. Esta entidade, ou tabela, simboliza qualquer testemunho de uma coleta, principalmente os espécimes das coleções botânicas, mas também citações da literatura, coletas não tombadas ou fotografias de indivíduos. Os seguintes atributos (campos) desta entidade foram utilizados nos cálculos:

- ‘iniciocriacao’: atributo que armazena automaticamente o momento, com precisão de microssegundos, em que um registro é aberto para inclusão;
- ‘datacriacao’: atributo que armazena automaticamente o momento final, com precisão de microssegundos, em que um registro é salvo no banco de dados. É considerada, para efeitos lógicos, a data e a hora de criação do registro. A diferença em relação ao campo anterior permite calcular o tempo exato necessário para a digitação do espécime;
- ‘criadopor’: registra automaticamente o código do usuário que digitou o espécime, sendo uma chave estrangeira da entidade ‘usuario’, comum a todos os sistemas de informação do JBRJ. No caso do Jabot, representa o código do digitador;

O fluxo de trabalho implantado no processo da digitação, que tornaria possível rastrear um lote de exsicatas, pode ser resumido nos seguintes passos:

1. um lote de exsicatas da família botânica sendo digitada seria retirado do herbário por um técnico e entregue ao supervisor;
2. o supervisor organizaria o material e abriria, por meio do Módulo Supervisor do sistema, um novo lote, informando o número de espécimes e o digitador responsável pela sua digitação, selecionado automaticamente

entre os digitadores ociosos, num sistema de fila única;

3. o digitador selecionado faria a transcrição dos dados das etiquetas para o banco de dados, excisada por excisada, afixando-lhes, na capa, etiquetas verdes (significando que foram digitados) e separando as duplicatas, que seriam distribuídas para outros herbários. Ao final da digitação, o lote era devolvido para o supervisor;

4. o supervisor daria baixa do lote no sistema e entregaria ao digitador um novo lote;

5. o supervisor imprimiria, por meio do sistema, etiquetas de código de barras para todos os espécimes do lote recém devolvido e encaminharia para o herbário o lote e as etiquetas;

6. um técnico do herbário se encarregaria de afixar nos espécimes as etiquetas de código de barras (contendo também seu número de tombo na coleção, para facilitar a indexação), antes de devolvê-los aos respectivos armários.

O período analisado foi de 29 de agosto de 2005, data oficial do início da digitação, até 12 de julho de 2007, data do último relatório de atividades disponível. Foram desconsiderados 313 dias com pouca ou nenhuma atividade, tais como os sábados, domingos, feriados, pontos facultativos, férias ou recessos dos digitadores, além de interrupções por motivos técnicos, como a fumigação anual do herbário, problemas com a rede de computadores, no software ou na rede elétrica.

O período estudado foi dividido em quatro fases. A primeira, de 29 de agosto de 2005 a 10 de fevereiro de 2006, contou com recursos financeiros plenos e envolvia profissionais do mercado como digitadores, que recebiam por produção (quantidade de espécimes digitados). A gestão de pessoal também foi terceirizada, ou seja, os digitadores eram contratados e pagos pela empresa prestadora do serviço, porém não havia um gerente de projeto dedicado, tarefa que foi realizada pelo coordenador do projeto. Em dezembro de 2005, quando se aproximava o final desta primeira fase, uma auditoria de avaliação qualitativa da digitação foi realizada. O número de erros encontrados foi considerado alto pela

coordenação do projeto, o que acarretou em mudanças no modelo de gestão até ali aplicado.

Na segunda fase, de 6 de março a 11 de julho de 2006, os digitadores passaram a ser predominantemente biólogos, recebendo salário fixo, e os recursos disponíveis permaneceram plenos. A gestão de pessoal continuou sendo realizada pela empresa contratada, mas introduziu-se a figura de um gerente de projeto.

Na terceira fase, de 12 de julho de 2006 a 10 de janeiro de 2007, os recursos financeiros foram suspensos, enquanto não se resolviam questões técnicas relacionadas à qualidade das imagens, no processo simultâneo de digitalização de parte do acervo. A curadoria do herbário, procurando evitar uma descontinuidade na digitação, assumiu gerência do projeto e a gestão de pessoal, mantendo com seus próprios recursos um número reduzido de digitadores, predominantemente biólogos.

