

**PROJETO DE PESQUISA:**

**SISTEMA DE INFORMAÇÃO DISTRIBUÍDO PARA COLEÇÕES BIOLÓGICAS:  
A INTEGRAÇÃO DO *SPECIES ANALYST* E SINBIOTA (FAPESP)**

## XXXXXXXXXXXXXXXXXXXX

Esta proposta trata da implementação de um sistema distribuído de informação sobre diversidade biológica (fauna, flora e microbiota) do Estado de São Paulo, associado a um sistema de previsão de distribuição geográfica de espécies, baseado em modelagem matemática.

O projeto pretende explorar os mais recentes avanços na área de bancos de dados distribuídos, protocolos de comunicação, conectividade (especialmente a Internet 2 e inteligência artificial, para atingir os três objetivos a seguir: i) desenvolver e implantar um sistema distribuído de informação sobre biodiversidade integrando várias coleções biológicas (dados históricos) do estado, o sistema de informação ambiental do programa Biota/Fapesp (dados atuais) e a rede do projeto *Species Analyst*; ii) desenvolver e implementar modelos matemáticos para previsão de distribuição geográfica de espécies, usando como entrada os dados obtidos na rede descrita no item anterior; iii) desenvolver aplicativos específicos para a resolução de problemas em biodiversidade e meio-ambiente, tais como estudos de espécies invasoras, mudanças climáticas, proteção de espécies ameaçadas, planejamento de áreas de conservação, entre outros, utilizando os dados do sistema distribuído (item i) e os modelos de distribuição geográfica de espécies (item ii) e torná-las disponíveis através da Internet, para a comunidade científica, comunidade educacional, formuladores de políticas ambientais e tomadores de decisão na área de conservação ambiental.

O projeto conta com a participação inicial de doze coleções biológicas (3 herbários, 2 coleções de ácaros, 3 de peixes, 1 de algas e 3 de microrganismos) e prevê a repatriação de dados de subcoleções específicas fora do estado de São Paulo (no Brasil e no exterior).

## XXXXXXXXXXXXXXXXXXXX

O Brasil está definindo mecanismos para cumprir com as obrigações da Convenção sobre a Diversidade Biológica (ONU, 1992). A disponibilidade e o acesso à informação de qualidade é fundamental para a tomada de decisão e para viabilizar trabalhos colaborativos e o processo de avaliação e participação da sociedade em geral.

O programa Biota/Fapesp foi implementado com esta visão e missão. Tem atuado na integração da comunidade científica e de seus dados, explorando as competências existentes e desenvolvendo tecnologias inovadoras para coletar, gerenciar, integrar, analisar e compreender a base de informação. A primeira fase do programa foi a de avaliar a base de conhecimento sobre os diferentes grupos taxonômicos que ocorrem no Estado de São Paulo. Como resultado temos a publicação de 7 volumes da série "Biodiversidade do Estado de São Paulo - Síntese do Conhecimento ao Final do Século XX" em 1999.

Em 1999, com o lançamento oficial do programa Biota/Fapesp, O Instituto Virtual da Biodiversidade, teve início o projeto responsável pelo desenvolvimento do sistema de informação ambiental do programa, o *SinBiota* (CRIA-BDT, 1999), que foi desenvolvido pelo Centro de Referência em Informação Ambiental - BDT e pelo Instituto de Computação da Unicamp. A comunidade científica adotou uma ficha padrão de coleta e está inserindo *online*, via internet, os seus dados de coleta com a lista de espécies associada em um sistema central, mantido pelo CRIA-BDT. O Instituto Florestal está digitalizando toda a base cartográfica do estado de São Paulo em escala de 1:50.000 que está sendo auditada pela Unicamp.

