

**Ministério do Meio Ambiente – MMA**  
**Secretaria de Biodiversidade e Florestas - SBF**  
**Departamento de Conservação da Biodiversidade - DCBio**  
SEPN 505 – Av W2N – Bl B – Ed. Marie Prendi Cruz

**GRUPO DE TRABALHO**  
**INTEGRAÇÃO DE SISTEMAS DE INFORMAÇÃO E BANCOS DE DADOS SOBRE**  
**BIODIVERSIDADE NO ÂMBITO DO MMA**

Brasília, DF - 22 a 24 de novembro de 2011

**GRUPO DE TRABALHO**  
**INTEGRAÇÃO DE SISTEMAS DE INFORMAÇÃO E BANCOS DE DADOS SOBRE**  
**BIODIVERSIDADE NO ÂMBITO DO MMA**

**AGENDA**

**Dia 1 - 22 11 2011**

**08:30 – Café da manhã de boas vindas**

**09:00 – Abertura**

Integração de dados para gestão da biodiversidade

Histórico

- Principais demandas de informação para gestão
- Objetivos, expectativas e estrutura da Oficina de Trabalho

**09:50 – Intervalo**

**10:10 – Apresentações (20min) e perguntas**

- Repositórios de dados (Flávia Costa – INPA)
- Proposta de campos mínimos pelo GTSIB (Keila Juarez – MMA)
- Padrões internacionais na estruturação de dados em Biodiversidade (Rafael Fonseca – GIZ)
- Metadados para o gerenciamento de dados ecológicos (Debora Drucker – Embrapa)

**12:30 – Almoço**

**14:00 – Discussão em grupos**

**Tema 1: Diretrizes para estruturação de bancos de dados em biodiversidade**

**Questões a serem respondidas:**

1. Quais temas mais relevantes para a gestão da biodiversidade?
2. Quais os dados mínimos (em biodiversidade) para a integração de sistemas diversos? (deverão ser observadas as “*informações mínimas necessárias para a integração de bancos de dados de espécies*”, abaixo, de acordo com o Relatório do seminário de 2010).
  - Registro único para os espécimes coletados e locais de depósito;
  - Nome científico;
  - Localidade (coordenadas geográficas, altitude/profundidade, UF, bioma, sítio de coleta, tipo de vegetação, habitat);
  - Data e hora;
  - Equipe (responsável pela expedição, coletor, assistente); método de coleta, captura ou marcação; esforço de coleta; determinador da espécie (pessoa que identificou o material);
  - Cada registro deve conter atributos do sítio de coleta;
  - O sistema deve possuir campos estruturados e relacionar o nome científico com toda a hierarquia taxonômica.
3. Como estes dados devem estar descritos e disponíveis?
4. Deverão ser definidos modelos de tabelas (planilhas) de dados? Em caso positivo, estes modelos deverão ser apresentados.

**Produtos esperados:**

1. Documento com diretrizes para estruturação de bancos de dados em biodiversidade.
2. Subsídios para a execução dos produtos dos consultores referentes à avaliação dos sistemas no âmbito do MMA e vinculadas:

**18:00 – Encerramento**

## **Dia 2 - 23 11 2011**

**08:30 – Café da manhã**

**08:45 – Apresentações (20min) e perguntas**

- Estruturação de dados da biodiversidade obtidos no processo de licenciamento ambiental de empreendimentos junto à Petrobrás, e mapeamento em ambiente SIG. Iniciativa Petrobrás (Ricardo Araújo Machado - CGPEG/DILIC/IBAMA)
- Estrutura de dados em Biodiversidade na UFSCar (Sérgio Morbiolo – UFSCar)
- IABIN/PTN: Arquitetura de Sistemas para Integração de Dados de Polinizadores das Américas (Pedro Luiz Pizzigatti Corrêa – Escola Politécnica USP, Projeto Polinizadores)
- Sistemas de Informação do JBRJ e integração (Eduardo Dalcin – JBRJ)

**11:15 – Intervalo**

**11:30 – Continuação da discussão em grupos**

**12:30 – Almoço**

**14:00 – Apresentação resultados dos grupos – Tema 1**

**14:30 – Discussão plenária e consolidação de encaminhamentos**

**15:30 – Intervalo**

**15:45 – Discussão em grupos**

**Tema 2: Metadados: Propostas de encaminhamentos para adoção de metodologias pelas instituições vinculadas do MMA.**

- Definição de metodologias (ferramentas);
- Avaliação da capacidade das instituições na implementação das ferramentas sugeridas;
- Sugestões para capacitação.

## **Dia 3 - 24 11 2011**

**08:30 – Café da manhã**

**08:45 – Apresentação resultados dos grupos – Tema 2**

**09:15 – Discussão plenária e consolidação de encaminhamentos**

**10:15 – Intervalo**

**10:30 – Discussão em grupos – Tema**

**Tema 3: Padrões e protocolos de comunicação**

- Quais recomendações para adoção de padrões e protocolos de interoperabilidade?
- É necessário o estabelecimento de um protocolo único?
- Quais requisitos mínimos para integração de dados no âmbito do MMA?