Na quarta e última fase da digitação, de 11 de janeiro a 12 de julho de 2007, com a digitalização concluída e as imagens disponibilizadas na Internet, os recursos retornaram, a gestão de pessoal permaneceu sendo realizada pela curadoria, que contratou um gerente biólogo para o projeto. O perfil dos digitadores continuou sendo predominantemente de biólogos.

RESULTADOS

Tempo de digitalização por espécime

Foram digitados 291.630 espécimes (TE) nos 369 dias de trabalho do período estudado, cerca de 810 por dia. A média de digitação variou muito ao longo do tempo (Fig. 2), estando diretamente relacionada a dois fatores principais: à quantidade e à produtividade dos digitadores.

Embora o projeto tenha previsto 30 digitadores por dia (15 por turno), quase nunca se trabalhou com a capacidade total, uma decorrência da alta rotatividade entre estes profissionais. Levando-se em conta apenas as fases 1 e 2, que contaram com recursos financeiros plenos e gestão privada de pessoal, observam-se algumas diferenças em termos de rotatividade: a fase 1 apresentou uma média de 20 digitadores profissionais por dia, com

desvio padrão (dp) de 4, enquanto que a fase 2 obteve uma média de 17 digitadores biólogos/dia, com dp de 3.

Quanto à produtividade da digitação, pode-se calculá-la de duas formas, representando dois pontos de vista. A primeira se apóia na diferença entre os momentos inicial e final da criação de um registro, um ponto de vista mais mecanicista. A segunda, divisão da janela de tempo disponível para a digitação pela quantidade de espécimes efetivamente digitados, fornece uma dimensão mais humana do processo.

Pela avaliação mais precisa, obtêm-se uma média geral de 3'09" (três minutos e nove segundos) para a digitação de um espécime. Entre os digitadores profissionais (fase 1), a média é 2'36", e entre os predominantemente biólogos (fases 2, 3 e 4), 3'50", sendo 3'37" para a fase 2, 8'29" para a fase 3 e 3'42" para a fase 4. Comparando-se as primeiras semanas das fases 1 e 2 (Fig. 3), observa-se que os digitadores profissionais não têm apenas uma produtividade média melhor que os biólogos, eles sempre foram mais velozes.

Para se calcular a produtividade pelo método menos mecanicista, foi preciso antes

calcular-se o tempo total disponível para a digitação (TDD), que depende do número de digitadores por turno: 1.188.240 minutos. Dividindo-se então TDD por TE (total de espécimes digitados, 291.630), obtêm-se 4'04" (quatro minutos e quatro segundos). Este é o tempo médio real de digitação de um espécime, que engloba as questões em torno do serviço, além da digitação propriamente dita (3'09"). Portanto foram gastos 55" (aproximadamente 23% do tempo total) com o fluxo do processo e a resolução de problemas técnicos e científicos. A Tabela 1 mostra o tempo real para as diversas fases do processo.

Tomando-se quatro minutos como um tempo médio para digitação de um espécime (15 por hora, 90 por dia), é factível supor que em um mês (20 dias), trabalhando 6 horas por dia, um digitador seria capaz de digitar, em média, 1800 espécimes. Quanto à digitalização, com dois técnicos trabalhando simultaneamente, foram obtidas 10.646 imagens de espécimes em três meses, uma média próxima a 1800 espécimes digitalizados por mês por profissional. Digitalizar um espécime, então, requereu o mesmo tempo necessário para digitar os dados de sua etiqueta de tombamento.