Em dezembro de 2000 foi lançado o Atlas Biota/São Paulo (CRIA-BDT, 2000), o primeiro componente do *SinBiota* a integrar a informação sobre biodiversidade fornecida pelos mais de 20 projetos temáticos do programa com a informação espacial contida na base cartográfica. Este sistema consiste em uma ferramenta baseada num Sistema de Informação Geográfica (Câmara et al. 1996), disponível através da Internet, em que é possível realizar consultas espaciais sobre as coletas armazenadas, observar sua distribuição espacial no Estado em contraste com outras informações contidas no banco de dados textual (autor, local da coleta) ou espacial (uso do solo e cobertura vegetal, por exemplo).

Apesar do Programa Biota/Fapesp e do *SinBiota* virem contribuindo de maneira decisiva para o aumento do conhecimento sobre a biodiversidade do estado, uma parte importante dos dados existentes ainda permanece inacessível. São os dados históricos que foram levantados nos últimos 300 anos de exploração científica da biodiversidade do planeta, iniciada pelas expedições européias, seguidas por pesquisadores americanos e mais recentemente, por

instituições locais, que hoje se encontram em diversos museus de história natural, herbários e coleções de culturas no mundo todo.

Os dados históricos sobre a diversidade de espécies estão disponíveis em diversos sistemas heterogêneos, informatizados ou não, baseados em fronteiras nacionais ou institucionais. Em geral, a comunidade científica está disposta a fornecer a informação, mas os sistemas geralmente são ineficientes e de difícil acesso. Esta informação freqüentemente não está em formato digital e é considerada "propriedade" das instituições individuais. Assim, o acesso à informação de cada coleção é feita em base individual, tornando o acesso à totalidade da informação praticamente impossível.

Para ilustrar a importância de um sistema de informação integrado temos o exemplo do banco de dados sobre aves do México (*Peterson et al., 1998*). Dados de 43 coleções científicas foram coletados e armazenados em um banco de dados centralizado, totalizando mais de 300 mil registros de espécimens para o país, num esforço que levou 4 anos para ser completado. Dado que a maior coleção representava individualmente somente 16% da totalidade dos espécimens, os benefícios de se obter um banco de dados completo não puderam ser notados até que todas as coleções tivessem sido integradas. Uma vez agregado, o sistema de informação tem sido instrumento em avanços nas áreas de conservação, planejamento do uso da terra, no desenho de sistemas de áreas de proteção e para biologia básica.

Considerando os recursos atualmente disponíveis, o procedimento adotado pelo México para integrar os dados de um único grupo taxonômico de várias coleções biológicas, hoje é considerado altamente ineficiente. O grande desafio está na atualização dos dados. A qualidade do sistema começa a decair a partir do exato momento de sua conclusão, uma vez que não foram criados mecanismos de atualização.

Os dados sobre biodiversidade, embora existentes em quantidades suficientes para vários grupos taxonômicos, não são acessíveis, e portanto, não são incorporados de maneira integrada. Via de regra, os estudos não levam em conta o poder analítico e a completude da informação, refletindo na qualidade de seus resultados.

Hoje, com a Internet 2 e com o desenvolvimento de novas tecnologias, principalmente no tocante à interoperabilidade de sistemas, temos a real oportunidade de explorar o desenvolvimento de sistemas distribuídos para integrar a informação existente nas coleções. O presente projeto de pesquisa visa integrar dados históricos contidos nas coleções, com dados atuais (levantamento dos projetos Biota/Fapesp) e dados históricos de coleções do exterior (*Species Analyst*, a ser citado a seguir).

A principal iniciativa para integração de dados históricos de coleções biológicas no contexto internacional, baseada em um modelo distribuído, é o projeto *Species Analyst* ([www.speciesanalyst.net](http://www.speciesanalyst.net)). Trata-se de um sistema de informação sobre biodiversidade distribuído na internet. Bases de dados com informação sobre ocorrência de espécies são integradas usando protocolos de recuperação da informação baseado em Z39.50 (ANSI/NISO, 1992) e HTTP/XML (W3C, 1997). O sistema hoje integra 19 bancos de dados distribuídos em 9 instituições. São cerca de 14 milhões de espécimens, sendo que existem mais 38 instituições que já manifestaram o seu interesse em participar da rede, o que deverá aumentar este número para cerca de 50 milhões de espécimens. Estima-se que o número total de espécimens em museus de história natural seja de aproximadamente 3 bilhões e que cerca de 5% desse total estão computadorizados. Se esta estimativa estiver correta, significa de cerca de 30% dos registros existentes farão parte do *Species Analyst*.

