

# 61º CONGRESSO NACIONAL DE BOTÂNICA

Manaus, 5-10/09/2010

Palestra proferida em Simpósio promovido pelo  
INCT-Herbário Virtual da Flora e dos Fungos

## INTEGRAÇÃO DINÂMICA DE DADOS SOBRE A FLORA DO BRASIL

Dora Ann Lange CANHOS<sup>1</sup>, Alexandre MARINO<sup>1</sup>, Diego ROCHA<sup>1</sup>, Sidnei de SOUZA<sup>1</sup>,  
Vanderlei Perez CANHOS<sup>1</sup>

**Resumo (máximo 200 palavras)** – O trabalho apresenta alguns resultados do projeto Herbário Virtual de Plantas e Fungos do Brasil, um dos Institutos Nacionais de Ciência e Tecnologia apoiados pelo CNPq. É destacada a base informacional de acesso público e as ferramentas desenvolvidas para facilitar o acesso e visualização dos dados e para dar elementos para análises, tanto dos curadores, como dos especialistas.

**Palavras chaves** – Biodiversidade, Sistema de Informação, Internet.

A disseminação pública de dados e informações botânicas no Brasil deu um grande salto nos últimos cinco anos. Um marco nesse contexto de crescimento foi o trabalho realizado em 2005-2006 que gerou o documento *Diretrizes e estratégias para a modernização de coleções biológicas brasileiras e a consolidação de sistemas integrados de informação sobre biodiversidade* (Egler *et al.* 2006). Esse trabalho reuniu as sociedades brasileiras de botânica, de zoologia e de microbiologia e o CRIA – Centro de Referência em Informação Ambiental que, coordenados pelo Centro de Gestão e Estudos estratégicos, apresentaram um planejamento de dez anos para coleções biológicas e sistemas de informação associados. Este documento foi apresentado à comunidade científica e apresentado e ratificado pelo governo brasileiro durante a Conferência das Partes da Convenção sobre Diversidade Biológica (COP-8) realizado em Curitiba. Foi também incluído pelo Ministério da Ciência e Tecnologia em seu plano de ação para 2007-2010. A CONABIO após incorporar sugestões ao documento o aprovou, em 2008, tornando-o um plano de ação do governo brasileiro para as coleções biológicas e a taxonomia (Marinoni & Peixoto, 2010).

Uma das metas gerais desse documento envolvendo as três áreas de conhecimento – Botânica, Zoologia e Microbiologia – foi a *consolidação de um sistema compartilhado de*

---

<sup>1</sup> Centro de Referência em Informação Ambiental, Av. Romeu Tórtima 388, Barão Geraldo, 13084-791, Campinas, São Paulo, dora@cria.org.br

*dados e informações sobre espécies e espécimes.* As diretrizes estabelecidas serviram de base para o plano estratégico do CRIA referente ao período 2007-2010.

Embora o país não tenha ainda implementado um grande projeto para atender as necessidades identificadas - desde a formação e valorização dos recursos humanos à infraestrutura das coleções, o apoio à digitação e digitalização dos dados e aos sistemas de informação *on-line* de interesse público - o diálogo aberto entre os diferentes atores e o estabelecimento de diretrizes e metas comuns resultaram nos avanços que agora se constata.

Uma iniciativa que merece destaque para a botânica é o projeto Herbário Virtual de Plantas e Fungos do Brasil, um dos Institutos Nacionais de Ciência e Tecnologia apoiados pelo CNPq, que adotou a rede *speciesLink* como a base do seu sistema informacional.

A proposta apresentada ao CNPq envolveu 25 instituições, a Universidade Federal de Pernambuco como coordenadora, 22 herbários (aqui incluindo também as coleções associadas aos herbários, como xilotecas, bancos de DNA e outras), o CRIA e a RNP (Rede Nacional de Ensino e Pesquisa). As instituições com herbários, no início dos trabalhos possuíam um acervo de 1.678.674 exemplares com 893.460 registros disponíveis *on-line* na rede *speciesLink*. Ainda como *marco zero* do projeto, participavam da rede *speciesLink* mais 18 herbários do país e dois do exterior, que repatriavam os dados dos exemplares coletados no Brasil. Somando os dados, o marco zero do Herbário Virtual era um acervo estimado em 2.578.086 exemplares, disponibilizando 1.448.274 registros *on-line*. Como o número estimado de exemplares depositadas em herbários brasileiros é de 6 milhões (Peixoto *et al.* 2009), significa que o INCT Herbário Virtual teve no seu início cerca de 43% do acervo nacional direta ou indiretamente envolvido. Iniciamos o projeto com um índice de georreferenciamento de 34%, representando cerca de 497 mil registros *on-line*.