**12:30 – Almoço**

**14:00 – Apresentação resultados dos grupos**

**15:00 – Discussão plenária e consolidação de encaminhamentos**

**15:30 - Encerramento**

# **GRUPO DE TRABALHO INTEGRAÇÃO DE SISTEMAS DE INFORMAÇÃO E BANCOS DE DADOS SOBRE BIODIVERSIDADE NO ÂMBITO DO MMA**

## **RELATORIA - APRESENTAÇÕES**

**Dia 1 - 22 11 2011**

### **Apresentação dos participantes.**

#### **Abertura da Oficina (Keila Juarez - MMA)**

Objetivos da Iniciativa  
Histórico  
Objetivos da Atividade  
Programação

#### **Apresentação – O que é necessário para um repositório de dados funcionar? (Flávia Costa – INPA)**

Ao invés de apresentar a importância, discutiu os caminhos factíveis para sua execução.

Empty archives – artigo de Bryn Nelson discutindo porque os artigos continuam vazios: sistemas complicados demais e dificuldade de acesso pelos usuários; conclusão: “precisamos de coisas fáceis”. (<http://www.nature.com/news/2009/090909/full/461160a.html>)

Matt Jones (NCEAS) – afirma ser praticamente impossível ter um único banco de dados para tudo; cada tema, área tem características muito específicas... Mas é possível ter um repositório para tudo, tendo um repositório global é possível atender a todos os usuários.

A palavra chave é flexibilidade (acomodar coisas imprevistas) sistemas perfeitos para uma situação provavelmente não serão bons no futuro! Não só flexibilidade computacional, ou de interoperabilidade (preocupação natural dos engenheiros de sistemas), mas para os usuários, dos campos (atributo dos registros), etc. O sistema deve estar apto a agregar novos campos no futuro (mesmo com campos mínimos obrigatórios desde o início). O sistema deve permitir que os nomes dos campos variem (desde que a idéia original possa sempre ser resgatada).

Sistemas muito complexos normalmente não têm sucesso, é necessário ter um repositório de dados global, com metadados bem estruturados. Estes precisam incluir mais do que apenas a descrição dos atributos dos registros; mas também, p.ex.: equipe envolvida, localização espacial e temporal da amostra, descrição dos métodos de amostragem e esforço de coleta (cruciais para a escolha das tabelas, linhas que serão utilizadas em determinado estudo), hierarquia taxonômica.

EML (Ecological Metadata Language) documentação básica de metadados em estudos de monitoramento. Como está sendo implementado no PELD hoje é bastante flexível, às vezes tabelas complementares se fazem necessárias (p.ex. para a documentação das unidades amostrais: área, forma, localização...).

Os campos da tabela de metadados precisam ser padronizados, permitindo buscas rápidas.

As tabelas de dados que entram nos repositórios, em princípio podem ter qualquer formato. Entretanto, se desejarmos que elas componham um banco de dados, estas precisam pelo menos ser compatíveis com um modelo relacional (registros individuais a cada linha). Os campos da tabela precisam minimamente ser padronizados (facilitando a integração de informações de diferentes localidades ou diferentes grupos taxonômicos), lembrando sempre da flexibilidade.

A preocupação com validação após a entrada dos dados é importante e justificada, entretanto é importante incluir nas tabelas um campo que indique o grau de confiança nos dados (p. ex. observação adicionada de foto, de vocalização, de coleta de indivíduo/amostra; conferindo confiança à amostra).

O formato de entrada de dados no repositório não pode ser de um registro por vez (p.ex. um formulário por registro!), mas por tabelas (p. ex. em formato ascii, que permite a intercomunicação com praticamente qualquer sistema).

Tabelas de entrada de dados devem ser individualizadas por temas (tabela tempo, tabela dados bióticos, tabela pessoas – todas ligadas por uma coluna comum, p.ex. amostras). Tabelas de campo não são tabelas de repositórios.

Gerenciamento do sistema – é imprescindível que haja pelo menos um gerente de dados em cada instituição envolvida em regime de dedicação exclusiva (buscando informações para entrada, fazendo a primeira validação, treinando usuários). Devem-se assegurar recursos para manutenção deste elemento chave.

É necessário um fortíssimo esforço de capacitação em todos os níveis para que haja entendimento de como fazer metadados.

Estruturação da ligação **repositório – banco de dados – resposta aos gestores:**

- repositório nacional/regional que recebe dados/metadados em forma de tabelas ascii (campos mínimos definidos pelo MMA e chave primária adequada);
- cada tabela vinculada a metadados estruturados (se necessário acompanhados de tabelas complementares);
- tabelas adequadas para compor cada banco de dados específico são selecionadas a partir dos metadados;
- bancos específicos podem ser consultados para atender às perguntas dos usuários (ex. repositórios do PELD – há uma interface para busca de dados por palavra de interesse).

**Perguntas/comentários:**

Rodrigo Koblitz (DILIC) - Na discussão sobre campos mínimos no trabalho da DILIC, não encontraram nenhum atributo que fosse comum a todos, sugere um dicionário comum (que identifique o que é o campo e a que tipo de informação ele se refere).

Flávia Costa (INPA) – o fato de o nome variar é bastante provável, sendo necessário haver uma ontologia (o dicionário) que permita saber exatamente o que está realmente sendo referido por um dado.

Pedro Pizzigatti (Poli USP) – cita referência bibliográfica sobre desenvolvimento de dicionário (pode ser utilizado como start, será disponibilizada ao grupo)

**Apresentação - GTSIB** (Keila Juarez - MMA)

Grupo de Trabalho Sistema de Integração de Informação de Biodiversidade.