O *Species Analyst* oferece um conjunto de ferramentas para modelagem de distribuição de espécies, dentre as quais se destaca o *GARP* (*Genetic Algorithm for Rule-Set Prediction*) (Stockwell e Noble, 1992). Este algoritmo permite previsões bastante precisas da distribuição geográfica de uma dada espécie, a partir dos dados de seus pontos de ocorrência (fornecidos pelo banco de dados distribuído), e de um conjunto de coberturas geográficas (temperatura, vegetação, precipitação, etc.) relevantes para a espécie em questão.

O Brasil tem participado desta iniciativa através da colaboração entre o CRIA-BDT e o Museu de História Natural da Universidade do Kansas, nos Estados Unidos, através do desenvolvimento conjunto de software e aplicativos e do intercâmbio de membros de equipe.

## XXXXXXXXXXXXXXXXXXXX

O foco principal desse projeto é resolver o problema da integração dos dados históricos contidos nas coleções biológicas, tornando-os amplamente disponíveis para a comunidade científica e educacional. Além disso, pretende-se desenvolver aplicativos que façam uso dos dados integrados, visando criar ferramentas acessíveis a um público mais amplo.

Em resumo, o projeto tem três objetivos principais:

- i) Implementar uma rede distribuída de informação sobre biodiversidade para o Estado de São Paulo, explorando as novas tecnologias e melhor conectividade (Internet 2) para a integração de bases de dados heterogêneas e independentes. As bases de dados a serem integradas incluirão 12 coleções biológicas do Estado de São Paulo, a rede *Species Analyst* e o *SinBiota*;
- ii) Integrar ferramentas de modelagem de nicho ecológico e de previsão de distribuição geográfica de espécies ao sistema, utilizando dados da rede distribuída descrita no item i);
- iii) Desenvolver aplicativos que façam uso dos dados integrados na rede de informação distribuída descrita no item i) e dos modelos matemáticos de distribuição no item ii) para dar suporte à comunidade científica, educacional, tomadores de decisão e formuladores de políticas ambientais.

## XXXXXXXXXXXXXXXXXXXXXXX

O procedimento utilizado para converter dados primários de ocorrência em mapas de distribuição indicando presença e ausência (provável) das espécies é uma etapa crítica para o sucesso dos produtos a serem obtidos. No entanto estes procedimentos são significativamente distintos entre si e as suas diferenças têm implicações importantes. A exploração dessas diferenças pode permitir a elaboração de um conjunto de critérios desejáveis para uma metodologia ideal.

No desenvolvimento de um modelo que resulta na distribuição geográfica provável de uma espécie, dois erros são possíveis: omissão e sobreprevisão. Omissão significa não incluir áreas que de fato são habitadas e sobreprevisão é incluir na previsão áreas que de fato não são habitadas ou onde a espécie comprovadamente não ocorre. A meta é atingir um algoritmo ideal que minimize esses erros simultaneamente. A qualidade e a quantidade de dados são a base para um modelo mais preciso.

O nicho ecológico fundamental de uma espécie pode ser definido como sendo um conjunto de condições ecológicas com as quais as populações conseguem se manter (Begon et al, 1998). É definido como sendo um espaço ecológico/ambiental multidimensional. O nicho fundamental, que é o foco da modelagem, deve ser cuidadosamente distinguido do nicho de ocorrência da espécie para manter os esforços de modelagem focados nas dimensões ecológicas importantes para a espécie em particular.

Um primeiro exemplo de algoritmo de modelagem é o BIOCLIM - *bioclimatic envelope* (Nix, 1986, Busby, 1991, McMahon et al. 1996). Este algoritmo tenta determinar a região geográfica que contém as características climáticas similares às de onde a espécie é encontrada. Em geral, este algoritmo tende a apresentar muita sobreprevisão.