Com relação aos dados repatriados, iniciamos o projeto com 382.148 registros repatriados dos Jardins Botânicos de Nova Iorque e Missouri, sendo 135.628 (35%) georreferenciados.

A tabela a seguir mostra a evolução do conteúdo da rede, comparando o marco zero com dados do dia 23 de julho de 2010.

Item	Marco Zero					Julho 2010				
	Acervo	On-line	%	Georref	%	Acervo	On-line	%	Georref	%
<b>Colaboradores e Associados INCT</b>	1.678.674	893.460	53%	442.159	49%	2.830.661	1.423.467	50%	1.171.162	82%
<b>Coleções não associadas</b>	899.412	554.814	62%	82.720	15%	1.041.658	681.644	65%	546.208	80%
<b>SubTotal</b>	2.578.086	1.448.274	56%	524.879	36%	3.872.319	2.105.111	54%	1.717.370	82%

<b>Repatriamento</b>	382.148	382.148	100%	151.966	40%	463.663	463.663	100%	243.128	52%
<b>Total</b>	<b>2.960.234</b>	<b>1.830.422</b>	<b>62%</b>	<b>676.845</b>	<b>37%</b>	<b>4.335.982</b>	<b>2.568.774</b>	<b>59%</b>	<b>1.960.498</b>	<b>76%</b>

O acervo total do INCT Herbário Virtual aumentou 46% e os dados *on-line* 40%, sendo que para os registros georreferenciados, somando os dados georreferenciados na origem com aqueles georreferenciados por aplicativo, o aumento foi de 190%.

A oportunidade de desenvolver trabalhos em parceria com os herbários, além do aumento na oferta de dados de acesso público, deu início a um processo de uso e análise das coleções, visando melhorar a qualidade dos dados disponíveis *on-line*.

Uma nova interface de busca foi desenvolvida visando dar maior aplicabilidade aos dados e, mais especificamente, oferecendo as ferramentas ao Comitê Gestor do INCT para recuperar da rede as exsicatas identificadas somente até família para orientar o seu programa de visita de especialistas. Acessando a página de busca do INCT ([inct.splink.org.br](http://inct.splink.org.br)) e optando pela busca avançada pode-se realizar a busca seguindo o seguinte critério: *família diferente de branco e gênero igual a (fonético) sp ou gênero igual a branco* (figura 1).

**Figura 1. Interface de busca avançada do Herbário Virtual da Flora e dos Fungos**

Obtém-se, desta busca, cerca de 83 mil registros que podem ser apresentados de diferentes formas. Os exemplos a seguir são as respostas, com os registros apresentados por coleção, por família e por estado (figuras 2 a 4).

Os registros são apresentados por coleção. Escolha o tipo de conteúdo, formato e saída desejada textual, mapas e Google maps

Coleção	Georeferenciados			Conteúdo	Formato	Saída
	Registros	Original	Automático			
ASE	1344	129	0	Resumido	HTML	ver
EAC	751	96	514	Completo	MS-Excel	ver
EAN	480	5	368	Resumido	XML	ver
FUEL	5009	289	4352	Resumido	HTML	ver
FURB	1625	1403	202	Resumido	HTML	ver
HST	1290	44	0	Resumido	HTML	ver
HTSA	246	51	116	Resumido	HTML	ver
HUEFS	8767	6575	0	Resumido	HTML	ver
HVASF	48	0	0	Resumido	HTML	ver

Figura 2. Resultados apresentados por coleção

Informações **Inventário**

**Coleções:**  
 ASE, EAC, EAN, FUEL, FURB, HST, HTSA, HUEFS, HVASF, INPA-Fungos, INPA-Herbario, IPA, JBRJ\_RB, JBRJ\_RBdna, JBRJ\_Rbw, JPB, MAC, MOSS, R, SP, SP-Algae, SPF, SPF-Algae, SPFw, TEPB, UB, UFP, UFP-Carpoteca, UFRN, UFRN-Fungos, URM

família	total
	10085
#	1
?	21
ACANTHACEAE	997
ACERACEAE	1
Achariaceae	1
ACHATOCARPACEAE	2
ACORACEAE	1
ACTINDIACEAE	8
ACTINIDACEAE	1
ADELANTHACEAE	2
Adiantaceae	25
Agaricaceae	433
Agaricales	3
Agavaceae	11
AIZOACEAE	18
ALISMATAACEAE	67
ALLIACEAE	9
ALOACEAE	1
Alstroemeriaceae	131
ALSTROEMIRIACEAE	1
ALSTROMERACEAE	1
AMANITACEAE	1