Justificativa: Percepção da necessidade de integração de informações para responder às atribuições dos órgãos ICMBIO e IBAMA (p.ex. definição de áreas prioritárias, mapas de sensibilidade, zoneamento, etc).

Objetivos: definir diretrizes para entrada de novos dados e criação de novos sistemas.

Descrição dos levantamentos realizados e resultados dos diagnósticos dos sistemas

- Identificados 34 sistemas/bancos: implantados, em implantação ou previstos (em 2008);
- Objetivos variados (muitas vezes duplicando esforços);
- Diagnóstico mínimo (tipo de banco e linguagem de programação utilizada, forma de acesso, local de hospedagem, integração com outros sistemas, lista taxonômica utilizada e sistema de geoprocessamento dos dados)

Proposição de arranjo para um sistema integrador: cadastro, locais de registro, táxons registrados e associação entre os campos.

Vantagens: possibilidade de estimativas de riqueza e, em alguns casos, de abundância; disponibilização de dados ambiental organizados; construção participativa; atualização constante; e busca refinada de informações.

Desafios: definição de instituição responsável pelo sistema; gerenciamento do sistema integrador; compilação de dados pretéritos; padronização de dados futuros; possíveis readequações; garantia orçamentária.

**Perguntas/comentários:**

Joseílson Costa (ICMBio)/Flávia Costa (INPA) - Discussão quanto à manutenção de um banco de dados que atenda a um grupo de instituições que tenham uma pergunta, ou a manutenção de repositórios que receba os bancos destas instituições. Vantagens e desvantagens.

Rodrigo Koblitz (DILIC) – histórico da negociação da DILIC com os empreendedores para estruturação de bancos de dados que tenham um dicionário com atributos mínimos.

Flávia Costa (INPA) – importância em se prever a sobrevivência do sistema, física e computacional.

**Apresentação – Importância da adoção de padrões para estruturação de dados em biodiversidade** (Rafael Fonseca - GIZ)

Primeira questão: “Ao que nos referimos com: integração de sistemas...?”

- No âmbito do MMA;

- Voltada à gestão da biodiversidade;
- Em escala nacional.

Dados em biodiversidade – qualquer elemento identificado em sua forma bruta que, por si só, não conduz a uma compreensão de determinado fator.

Informação – dado trabalhado, conjunto de dados conjugados que possuem relevância e propósito.

Conhecimento de biodiversidade – conjunto organizado e estruturado de informações a respeito do mundo que permite simulações. Além dos dados organizados requer um conhecimento associado.

Estrutura de ação (apresentação de fluxograma: da coleta à estratégia)

Conhecimento em biodiversidade – o que temos de bases de dados brutos? Como estão estruturadas? Na maioria das instituições as bases não estão estruturadas, qual é a estratégia de ação?

Dados primários talvez possa ser o investimento possível agora, mesmo não sendo o ideal. Pode-se pensar numa estratégia modular, que seja desenvolvida em longo prazo.

Sistemas distribuídos X Sistemas centralizados.

Integração de dados e características específicas (estruturação dos dados na base).

Sistemas centralizados: segurança, homogeneidade da base, fácil acesso, estrutura única.

Sistemas distribuídos: busca de dados a partir de consulta aos repositórios em cada instituição.

Atualização e garantia do dado é responsabilidade do portador. Tradutor garante comunicação entre sistemas com linguagens distintas (técnicas e de software)

Como integrar?

Modelos/estrutura de dados - adaptação das bases aos campos mínimos do tradutor (modelo de dados);

Protocolos de Comunicação – garantir a organização de forma que sejam entendidos na recepção;

Metadados – auxiliam nas buscas e integração dos sistemas;

Darwin Core – tem como objetivo facilitar estruturação e intercâmbio de dados de ocorrência geográfica de espécies (especialmente ligando as informações à taxonomia);

ABCD – Access to Biological Collection Data: permite buscas unificadas, suporte internacional. Desvantagem necessidade de certa reestruturação dos bancos de dados existentes;

Protocolos de interoperabilidade – tradutor que permite que cada base de dados adote o seu;

DIGIR - Distributed Generic Information Retrieval – utiliza padrões e protocolos abertos, é muito difundido no Brasil e no mundo;

TAPIR - Access Protocol for Information Retrieval – independência da estrutura lógica do banco.

Para pensar:

- Diretrizes para modelo de dados. Precisamos de um único?
- O que fazer com os bancos rígidos já desenvolvidos, ou em desenvolvimento?
- Sugestão de modelo saída de dados para sistemas já desenvolvidos?
- Adoção de protocolo de integração de dados?
- Vantagens e desvantagens entre sistemas distribuídos e centralizados;
- Custo e viabilidade – complexidade, capacitação, manutenção (física e técnica).

### **Perguntas/considerações:**

Joseílson Costa (ICMBio) – reafirma sua impressão de que para algumas instâncias no âmbito do MMA talvez haja possibilidade dum sistema único.

Rafael Fonseca (GIZ) – necessidade de se focar na realidade das instituições ao pensar em qual caminho seguir.

Flavia Costa (INPA) – os dados relevantes de monitoramento apresentam, normalmente, dinâmica temporal (diferentemente dos registros de ocorrência, ou dados de coleção), são complexos e necessitam que a descrição desta dinâmica esteja associada aos dados.