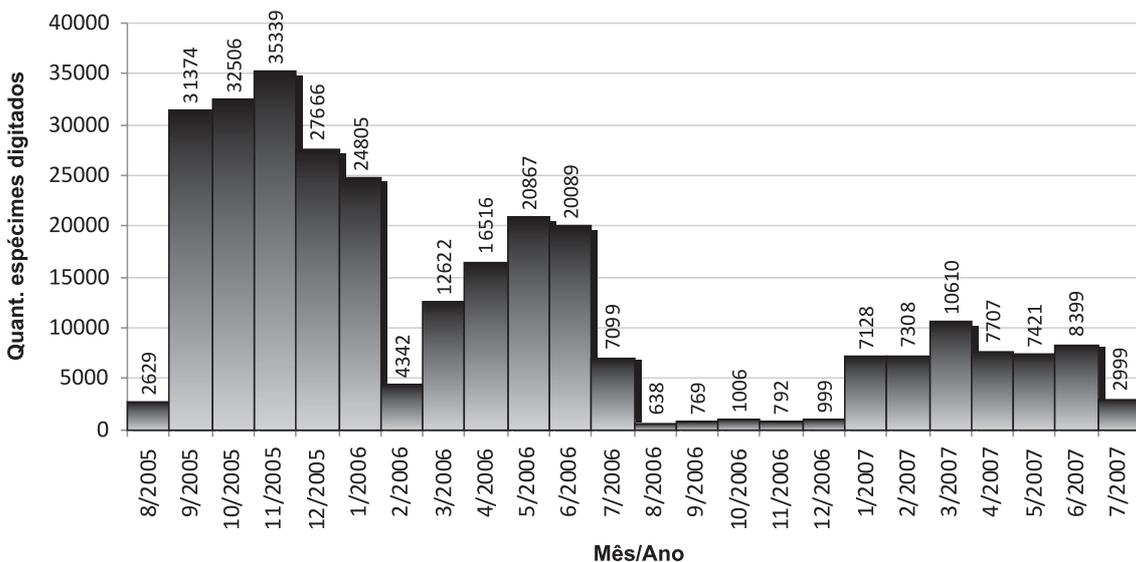


Figura 2 – Espécimes digitados por mês durante as quatro fases da primeira etapa do projeto de informatização do herbário do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro.

Figure 2 – Specimens typed per month during the four phases of first stage of the project of digitization of the Instituto de Pesquisas Jardim Botânico do Rio de Janeiro herbarium.

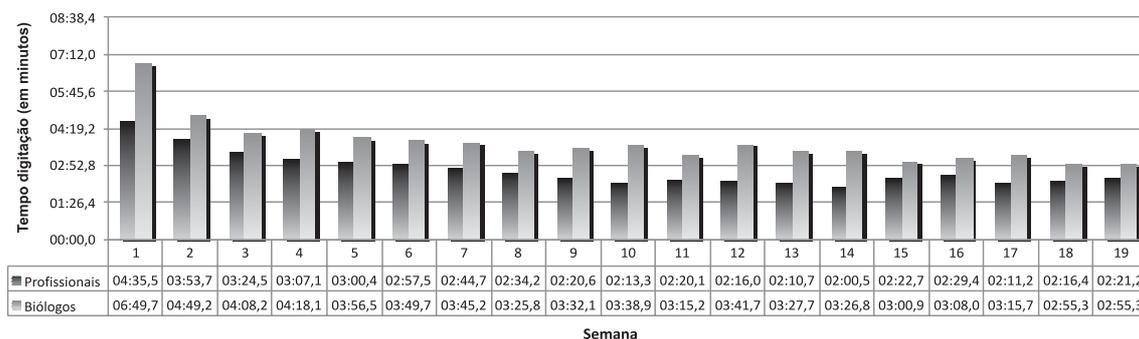


Figura 3 – Tempo médio da digitação de um espécime nas 19 primeiras semanas das Fases 1 e 2.

Figure 3 – Average time of entering a specimen in the first 19 weeks of the phases 1 and 2.

Para estimar o esforço de verificar e corrigir dados de espécimes botânicos digitados, pode-se apoiar na auditoria realizada ao final da Fase 1. Nesta experiência, de um mês, um botânico avaliou 983 espécimes, aproximadamente a metade da quantidade que, em média, poderia ser digitada num mesmo período de tempo. Validar dados, portanto, parece requerer o dobro do tempo necessário para digitá-los.

Custo por espécime digitado e digitalizado

Dos recursos investidos na primeira etapa do projeto (R\$ 1.455.814,78), os gastos com informatização propriamente dita, incluindo equipamentos, infraestrutura, pessoal, administração e impostos envolvidos com a digitação dos dados de 291.630 espécimes e a digitalização das 10.646 exsiccatas, consumiram R\$ 523.443,29 (36%) do projeto. A Tabela 2 mostra como foram investidos tais recursos; e a Tabela 3, apresenta o custo do espécime digitado e digitalizado.