Outro exemplo é o GARP - *Genetic Algorithm for Rule-set Prediction* (Stockwell e Noble, 1992). Este algoritmo inclui numerosos algoritmos individuais (isto é, BIOCLIM, regressão logística) e usa a combinação de sua habilidade analítica para produzir regras dos componentes em um conjunto de regras (*rule-set*) mais amplas. Assim, parte da distribuição das espécies poderá ser determinada como dentro ou fora do nicho, baseada em regras diferentes de algoritmos diferentes. Testes extensivos com GARP indicam excelente habilidade de previsão de distribuição de espécies.

O método de modelagem GARP vem sendo aprimorado de forma colaborativa entre as equipes da Universidade de Kansas e do CRIA-BDT, de forma a permitir o uso do conjunto de dados da rede distribuída nos programas de previsão e análise.

## XXXXXXXXXXXXXXXXXXXXXXX

O sistema proposto neste projeto será dividido em três módulos relativamente independentes:

- a rede de coleções biológicas;
- a modelagem para distribuição geográfica de espécies; e,
- os aplicativos.

A figura 1 mostra um diagrama simplificado do sistema proposto. Na figura, cada módulo está representado por um retângulo e as setas representam o fluxo de dados. Um retângulo adicional representa o usuário final que acessa o sistema através da Internet.

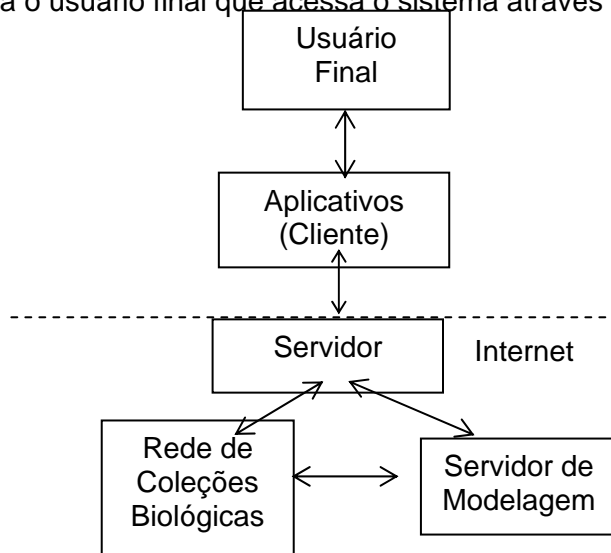


Figura 1: Diagrama simplificado do sistema proposto.

O **primeiro componente** do sistema consiste na rede responsável pela integração dos dados existentes nas instituições participantes. Este é um dos componentes mais importantes, considerando o modelo proposto, uma vez que permite que cada coleção tenha total domínio sobre os seus dados, que serão gerenciados e armazenados em sistemas proprietários, em diferentes formatos, atendendo às necessidades locais e, ao mesmo tempo, serão integrados e disponibilizados em um formato padrão "pronto-para-uso" (*ready-to-use format*). Assim, esta estrutura não interfere na autonomia das coleções quanto à sua política de informação e, ao mesmo tempo, permite a utilização dos dados de forma integrada com os dados de outras coleções, de uma forma transparente para o usuário final.

O **segundo componente** consiste em uma série de ferramentas que convertem dados primários de ocorrência de espécies em mapas contínuos que indicam a probabilidade de ocorrência ou não-ocorrência de cada espécie. Essas ferramentas, baseadas em modelagem estatística do nicho ecológico de espécies, têm um grande impacto na qualidade dos resultados



obtidos em aplicações utilizadas para estudos de biodiversidade, especialmente quando a quantidade de dados é pequena (e.g., para espécies raras ou ameaçadas e espécies pouco estudadas). No caso do Brasil, essas ferramentas são ainda mais críticas dadas as restrições impostas pela falta de recursos, pela dimensão continental do país e pela sua rica diversidade biológica.