Eduardo Dalcin (JBRJ) – analogia para as necessidades do MMA.

Produtor X consumidor: o MMA encontra-se em um restaurante querendo consumir um prato (informação em biodiversidade para tomar uma decisão) este prato vem de uma cadeia: produção (coleta de dados), distribuição no mercado (entrada em bancos de dados), empacotar no modo de usar (disponibilização), cozinha (knowledge workers preparam a informação para consumo). A lata

do milho é o Darwincore, o abridor é o DIGIR, existem pratos que são consumidos crus, enquanto outros devem ser combinados. O que precisamos construir é um super-mercado, não discutir se o agricultor rega duas vezes o milho, nem o cardápio que é particularidade do cozinheiro. Queremos falar das vitrines para expor essas comidas (dados). Dando dois passos atrás, discutirmos como será o mercado, quais prateleiras, como abrir os produtos. Como processá-los para elaborar pratos já é uma outra discussão.

Rafael Fonseca (GIZ)– definir se queremos ceasa, carrefour, delivery...

### **Apresentação – Metadados para o gerenciamento de dados ecológicos** (Debora Drucker - Embrapa)

Publicação Michener – entropia dos dados, ao longo do tempo após a coleta, os detalhes específicos a respeito da coleta vão sendo perdidos. No caso dos dados ecológicos existe uma grande heterogeneidade em todos aspectos relacionados a estes dados.

([www.esajournals.org/doi/pdf/10.1890/1051-0761\(1997\)007%5B0330:NMFTES%5D2.0.CO%3B2](http://www.esajournals.org/doi/pdf/10.1890/1051-0761(1997)007%5B0330:NMFTES%5D2.0.CO%3B2))

Pirâmide da quantidade de dados sobre um ponto contra escalas mais amplas.

Integração e compartilhamento, quais informações são úteis para que as pessoas entendam os dados: metadados - fornecem informações sobre que entidades, processos ou medidas são representados nos dados. São a sua documentação, permitindo seu encontro, acesso e reutilização. Um padrão de metadados fornece estrutura e consistência para a documentação dos dados, em linguagem que os computadores podem ler. Darwincore é um padrão amplamente adotado para descrever dados de espécimes. Para documentar dados ecológicos, um padrão muito usado é o

EML – Ecological Metadata Language - desenvolvido para conjuntos de observações.

Softwares para gerenciamento de dados ambientais desenvolvidos pelo KNB (Knowledge Network for Biocomplexity):

Metacat - metadata catalog - um repositório de dados e metadados

Morpho – editor de metadados em linguagem EML

Publicação: Leinfelder et al, 2011, Using Semantic Metadata for Discovery and Integration of Heterogeneous Ecological Data (<https://semtools.ecoinformatics.org/repository/docs/pubs/EIM-2011/main.pdf>) – um exemplo de iniciativa de desenvolvimento de ferramentas de web-semântica para consultar bases de metadados. Ferramentas como essas auxiliam a associação de termos de acordo com os significados das entidades medidas. Assim, é possível contornar, por exemplo, o problema de a mesma entidade ser definida por diferentes termos, tornando os resultados das buscas em catálogos de metadados mais eficazes.

### **Perguntas/comentários:**

Disponibilização só no servidor da instituição, ou no servidor da KNB?

Sinonímia, ao que se lastra a taxonomia do banco?

Transparência dos dados no servidor local e no do KNB?

Customização?

Suporte para problemas?

## **Dia 2 – 23 11 2011**

### **Apresentação – Plano de Caracterização Regional da Bacia de Campos – PCR-BC** (Ricardo Machado - DILIC)

Objetivo, dar suporte ao processo de licenciamento ambiental para atividade de exploração de petróleo.

Atender previamente aos requisitos dos TRs para os EIAs na atividade petrolífera.

Subsidiar o processo de licenciamento com informações disponíveis em Bancos de dados georreferenciados com informações sobre biodiversidade das áreas sob estudo.

### **Perguntas/comentários:**

Discussão quanto ao histórico e possibilidades de replicar essa experiência com outros setores de empreendimentos.

### **Apresentação – SITAMAR** (Alexsandro Santos- TAMAR)

Objetivo, garantir a qualidade dos dados coletados.

Estabelecimento de protocolo para coleta de dados: supervisão dos dados coletados por equipe técnica; capacitação; padronização de manual e planilha para coleta dos dados; entrada dos dados imediata à coleta; checagem e validação previa à disponibilização ao público.

Características do sistema: online, software livre; usuário externo não tem acesso aos dados brutos (somente por demanda justificada).

Estabelecimento de uma política de dados e de um manual de padrões.

#### **Apresentação - SISTAXON** (Joseílson Costa - ICMBio)

Objetivo: armazenar, processar e disponibilizar informações de classificação taxonômica, distribuição, biologia, ecologia e conservação dos táxons da biodiversidade brasileira.

Objetivos específicos: disponibilizar catálogo de nomes dos seres vivos (nome científico, comum e sinônímia); prover listas das espécies ameaçadas de extinção e status de conservação; prover informações sobre distribuição, biologia, ecologia e conservação da biodiversidade brasileira; facilitar revisão e atualização da nomenclatura taxonômica; manter acervo bibliográfico digital sobre a biodiversidade nacional; e subsidiar processos de gestão (revisão da lista de espécies ameaçadas, elaboração de políticas públicas, análises de viabilidade populacional, mapeamento e previsão de distribuição populacional de espécies).