Efetividade: apropriação pela sociedade

O banco de dados Jabot dispunha, em novembro de 2008, de quase 410.000 registros de espécimes botânicos, tanto de coleções de herbário, quanto indivíduos vivos, 77% deles obtidos com recursos do projeto de informatização do herbário, 21% por meio da importação de bases de dados que, historicamente, estavam dispersas na instituição, e 2% digitados diretamente no aplicativo web do sistema Jabot. Os dados podem ser conferidos no sítio

eletrônico desenvolvido para o projeto (<http://jabot.jbrj.gov.br>) e compõem, também, outros projetos de integração de dados de coleções botânicas, taxonomia e conservação. Para algumas iniciativas, o Jabot fornece visões, filtradas por tipo de coleção, táxon, procedência e/ou coletor, tanto internas, como o sítio do banco de DNA do Jardim Botânico, quanto de instituições parceiras, como o Parque Nacional de Itatiaia (<http://www.jbrj.gov.br/jabot/pni>).

Um convênio importante foi o estabelecido em 2005 com o *speciesLink* (<http://splink.cria.org.br>), um sistema distribuído de informação financiado pela FAPESP e desenvolvido pelo Centro de Referência em Informática Ambiental (Canhos *et al.* 2005), que integra, em tempo real, dados primários de coleções científicas de quase 160 coleções e sub-coleções biológicas brasileiras. O JBRJ foi a primeira instituição fora de São Paulo a fornecer dados para o portal e o primeiro provedor a conectar-se *on line* com o *speciesLink*. Os dados das três coleções do JBRJ (herbário, xiloteca e banco de DNA) disponibilizadas no portal representavam, em novembro de 2008, 18% de todos os registros sobre plantas e fungos macroscópicos e 12% de todos os registros do portal. Esta parceria vem permitindo importante intercâmbio de tecnologias para administração e disseminação da grande massa de dados, nas áreas de rede de informações, limpeza de dados, padrões e protocolos para coleções digitais, além de projetos de modelagem e análise ambiental com base nestes dados (Gonzalez 2007).

Tabela 1 – Diferenças entre as quatro fases da primeira etapa do projeto de informatização do acervo do herbário do Instituto de Pesquisas Jardim Botânico (RB), e os resultados obtidos em termos de tempo exato (a) e tempo real (b) para digitação de um espécime, além da relação entre ambos.

Table 1 – Differences among the four phases of the first stage of the project of digitization of the Instituto de Pesquisas Jardim Botânico do Rio de Janeiro herbarium (RB), and the results obtained in terms of exact (a) and realistic time (b) for entering a specimen, and the ratio between both.

| Fase | Perfil da equipe de digitadores | Tipo de salário | Recursos financeiros | Gestão de pessoas | Gerente de projeto | (a) | (b) | (b)/(a) |
|------|---------------------------------|-----------------|----------------------|-----------------------|--------------------|-------|--------|---------|
| 1 | Profissionais de mercado | Produtividade | Plenos | Empresa terceirizada | Não | 2'36" | 3'10" | 22% |
| 2 | Predominância de biólogos | Fixo | Plenos | Empresa terceirizada | Sim | 3'37" | 4'57" | 36% |
| 3 | Predominância de biólogos | Fixo | Limitados | Curadoria do Herbário | Não | 8'29" | 19'08" | 125% |
| 4 | Predominância de biólogos | Fixo | Plenos | Curadoria do Herbário | Sim | 3'42" | 5'58" | 61% |

Tabela 2 – Resumo dos investimentos na informatização do herbário do Jardim Botânico, em reais e em dólares, conforme JBRJ (2003). Por “custos” (coluna 1), entende-se os fixos e os variáveis, incluindo infraestrutura. Para cálculo dos gastos com administração por rubrica, aplicaram-se as mesmas percentagens da coluna 3 ao totais gasto com administração da informatização — leia-se coordenação, produção executiva, gerência, elaboração, material de consumo e outras despesas administrativas relacionadas. Para conversão para dólares, foi utilizada a taxa US\$ 1 para R\$ 2,25, obtida pela média aritmética do valor dessa moeda em abril de 2005 (segundo Brasil 2005) e julho de 2007 (Brasil 2007).