O **terceiro componente** consiste em uma série de aplicativos, cada um com um objetivo específico, destinado a atender a demandas específicas. Aplicativos são baseados no paradigma cliente/servidor. No aplicativo "cliente" é implementada uma interface com o usuário e, em muitos casos, algum processamento de dados pode ser oferecido. Clientes podem ser navegadores internet (com interface web), ou aplicativos específicos, como planilhas (MS-Excel, por exemplo), bancos de dados (MS-Access) e sistemas de informação geográfica (como ArcView GIS). O segundo retângulo representa o servidor de aplicativos onde a maior parte dos dados é processada.

Estes últimos dois módulos (modelos e aplicativos) são independentes do primeiro componente, no sentido que eles consideram a entrada de dados como sendo proveniente de uma fonte homogênea única. Não é necessário considerar a complexidade da rede de onde os dados são obtidos. É evidente que a qualidade dos resultados obtidos dos aplicativos depende da qualidade dos dados do sistema, ou seja, depende da informação fornecida pela rede de coleções biológicas.

## **X.1 Rede de Coleções Biológicas**

Para atingir os resultados desejados, a rede deverá preencher os seguintes requisitos:

- Disponibilidade: a rede deve estar disponível, em funcionamento sempre;
- Tolerância a falhas: a rede deve ser capaz de operar quando alguns nós não estiverem funcionando;
- Desempenho e Robustez: a rede deve ser capaz de processar e responder em um tempo razoável um grande número de solicitações simultâneas;
- Descoberta de novos nós: a rede deve ter mecanismos automáticos ou manuais para inclusão de novos nós;
- Independência de plataforma: o software deve operar em ambientes computacionais heterogêneos; e,
- Escalabilidade: a rede deve ter capacidade para aumentar o número de nós, sem que seja necessário desenvolver novos componentes de *software*.

A figura 2 a seguir mostra uma possível arquitetura para a rede.

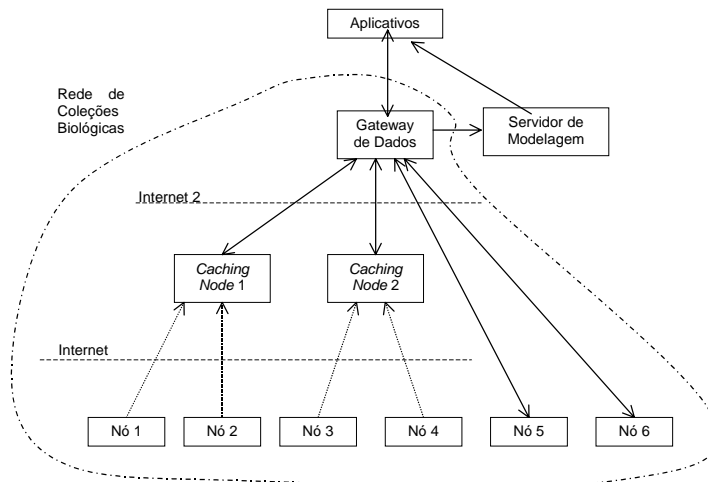


Figura 2: Diagrama da rede de coleções biológicas.

O **gateway de dados** é responsável pela integração e disponibilidade dos dados aos demais módulos do sistema. Este recebe as requisições ou consultas dos diversos aplicativos e as envia para cada nó da rede. Cada nó processa a requisição e retorna um conjunto de dados ao *gateway*, que irá integrá-los em um conjunto único, homogêneo, que é retornado ao aplicativo.

Para permitir a integração dos dados pelo *gateway*, é necessário determinar um conjunto padrão de campos que atendem às necessidades dos aplicativos e que são fornecidos pela rede. Inicialmente, o conjunto de campos que será servido pela rede é o descrito em ZBIG, 1999. Além disso, é necessário utilizar um protocolo de comunicação que permite que o *gateway* envie suas consultas ao nó e que seja capaz de receber as respostas. É importante escolher um protocolo padrão de largo uso.

Para utilizar a rede de uma forma mais racional, a sua arquitetura deve prever mecanismos de gerenciamento de metadados que permite ao *gateway* identificar quais os nós (ou *caching nodes*) da rede deverão ser ativados para responder a uma determinada consulta, evitando acesso aos nós que não têm a informação desejada.