#### **Apresentação - SISBIO** (Joseílson Costa - ICMBio)

Sistema de Autorização e informação em biodiversidade.

Proposta original: instrumento que agregasse as normas estabelecidas para autorização de coleta de material biológico para fins científicos, dentro e fora de UCs federais, incluindo a informatização do processo de solicitação.

Processo: consulta interna quanto às necessidades institucionais, consulta pública, definição de marco legal, definição de comitê gestor.

Avanços.

#### **Apresentação - Sistemas de Informação do JBRJ e integração com outros sistemas** (Eduardo Dalcin - JBRJ)

Três frentes: JABOT (organização de informações de coleções científicas), Lista da Flora do Brasil (desenvolvido e hospedado pelo CRIA) e CNCFlora - Lista de Espécies Ameaçadas (desenvolvimento de um catálogo de metadados possibilitando a avaliação do status de conservação e o risco de extinção das espécies). Atualmente estes três sistemas não são interoperáveis.

Cenário institucional futuro: a partir de parcerias com o Museu Nacional de História Natural (Paris) e Kew Gardens (Londres), no projeto REFLORA (Plantas do Brasil: resgate histórico e herbário virtual para o conhecimento e conservação da Flora Brasileira), pretende-se desenvolver um novo sistema, o JABOT 2 (desenvolvimento: COPPE-UFRJ; infraestrutura: LNCC).

JABOT 2 - um núcleo taxonômico que fará a interface entre três sistemas: coleções científicas, espécies ameaçadas, e floras (nacional, estaduais, locais).

Dentro do escopo do mesmo REFLORA está em curso uma reestruturação da infraestrutura computacional do JBRJ, em suas três frentes: processamento, armazenagem e conectividade. Em um projeto coordenado pelo Laboratório Nacional de Computação Científica (LNCC), estamos capacitando a instituição à armazenar, processar e disponibilizar uma grande quantidade de dados sobre a biodiversidade vegetal brasileira.

#### **Apresentação - Estrutura de dados em Biodiversidade na UFSCar** (Sérgio Morbiolo - UFSCAR)

Desenvolvimento em software livre/código aberto, com possibilidade de sua utilização em todos os níveis (implementação, hospedagem, edição/visualização, manutenção do banco, manutenção do sistema).

Níveis diferenciados de utilização por usuário, garantia contra acidentes por erro ou conduta danosa. Sistema de permissões mediante cadastro (gerente, curador, especialista, anônimo, etc).

Validação por especialistas em diferentes níveis: taxonomia, grafia, etc...

Gerenciamento diferenciado para cada banco do sistema, concatenação de informações em banco de referência (possibilita busca em vários bancos ao mesmo tempo no mesmo servidor).

Servidor web intermédia o acesso do usuário (impossibilita acesso direto ao bando de dados pelo usuário remoto, possibilita a conversão de dados em formato conveniente: CSV, HTML).



Entrada de dados por formulário web, ou planilha CSV; condicionada a cadastro de usuário.  
Possibilidades de busca: por dados relacionados; entre bancos de dados; em qualquer campo; por localidade ou polígono georreferenciado.  
Funcionalidades: exportação para planilhas em padrões internacionais, criação de relatórios, impressão de etiquetas.  
Flexibilidade: possibilidade de alteração dos formulários de entrada (restrita aos gerenciadores); possibilidade de configuração dos módulos do formulário de entrada (conforme necessidade específica da coleção, sem alterar a estrutura do banco)

### **Apresentação - IABIN/PTN: Arquitetura de Sistemas para Integração de Dados de Polinizadores das Américas** (Pedro Correa - USP)

Arquitetura de tecnologias de informação/comunicação e padrões utilizados na Rede Temática de Polinizadores da IABIN (Rede de Informação em Biodiversidade Interamericana).

Objetivo: desenvolver rede integrada para compartilhamento dos principais dados sobre polinizadores através de um conjunto padrão de metadados e protocolos de compartilhamento.

Padrões: - Portal rede temática de polinizadores (ferramentas de acesso a dados);  
- Dados distribuídos (web-based data portal, provedores de dados distribuídos);  
- Padrões (DarwinCore - ocorrências; DarwinCore interaction extension - registros de interação planta-polinizador; DublinCore e FGDC (catálogo IABIN) - provedores de dados, conjuntos de dados e referências bibliográficas; MRTG (catálogo IABIN) - esquema tentativa para metadados de recursos multimídia.

Protocolos: TAPIR (DiGIR), webservices.

Portal de dados: adaptações à ferramenta utilizada no GBIF, em especial quanto aos dados de interação.

Provedores de dados:

- Cada provedor define a ferramenta mais adequada;
- Conteúdos mínimos: dados de espécimes, dados de interação;

Ferramentas: Pollinator Data Digitizer (PDD) - sistema simples para o registro de espécimes e observações, planilhas simples para registro em lotes, utiliza padrões de dados para compartilhamento, módulo provedor de dados, multilingual, ferramenta livre.

Base de dados centralizada para Contatos de especialistas em Polinização e Polinizadores.

Ferramenta para georreferenciamento de dados, sua visualização em mapas e cruzamento com outras informações georreferenciadas.