Table 2 – Summary of investments in digitization of the herbarium in Real and in dollars, according to JBRJ (2003). “Costs” (column 1) means the fixed and variable ones, including infrastructure. To calculate expenses by item, the same percentages in the column 3 was applied to the total spent on administration of the digitization - coordination, executive production, management, development, supplies and other administrative expenses. The rate used for conversion to dollar was US\$ 1 to R\$ 2.25, obtained by the arithmetic mean of the values of that currency in april of 2005 (according Brasil 2005) and in july of 2007 (Brasil 2007).

| Rubrica de custos | Execução (R\$) | % Total da Execução | Administração | Total (R\$) | Total (US\$) |
|-------------------|----------------|---------------------|---------------|-------------|--------------|
| (a) Digitação | 378.709,52 | 87,98% | 81.837,43 | 460.546,95 | 204.687,53 |
| (b) Digitalização | 51.715,55 | 12,02% | 11.180,79 | 62.896,34 | 27.953,93 |
| Total | 430.425,07 | 100% | 93.018,22 | 523.443,29 | 232.641,46 |

A pesquisa interativa e integrada facilitou o cruzamento e reunião das informações permitindo a elaboração de notas, relatórios técnicos, pareceres, artigos científicos e de divulgação, que contribuíram com atividades de gestão da biodiversidade brasileira e resultaram em pesquisas que tratam de questões botânicas, como identificação de espécies novas para a ciência (Meyer 2008), florística (Oliveira & Godoy 2007), conservação de espécies, ecossistemas e biomas (Martinelli *et al.* 2008)

e temas relacionados à informatização de coleções (Vargas *et al.* 2007). Dois bons exemplos da utilização dos dados foram os eventos realizados para a Revisão da Lista de Espécies Ameaçadas de Minas Gerais, onde os pesquisadores trabalharam diretamente ligados ao acervo para obter informações sobre as espécies, e o Workshop sobre as Coleções Amazônicas, realizado em Manaus (<http://ppbio.inpa.gov.br/Port/colecoes/workshop>), ambos de 2006.

Tabela 3 – Custos, em reais e em dólares, por espécime digitado e espécime digitalizado. Para conversão para dólares, foi utilizada a taxa U\$ 1 para R\$ 2,25, obtida pela média aritmética do valor dessa moeda em abril de 2005 (segundo Brasil 2005) e julho de 2007 (Brasil 2007).

Table 3 – Costs, in Real and in dollars, for typed and scanned specimen. The rate used for conversion to dollar was US\$ 1 to R\$ 2.25, obtained by the arithmetic mean of the values of that currency in april of 2005 (according Brasil 2005) and in july of 2007 (Brasil 2007).

| | Gastos totais, incluindo administração (g) | Quantidade (q) | Custo em R\$ (g/q) | Custo em U\$ |
|----------------------------------|--|----------------|--------------------|--------------|
| Espécime digitado | R\$ 460.546,95 | 291.630 | R\$ 1,56 | U\$ 0,69 |
| Espécime digitalizado | R\$ 62.896,34 | 10.646 | R\$ 5,90 | U\$ 2,63 |
| Espécime digitado e digitalizado | | | R\$ 7,46 | U\$ 3,31 |

DISCUSSÃO

Os valores de tempo e custo de informatização ficaram dentro dos limites encontrados na literatura para empreitadas similares (Smith *et al.* 2003; Canhos *et al.* 2004a; Peixoto 2005), atingindo as metas previstas com os recursos disponíveis para período estabelecido.

Conforme os dados apresentados na Tabela 1, o arranjo de gestão aplicado na fase 1 — digitadores profissionais do mercado, recursos disponíveis plenos, gestão terceirizada de pessoal e gerência geral realizada pelo coordenador do projeto — foi o que apresentou a melhor produtividade no quesito tempo, tanto no cálculo mais preciso, quanto no mais realista, e também na razão entre eles, tornando-o o mais recomendado entre todos. Como os digitadores profissionais da fase 1 tinham seus salários vinculados à produtividade e os demais recebiam salário fixo, é possível que a forma de pagamento explique o melhor desempenho. A influência deste fator não pôde ser medida, pois não houve grupos de biólogos recebendo por produtividade nem digitadores profissionais com salário fixo.