No nível mais baixo da rede estão os *caching nodes* e os nós onde estão armazenados os dados primários. Esses nós, via-de-regra estão fisicamente localizados nas instituições que mantém as coleções biológicas. É no nível de nó que os dados do sistema de gerenciamento

das coleções (ambiente heterogêneo) são convertidos em um conjunto de campos padrão mantidos pela rede (ambiente homogêneo).

Uma característica importante que precisa ser considerada no projeto da rede é a heterogeneidade que existe entre as instituições participantes. Além da significância do acervo, temos que considerar a capacitação e infra-estrutura em *software* e *hardware* e a conectividade. Para minimizar o problema de coleções com acervos relevantes que têm capacidade limitada é que se projeta uma camada de componentes denominados *caching nodes*. Esses componentes serão responsáveis por armazenar cópias dos registros das coleções que não podem participar plenamente da rede devido a limitações de infra-estrutura de computação ou de comunicação. É importante sempre ter em mente que a velocidade de resposta da rede dependerá da velocidade de seu nó mais lento. Idealmente o *gateway* e os *caching nodes* deverão estar conectados à Internet 2.

## **X.2 Integração dos Nós**

O objetivo maior da proposta é a integração das diferentes fontes de informação sobre a biodiversidade do Estado de São Paulo. Isso requer, além da tecnologia adequada, o compromisso das diferentes instituições de compartilharem seus dados e informações, e de cooperarem e colaborarem com a rede. Dado esse desafio, limitamos o escopo deste projeto para conectar e integrar dados de determinadas instituições, sem nos responsabilizarmos pela informatização dos acervos. Acreditamos que esse projeto deverá estimular avanços significativos nessa área e iremos estabelecer algumas diretrizes que poderão orientar as coleções nesse sentido.

O trabalho da equipe consistirá em estudar o sistema de cada coleção *in loco*, instalar e configurar o *software* de conexão à rede, viabilizando a interoperabilidade dos sistemas. O objetivo de cada visita técnica a cada coleção será avaliar:

- o *software* e *hardware* utilizados pela coleção;
- formato em que os dados se encontram; e,
- sua infra-estrutura de comunicação.

## **X.3 Desenvolvimento do Modelo de Distribuição de Espécies**

Com relação aos modelos matemáticos de distribuição geográfica de espécies (GARP em especial), deverão ser trabalhados os seguintes pontos:

- teste de robustez e sensibilidade a uma série de fatores como tamanho da amostra de pontos de ocorrência conhecidos; riqueza de dados ambientais e efeitos de escala espacial dos *layers* ambientais;
- ajuste do parâmetro "*fitness*" no algoritmo genético para pesar os erros de omissão e sobreprevisão de forma distinta;
- permitir a escolha de algoritmos a serem usados
- permitir que se acrescente novas regras ao algoritmo genético

#### **X.4 Aquisição de Dados Espaciais**

Além de dados sobre a ocorrência de espécies, todos os modelos requerem um conjunto de dados ambientais. Para análises de pequena escala, os dados ambientais do Biota/Fapesp podem ser utilizados. No entanto, para análises de larga escala é necessário um novo conjunto de *layers* geográficos. Dada a sua natureza de larga escala, não precisam ser tão detalhados quanto o Biota/Fapesp, o que reduz o custo de sua aquisição. Os *layers* mais relevantes para análise são vegetação, precipitação, humidade, temperatura, solo e geologia, e seriam desenvolvidos a uma resolução de 1x1 km por pixel para todo o Estado de São Paulo.