WebServices: fornecimento de metadados para o catálogo do Serviço Geológico governo americano (referências e imagens sobre polinizadores, dados sobre provedores)

### **Apresentação - Tutorial Interoperabilidade** (Pedro Pizzigatti - USP)

### **Apresentação - Establishing the framework for a National Flora Endangered Species Information System** (Eduardo Dalcin - JBRJ)

# GRUPO DE TRABALHO

## INTEGRAÇÃO DE SISTEMAS DE INFORMAÇÃO E BANCOS DE DADOS SOBRE BIODIVERSIDADE NO ÂMBITO DO MMA

### DISCUSSÕES EM GRUPO

#### **TEMA 1 - G1: Diretrizes para estruturação de bancos de dados em biodiversidade**

1. Quais são os *DADOS* mais relevantes para a gestão da biodiversidade?

*Os temas já estão pré-definidos nos acordos e convenções que o país é signatário.*

*Levar em consideração de forma especial os dados que se referem às dinâmicas outras, que não somente a ocorrência pontual de espécies.*

- *Dados de ocorrência;*
- *Dados de abundância;*
- *Dinâmica de vegetação;*
- *Dinâmica de populações;*
- *Uso e cobertura de solo (do espaço, de áreas);*
- *Dados abióticos;*
- *Dados sobre interações;*
- *Etc.*

2. Quais os dados mínimos (em biodiversidade) para a integração de sistemas diversos? (deverão ser observadas as “*informações mínimas necessárias para a integração de bancos de dados de espécies (ou espécimes?)*”, abaixo, de acordo com o Relatório do seminário de 2010).

- Registro único para os espécimes coletados e locais de depósito;
- Nome científico;
- Localidade (coordenadas geográficas, altitude/profundidade, UF, bioma, sítio de coleta, tipo de vegetação, habitat);
- *DATUM SIRGAS 2000 - obrigatório para instituições do governo federal (norma do IBGE)*
- Data e hora;
- Equipe (responsável pela expedição, coletor, assistente); método de coleta, captura ou marcação; esforço de coleta; determinador da espécie (pessoa que identificou o material);
- Cada registro deve conter atributos do sítio de coleta;
- O sistema deve possuir campos estruturados e relacionar o nome científico com toda a hierarquia taxonômica.

*Estes "campos mínimos" estão adequados às espécies, mas não cobrem todos os aspectos das informações relacionadas à Biodiversidade.*

*Só é possível estabelecer campos mínimos para registros de ocorrência de espécimes/espécies. Para dados ecológicos (não específicos) a variabilidade impede a definição de campos mínimos, entretanto, estes devem ser considerados, descritos em repositórios, com metadados associados (conceito de repositório ( Flavia Costa - INPA) é comum ao conceito de catalogo de metadados Eduardo Dalcin - JBRJ).*

*Mais importante do que campos mínimos ou estrutura de dados são os metadados dos recursos de informação associados à biodiversidade.*

3. Como estes dados devem estar descritos e disponíveis?

*Em padrões de dados e metadados internacionalmente/nacionalmente aceitos e disponibilizados através de serviços (webservices)*

4. Deverão ser definidos modelos de tabelas (planilhas) de dados? Em caso positivo, estes modelos deverão ser apresentados.

*Não. Dependendo do tipo de dados, padrões internacionais podem ser adotados ou adaptados.*

*Perguntas específicas de cada instituição podem demandar definições de novos padrões.*

*Necessidade de interação entre 'coletores' e gerentes de bancos de dados.*

Produtos esperados:

1. Documento com diretrizes para estruturação de bancos de dados em biodiversidade.
2. Subsídios para a execução dos produtos dos consultores referentes à avaliação dos sistemas no âmbito do MMA e vinculadas:

*Diretrizes gerais devem traçar "melhores práticas" / requisitos básicos para criação de sistemas de informação/bancos de dados dentro das instituições / projetos vinculados ao MMA que possibilitem organização e acesso a dados de qualidade.*

*Busca e checagem de dicionários e ontologias existentes (p.ex. NBII, Sistema Cassia - Von Humboldt, trabalho do GBIF com termos usados em Biodiversidade).*

*O padrão de metadados deve ser comum entre os repositórios, que pode ser centralizado ou distribuído, não mutuamente excludente.*

*Existindo-se um padrão, siga-o ou adapte-o.*

## **TEMA 1 - G2: Diretrizes para estruturação de bancos de dados em biodiversidade**

1. Quais *TEMAS* mais relevantes para a gestão da biodiversidade - *instituições ou processos que já completam, ainda que parcialmente, estas lacunas?*

- *Monitoramento (alterações e tendências) – COMOB/ICMBio*
- *Espécies Ameaçadas – CGESP/ ICMBio e CNCFlora/JBRJ*
- *Identificação de ocorrência/ Inventário - Vários*
- *Mapas de ocorrência para áreas prioritárias - Vários*
- *Mapeamento de Fitofisionomias e cobertura vegetal – SNIF/SFB*
- *Ecossistemas*
- *Fatores de perda de biodiversidade – DETER, DETEX e PRODES/INPE*
- *Inventário Florestal – SNIF/SFB*
- *Dados populacionais/ abundância de espécies - CGESP/ ICMBio e JBRJ (para spp ameaçadas)*
- *Monitoramento de pós-licença*
- *Dados taxonômicos – CGESP/ICMBIO, Jardim Botânico*

*Diretrizes Gerais:*

- *Recomenda-se a apresentação dos sistemas citados para o grupo de trabalho do MMA;*
- *Padrão de interoperabilidade de Governo Eletrônico/ e-Ping;*
- *Diretrizes da infraestrutura nacional de dados espaciais – INDE/CONCAR;*
- *Adoção de protocolos e padrões internacionais para dados de biodiversidade*

2. Quais os dados mínimos (em biodiversidade) para a integração de sistemas diversos? (deverão ser observadas as “*informações mínimas necessárias para a integração de bancos de dados de espécies (ou espécimes?)*”, abaixo, de acordo com o Relatório do seminário de 2010).