Sem um instrumento que medisse a qualidade dos dados digitados em todas as fases, como o que fora proposto pela equipe de informática de tecnologia de informação (TI) na primeira versão do projeto (Gonzalez *et al.* 2003), não foi possível verificar se a produtividade na digitação era inversamente proporcional à qualidade dos dados, como se suspeitou no final da fase 1, após a auditoria.

CONCLUSÕES

A primeira etapa do projeto de informatização do RB não foi um processo perfeito, sendo a falta de mecanismos de monitoramento da qualidade dos dados seu ponto mais fraco. O projeto mostrou, por outro lado, que iniciativas com tal propósito são viáveis, além de oportunas e muito úteis para a sociedade.

O custo de digitação por espécime calculado para o projeto, que foi de R\$ 1,56, permite afirmar que o valor utilizado como referência no documento de Peixoto *et al.* (2006), que foi de R\$ 1,51, está bem estimado.

A meta de 332 mil espécimes de cerca de 17 instituições digitados por ano, também sugerida por Peixoto *et al.* (2006) parece subestimada, uma vez que a experiência do JBRJ mostrou ser factível digitar cerca de 150 mil espécimes por ano em apenas uma instituição. Se realizada simultaneamente em outras 16, seria possível digitarem-se cerca de 2,5 milhões de espécimes por ano. Em outras palavras, a meta para 10 anos, de 3,32 milhões, poderia ser atingida em dois. Se todos os espécimes fossem também digitalizados, seria necessário o dobro do tempo. E se além de digitados e digitalizados, os dados fossem também auditados quanto à qualidade, seriam necessários mais 4 anos, perfazendo um total de 8 anos — e ainda assim seria um prazo menor que o estimado em Peixoto *et al.* (2006).

AGRADECIMENTOS

A Arianne Peixoto, Claudia Barros e Renato Pizarro Drummond, pelas contribuições ao texto.