Figura 3. Inventário por família

INCT-Herbário Virtual da Flora e dos Fungos  
english | ajuda

FEDERAL DE SANTA CATARINA  
IPA  
USP  
Instituto de Botânica  
FURB  
UFPR  
UNICAMP  
UFPA  
UFERSA

Informações **Inventário**

**Coleções:**  
ASE, EAC, EAN, FUEL, FURB, HST, HTSA, HUEFS, HVASF, INPA-Fungos, INPA-Herbario, IPA, JBRJ\_RB, JBRJ\_RBdna, JBRJ\_RBw, JPB, MAC, MOSS, R, SP, SP-Algae, SPF, SPF-Algae, SPFw, TEPB, UB, UFP, UFP-Carpoteca, UFRN, UFRN-Fungos, URM

estado	total
	7627
311	1
?	8
AC	16
Acre	1594
Argentina	10
AL	15
Alagoas	4903
Alaska	5
AM	96
Amapá	31
Amapá	210
Amarante	1
Amazonas	8656
Angra dos Reis	1
Antarctica	13
Antioquia	24
Antártica Chilena	1
Aracaju	4
Araioses	1
Arizona	4
BA	178
Bahia	9788

**Figura 4. Inventário por estado**

Além de contribuir para a definição de uma estratégia de visita de especialistas aos herbários, essas ferramentas são importantes para identificar problemas com registro de dados em campos errados, a falta de padronização e erros de digitação, normalmente identificados nos relatórios de *data cleaning* disponível a cada coleção, além do uso de caracteres como “?” ou expressões como “não sei” que dificultam a análise preliminar via aplicativo. Esse aplicativo facilita e dá elementos para análises, tanto dos curadores, como dos especialistas.

O INCT Herbario Virtual vem investindo em duas frentes de trabalho em relação aos dados: aumento do número de registros de acesso livre e aberto disponível *on-line* e melhora na qualidade e completude dos dados.

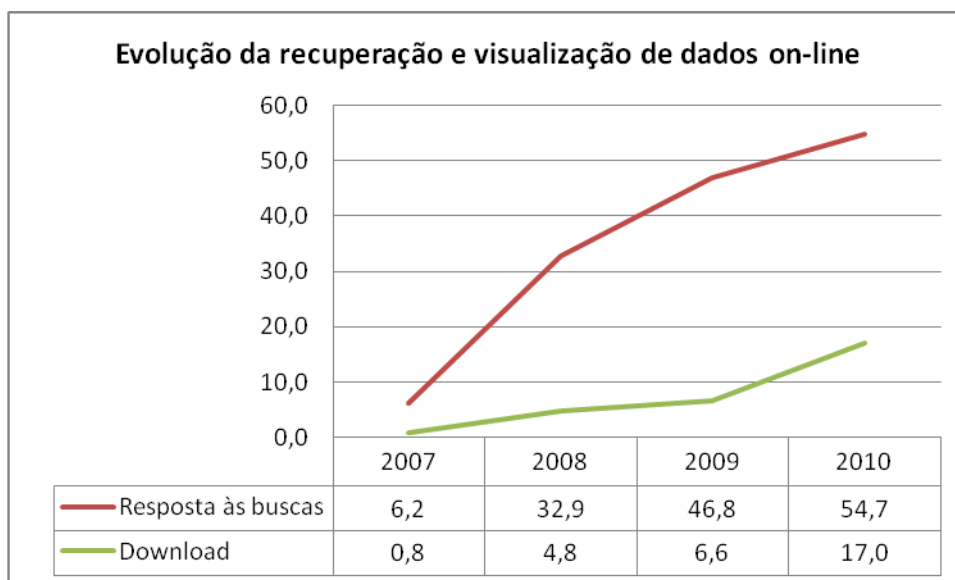
Uma importante ferramenta que deverá ser trabalhada em breve inclui o uso da Lista de Espécies da Flora do Brasil (Forzza *et al.* 2010). Serão desenvolvidos aplicativos para elaboração de relatórios por coleção, indicando a lista de nomes válidos, sinônimos e de nomes que não constam da Lista do Brasil. Com a análise desses dados os curadores podem tornar ainda mais dinâmicos os contatos com os especialistas em grupos taxonômicos. Deverão também ser desenvolvidas interfaces que possam indicar omissões na lista de ocorrência geográfica de espécies e de lacunas geográficas de dados nas coleções.