- *Registro único para os espécimes coletados e locais de depósito;*
- *Nome científico;*
- *Localidade (coordenadas geográficas, altitude/profundidade, UF, bioma, sítio de coleta, tipo de vegetação, habitat);*
- *Data e hora;*
- *Equipe (responsável pela expedição, coletor, assistente); método de coleta, captura ou marcação; esforço de coleta; determinador da espécie (pessoa que identificou o material);*
- *Cada registro deve conter atributos do sítio de coleta;*
- *O sistema deve possuir campos estruturados e relacionar o nome científico com toda a hierarquia taxonômica.*

*Sem a compreensão clara e detalhada do escopo do sistema/rede a ser desenvolvido o grupo não se sentiu à vontade em estabelecer campos mínimos, para não incorrer no erro de omitir campos importantes, ou restringir o preenchimento de dados complementares necessários.*

3. Como estes dados devem estar descritos e disponíveis?

- *Adoção de sistemas descentralizados respeitando as estruturas já em curso nas diferentes instituições*

- *Adoção de formato acessível/ disponível/ compatível de metadados nas instituições provedoras*

4. Deverão ser definidos modelos de tabelas (planilhas) de dados? Em caso positivo, estes modelos deverão ser apresentados.

*Não. Cada instituição possui autonomia em adotar o modelo de dados que atenda suas necessidades específicas e da integração proposta.*

## **TEMA 2 - Metadados: diretrizes para definição e adoção de padrões/ferramentas pelas instituições vinculadas do MMA.**

- *Padrões de metadados que tenham documentação e sejam amplamente utilizados, como os adotados pelo TDWG, GBIF (por exemplo DarwinCore, DublinCore, EML, Perfil de Metadados Geoespaciais determinado pela Infra-Estrutura Nacional De Dados Espaciais (o qual é baseado no padrão ISO 19115), MRTG para dados multimídia), atentando para as particularidades de cada aplicação;*
- *A implementação de padrões de metadados deve assumir a adoção de dicionários de dados e ontologias (levando em consideração modelos já existentes – RDF, OWL, <http://www.gbif.org/communications/resources/print-and-online-resources/bionomenclature/>) (a consultoria em tecnologia da informação deve contemplar durante o diagnóstico dos sistemas/bancos a criação de um glossário de atributos);*
- *Sugestões para capacitação – nivelamento de todos envolvidos no processo de integração, em relação aos padrões e ferramentas (oficinas de nivelamento hierarquizadas em função do grau de envolvimento com o processo);*
- *Treinamento institucional (nivelamento MMA e vinculadas) como gestor de informação em biodiversidade (provedor de dados) – estudar a possibilidade de criação de um mestrado profissional em gestão de informação em biodiversidade (tomando como modelo a demanda do ICMBio ao JBRJ para criação de um mestrado em monitoramento) (articulação com CAPES no âmbito do projeto SIBBR);*
- *Capacitação em gestão da informação sobre biodiversidade (graduação, pós-graduação - sugerir articulação com o MEC no âmbito do projeto SIBBR);*
- *Produção de material didático para público completamente leigo, divulgação de vídeo-aulas;*
- *Mapeamento dos candidatos à capacitação nas diferentes instituições e desenvolver exercício com os dados institucionais;*

## **Tema 3: Padrões e protocolos de comunicação**

- 1) Quais recomendações para adoção de padrões e protocolos de interoperabilidade?
- 2) É necessário o estabelecimento de um protocolo único?
- 3) Quais requisitos mínimos para integração de dados no âmbito do MMA?

- *Padrões de comunicação que tenham documentação e sejam amplamente utilizados, como os adotados pelo TDWG, GBIF (por exemplo DarwinCore, TAPIR, DIGIR), do E-ping, Perfil de Metadados Geoespaciais determinado pela Infra-Estrutura Nacional De Dados Espaciais (o qual é baseado no padrão ISO 19115);*
- *Utilizar webservices;*