REFERÊNCIAS BIBLIOGRÁFICAS

- Baumgratz, J. F. A.; Guedes-Bruni, R. R.; Pereira, T. S. & Andreato, R. H. P. 2001. O Herbário como fonte de conhecimento e desafios contemporâneos. *In*: Silva, N. M. F.; Carvalho, L. d'A. F. & Baumgratz, J. F. A. (eds.). O herbário do Jardim Botânico do Rio de Janeiro - um expoente na história da flora brasileira. Instituto de Pesquisas Jardim Botânico do Rio de Janeiro/Ministério do Meio Ambiente, Rio de Janeiro. Pp. 125-131.
- Brasil. 2005. Ato Declaratório Executivo Cosit nº 15, de 3 de maio.
- Brasil. 2007. Ato Declaratório Executivo Cosit nº 20, de 2 de julho.
- Canhos, D. A. L.; Canhos, V. P. & Souza, S. 2006. Coleções biológicas e sistemas de informação. *In*: MCT. Diretrizes e estratégias para a modernização de coleções biológicas brasileiras e a consolidação de sistemas integrados de informação sobre biodiversidade. Centro de Gestão de Estudos Estratégicos (CGEE)/Ministério de Ciência e Tecnologia (MCT), Brasília. Pp. 241-314.
- Canhos, D. A. L.; Chapman, A. & Canhos, V. P. 2004. Study on data-sharing with countries of origin - Commissioned study by GBIF (OCB Work Programme - contract no. GBIFS/2003/04). Global Biodiversity Information Facility, Copenhagen. 41p.
- Canhos, D. A. L.; Souza, S.; Giovanni, R.; Marino, A.; Siqueira, M. F.; Fonseca, R. L.; Cruz, B. A. & Canhos, V. P. 2005. Estudo de caso: Sistemas de informação on-line - a experiência do CRIA. Disponível em http://www.cria.org.br/cgee/documentos/ntcria_vs5.doc. Acessado em 20 aug 2008.
- Canhos, V. P.; Souza, S.; Giovanni, R. & Canhos, D. A. L. 2004. Global Biodiversity informatics: setting the scene for a "new world" of ecological modeling. *Biodiversity Informatics* 1: 1-13.
- Forzza, R. C.; Mynssen, C. M.; Tamaio, N.; Barros, C. F.; Franco, L. & Pereira, M. C. A. 2008. As coleções do Herbário. *In*: JBRJ. Jardim Botânico do Rio de Janeiro 1808-2008. Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, Rio de Janeiro. Pp. 45-56.
- Gaikwad, J. & Chavan, V. 2006. Open Access and biodiversity conservation: challenges and potentials for the developing world. *Data Science Journal* 5: 1-17.
- Gonzalez, M. 2007. Distribuição geográfica conhecida e potencial de *Hennecartia omphalandra* Poisson e *Macropeplus ligustrinus* (Tul.) Perkins (Monimiaceae). Dissertação de Mestrado. Escola Nacional de Botânica Tropical, Rio de Janeiro. 109p.
- Gonzalez, M.; Peixoto, F. L.; Silva, L. A. E. & Rocha, M. S. 2003. Proposta de metodologia para informatização do herbário RB, Instituto de Pesquisas Jardim Botânico do Rio de Janeiro (Resumo). Anais do 54º Congresso Nacional de Botânica, Belém.
- IBAMA. 2007. Instrução Normativa nº 160, de 27 de abril de 2007. Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis/Ministério do Meio Ambiente.
- JBRJ. 2003. Processo nº 02011.001085/2003-30: Informatização do Acervo do Herbário do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro.
- Lane, M. A.; Edwards, J. L. & Nielsen, E. S. 2000. Biodiversity Informatics: the challenge of rapid development, large databases, and complex data (keynote). *Proceedings of 26th International Conference on Very Large Data Bases* Morgan Kaufmann, Cairo. Pp. 729-732.
- Magalhães, C.; Santos, J. C. & Salem, J. I. 2001. Automação de coleções biológicas e informações sobre a biodiversidade da Amazônia. *Parcerias Estratégicas* 12: 294-312.
- Martinelli, G.; Vieira, C. M.; Gonzalez, M.; Leitman, P.; Piratininga, A.; Costa, A. F. & Forzza, R. C. 2008. Bromeliaceae da Mata Atlântica brasileira: lista de espécies, distribuição e conservação. *Rodriguésia* 59(1): 209-258.

- Meyer, F. S. 2008. O gênero *Tibouchina* Aubl. (Melastomataceae) no estado do Paraná, Brasil. Programa de Pós-graduação em Botânica/Setor de Ciências Biológicas: 114. Universidade Federal do Paraná, Curitiba.
- Oliveira, R. B. & Godoy, S. A. P. 2007. Composição florística dos afloramentos rochosos do Morro do Forno, Altinópolis, São Paulo. *Biota Neotropica* 7(2): 37-48.
- Peixoto, A. L. & Morim, M. P. 2003. Coleções botânicas: documentação da biodiversidade brasileira. *Ciência & Cultura* 55(3): 21-24.
- Peixoto, A. L.; Barbosa, M. R. V.; Menezes, M. & Maia, L. C. 2006. Diretrizes e estratégias para a modernização de coleções botânicas brasileiras tendo como foco a formação de taxonomistas e a consolidação de sistemas integrados de informação sobre biodiversidade. Disponível em <http://www.cria.org.br/cgee/documentos/botanica.doc>. Acessado em 14 de março de 2009.
- Peixoto, F. L. 2005. O processo de informatização de herbários: estudo de caso. *Escola Nacional de Botânica Tropical*: 79. Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, Rio de Janeiro.
- Smith, G. F.; Steenkamp, Y.; Klopper, R. R.; Siebert, S. J. & Arnold, T. H. 2003. The price of collecting life. *Nature* 422: 375-376.
- Suarez, A. V. & Tsutsui, N. D. 2004. The value of museum collections for research and society. *BioScience* 54(1): 66-74.
- Vargas, B. C.; Nakajima, J. N. & Romero, R. 2007. Acervo digital do Herbarium Uberlandense: Parque Nacional da Serra da Canastra, MG. *Horizonte Científico* 1(7): 14.