Com o aumento da quantidade e da qualidade dos dados disponíveis na Internet de forma livre e aberta, um resultado esperado é o aumento do uso dos dados. Desde 2007 o CRIA tem mantido um *log* do acesso à rede *speciesLink*. A tabela a seguir apresenta o número total dos dados de herbários *on-line* de 2007 a 2009 (média dos dados em dezembro de cada ano) e do dia 5 de julho de 2010. Acrescentamos à tabela dois dados importantes: (1) o número de registros recuperados em resposta às buscas e (2) o número de registros visualizados após a apresentação dos resultados. A coluna evolução sempre compara um ano com o seu antecessor.

Ano	2007	2008	Evolução	2009	Evolução	05/07/2010	Evolução
<b>Total On-line</b>	1.344.756	1.927.980	43%	2.272.152	18%	2.390.318	5%
<b>Resposta às buscas</b>	8.348.854	63.379.662	659%	106.409.823	68%	130.714.575	23%
<b>Download</b>	1.124.877	9.330.497	729%	14.901.505	60%	40.751.179	173%

Comparando julho de 2010 com dezembro de 2007 temos que o número total de registros *on-line* cresceu cerca de 78%, enquanto o número de registros que atendem ao critério de buscas do usuário cresceu 1.466% e o número de registros recuperados para análise 3.523%, observando que o acesso feito através do CRIA não é contabilizado.

Pode-se visualizar esses números dividindo o número de registros recuperados na busca e visualizados pelos usuários pelo número total de registros disponíveis, o que dá uma idéia de quantos registros são recuperados e vistos em relação ao acervo disponível (figura 5).



**Figura 5. Número relativo de registros recuperados e visualizados on-line**

Para entender melhor os números apresentados na figura 5, no primeiro semestre de 2010, mais de 130 milhões de registros atenderam os critérios de busca on-line e mais de 40 milhões de registros foram realmente acessados pelos usuários. Esse número, 40 milhões de registros “vistos” representa 17 vezes o acervo total disponível on-line. Esses números

comprovam a importância e o resultado significativo que essa iniciativa vem obtendo graças ao trabalho desenvolvido de forma cooperativa e em rede.

### **Referências Bibliográficas**

Egler, I., Santos; M.M., Canhos, V.P.; Peixoto, A.L.; Barbosa, M.R.V.; Menezes, M.; Maia, L.C.; Vazoleler, R.F.; Marinoni, L.; Canhos, D.A.L., 2006. Diretrizes e estratégias para a modernização de coleções biológicas brasileiras e a consolidação de sistemas integrados de informação sobre biodiversidade. Brasília, Ministério da Ciência e Tecnologia (MCT) & Centro de Gestão e Estudos Estratégicos (CGEE). 42p.

Disponível no endereço <http://www.mct.gov.br/index.php/content/view/66226.html>

Marinoni, L. & Peixoto, A.L. 2010. As coleções biológicas como fonte dinâmica e permanente de conhecimento sobre a biodiversidade. *Ciência & Cultura* 62(3): 54-57. (Deliberação em <http://www.mma.gov.br/estruturas/conabio/arquivos/deliberacao.pdf>)

Peixoto, A.L., Barbosa, M.R.V., Canhos, D.A.L. & Maia, L.C. 2009. Coleções Botânicas: Objetos e Dados Para a Ciência. In: Granato, M. & Rangel, M. (Orgs). *Cultura material e patrimônio da Ciência e Tecnologia*. Museu da Astronomia e Ciências Afins. Rio de Janeiro. (<http://inct.florabrasil.net/downloads/cole%C3%A7%C3%B5es-botanicas-capitulo-2009.pdf>)

Forzza, R.C.; Leitman, P.M.; Costa, A.F.; Carvalho Jr., A.A.; Peixoto, A.L.; Walter, B.M.T.; Bicudo, C.; Zappi, D.; Costa, D.P.; Lleras, E.; Martinelli, G.; Lima, H.C.; Prado, J.; Stehmann, J.R.; Baumgratz, J.F.A.; Pirani, J.R.; Sylvestre, L.; Maia, L.C.; Lohmann, L.G.; Queiroz, L.P.; Silveira, M.; Coelho, M.N.; Mamede, M.C.; Bastos, M.N.C.; Morim, M.P.; Barbosa, M.R.; Menezes, M.; Hopkins, M.; Secco, R.; Cavalcanti, T.B.; Souza, V.C. 2010. Introdução. in *Lista de Espécies da Flora do Brasil*. Jardim Botânico do Rio de Janeiro. Disponível em <http://floradobrasil.jbrj.gov.br/2010>.