## **GRUPO DE TRABALHO INTEGRAÇÃO DE SISTEMAS DE INFORMAÇÃO E BANCOS DE DADOS SOBRE BIODIVERSIDADE NO ÂMBITO DO MMA**

### **CONSIDERAÇÕES GERAIS (requisitos mínimos para todo processo)**

- Adotar: recomendações do E-ping e da INDE; ferramentas de código aberto; padrões de usabilidade e acessibilidade;
- Adotar padrões e protocolos nacional/internacionalmente aceitos;
- Considerar a Política de Informação do MMA (Portaria MMA 160/09);
- Trilogia: Recursos Humanos capacitados, Recursos computacionais e Ferramentas de implementação adequados, Governança do sistema;
- Importância do estabelecimento de uma sólida e transparente Política Nacional para disponibilização e utilização de informações em biodiversidade (garantindo a segurança necessária para dados sensíveis e reservados – articulações com ABIN) (importância no estabelecimento de políticas de dados em todos os níveis, MMA e vinculadas, preferencialmente em consonância com as discussões no âmbito do SIBBr-MCT);
- Possibilidade de um diagnóstico das políticas referentes ao estabelecimento de políticas de informação como subsídio ao estabelecimento da política do MMA para dados/informação de biodiversidade (discutir com SIBBr-MCT a possibilidade de financiamento para tal);
- Possibilidade de criação de um grupo de discussão em rede e um repositório experimental (tb na rede), como subsídio à discussão do grupo e ao trabalho do consultor (POLI USP);
- Criação de uma página no site do MMA para disponibilização das informações produzidas no GT e das que sirvam como subsídio às discussões;
- Institucionalização do GT visando sua continuidade, o comprometimento das instituições representadas e facilitação da participação dos componentes;

# **GRUPO DE TRABALHO INTEGRAÇÃO DE SISTEMAS DE INFORMAÇÃO E BANCOS DE DADOS SOBRE BIODIVERSIDADE NO ÂMBITO DO MMA**

## **PRÓXIMOS PASSOS**

- Criação do portal na POLI-USP (imediato)
- Envio da memória da reunião consolidada (MMA – data 28/12/2011)
- Revisão e sugestões ao documento (Todos – data 31/01/2012)
- Formalização do grupo e convite às instituições (MMA – tentativa até 31/01/2012);
- Disponibilização no portal na POLI-USP
  - (documentos-base para nivelamento dos participantes)
  - (documentos relacionados ao diagnóstico de maturidade institucional)
  - (documentos tutoriais – que subsidiem as discussões, iniciando com metadados)
  - (discussão do formato final às diretrizes – textos colaborativos, data tentativa 15/02/2012)
  - (avaliação institucional quanto ao grau de atendimento institucional às diretrizes definidas, data tentativa 15/02/12)

Próxima reunião (data tentativa – 05 a 09/03/12)

- - Apresentação e validação de uma proposta de funcionamento do GT - MMA;
- - Propostas para evolução/desenvolvimento das provas conceituais (clima/biodiversidade) Eduardo Dalcin (JBRJ) e Pedro Pizzigatti (Poli USP);

**GRUPO DE TRABALHO**  
**INTEGRAÇÃO DE SISTEMAS DE INFORMAÇÃO E BANCOS DE DADOS SOBRE**  
**BIODIVERSIDADE NO ÂMBITO DO MMA**

**PARTICIPANTES**

<b>PARTICIPANTES</b>	<b>INSTITUIÇÃO</b>	<b>TELEFONE</b>	<b>EMAIL</b>
Alessandra Andrade	ICMBIO	61 33419389	alessandra.andrade@icmbio.gov.br
Alexsandro Santos	TAMAR/ICMBio	71 36761045	alex@tamar.org.br
Ana Cristina Lacerda	SFB	61 20287571	ana.lacerda@florestal.gov.br
Bernardo Wernick	ICMBIO		bernardo.wernick@icmbio.gov.br
Débora Drucker	EMBRAPA	19 32116200	debora@cnpem.embrapa.br
Eduardo Dalcin	JBRJ	21 32042116	edalcin@jbrj.gov.br
Fabíola Ladeira	DILIC/IBAMA		fabiola.ladeira@ibama.gov.br
Flávia Costa	INPA	92 36431892	flaviacosta001@gmail.com
Graziele Batista	DBFLO/IBAMA	61 33161165	graziele.batista@ibama.gov.br
Helio Jorge	PROBIO II/MMA	61 85802030	helio.cunha@mma.gov.br
Joseílson Costa	ICMBIO	83 32455001	joseilson.costa@icmbio.gov.br
Keila Macfadem	MMA	61 20282027	keila.juarez@mma.gov.br
Luis Felipe Correa	DILIC/IBAMA	61 33161972	luis.correa@ibama.gov.br
Maria Cristina Braga	SEPED/MCTI	61 33178610	maria.braga@mct.gov.br
Pedro Pizzigatti Corrêa	POLI/USP	11 30915267	pedro.correa@poli.usp.br
Rafael Fonseca	GIZ	61 82860022	rafael.fonseca@giz.de
Renata Araujo	CTI/ICMBio		renata.araujo@confidereit.com.br
Ricardo Machado	DILIC/IBAMA	21 71809204	ricardo.ibama@gmail.com
Ricardo Otoni	DILIC/IBAMA	79 37127450	otoni.52@gmail.com
Roberto Silva	IBAMA		roberto-victor.silva@ibama.gov.br
Rodrigo Frolita	IBAMA		rodrigo.frolita@ibama.gov.br
Rodrigo Jorge	ICMBio		rodrigo.jorge@icmbio.gov.br
Rodrigo Koblitiz	DILIC/IBAMA	61 33161595	rodrigo.koblitiz@ibama.gov.br
Rosemary Oliveira	COMOB/ICMBio	61 33419391	rosemary.oliveira@icmbio.gov.br
Sergio Morbiolo	UFSCAR	15 91356802	morbiolo@yahoo.com