#### **X.5 Desenvolvimento de Aplicativos**

A informação que estará disponível na rede poderá ter inúmeros usos, muitos de clara importância para a economia, agricultura, ciência e conservação, incluindo:

- Biologia básica - O estudo das diferentes facetas da biologia e história natural das espécies é facilitado pelo acesso fácil a dados sobre espécimens. Particularmente, detalhes sobre distribuição e ecologia, bem como aspectos sobre biogeografia, dependem do conhecimento acurado sobre a distribuição geográfica das espécies. A inclusão de uma metodologia "por inferência" que tem por objetivo a modelagem de nichos ecológicos das espécies, oferece grandes possibilidades de nova compreensão ou conhecimento da espécie.
- Conservação de espécies ameaçadas - A ocorrência de espécies fornece informação importante sobre a distribuição de espécies a níveis populacionais criticamente baixos. Para estes grupos, novos estudos talvez não sejam possíveis e os dados históricos podem ser a única fonte de informação sobre sua distribuição e ecologia (Godown and Peterson, 2000; Peterson and Brisbin, 1998).

- Conservação da biodiversidade - A possibilidade de usar uma grande quantidade de dados e descrever a distribuição da diversidade de espécies pode permitir a designação de sistemas que otimizam a proteção da diversidade (Peterson et al, 2000).
- Espécies invasoras - As novas aproximações de modelagem de nichos ecológicos de espécies, baseadas em dados de ocorrência, oferecem a oportunidade de perguntar se determinadas espécies encontrariam condições apropriadas para a sua sobrevivência fora de sua distribuição geográfica corrente. Esta pergunta poderia também ser formulada para saber se uma espécie poderia "potencialmente" invadir uma região (Peterson and Vieglais, no prelo).
- Mudanças climáticas - Apesar de sabermos que mudanças climáticas drásticas estão ocorrendo, sabemos muito pouco a respeito do efeito destas mudanças nas espécies biológicas e suas comunidades. Podemos também perguntar onde as espécies encontrarão condições apropriadas sob um cenário de mudança climática.
- Saúde pública - doenças infecciosas são elementos biológicos que possuem nichos ecológicos exatamente iguais a organismos vivos. Assim, tanto o -organismo doente como as espécies que o carregam (vetores, espécies afetadas) podem passar por modelagem e ter sua distribuição prevista.
- Aplicações para a agricultura - Ecossistemas agrícolas são grandes coleções de espécies, algumas cultivadas outras não. O sucesso do empreendimento agrícola depende da distribuição ecológica dessas espécies que interagem. Esforços iniciais de modelagem indicam grandes possibilidades em utilizar dados da biodiversidade para otimizar a produção agrícola (Sánchez-Cordero e Martínez-Meyer, 2000)

Cada uso dos dados requer o desenvolvimento de um aplicativo específico. Um conjunto de aplicações (como as descritas acima) será implementado como parte do projeto.

A metodologia para o desenvolvimento de aplicativos será a seguinte:

- entrevistas com usuários especialistas para identificar possíveis aplicações;
- definição da lista dos aplicativos mais relevantes a serem desenvolvidos;
- lista das necessidades em termos de sistema para cada aplicativo;
- seleção dos aplicativos que serão desenvolvidos baseada em viabilidade e relevância; e,
- utilização de metodologia de desenvolvimento padrão de *software* para todas as aplicações selecionadas.

Devido à complexidade do sistema, todo o desenvolvimento de *software* seguirá o método padrão aceito de engenharia de *software*.

## **X.6 Coleções Biológicas**

As seguintes coleções foram contactadas e concordaram em participar do projeto:

- Herbário do Instituto de Botânica
- Herbário do Depto. de Botânica do Instituto de Biociência da USP
- Herbário do Depto. de Botânica do Instituto de Biologia da Unicamp
- Coleção de Ácaros do Depto. de Entomologia, ESALQ, USP
- Coleção de Ácaros do Depto. de Zoologia, UNESP São José do Rio Preto
- Coleção de Peixes, do Museu de Zoologia da USP
- Coleção de Peixes, do Depto. de Zoologia da USP Ribeirão Preto
- Coleção de Peixes, do Depto. de Zoologia da UNESP São José do Rio Preto
- Coleção de Microrganismos do Instituto de Botânica
- Coleção de Microrganismos do Instituto Biológico de Campinas
- Coleção de Microrganismos da Coleção de Culturas Tropical, Fundação André Tosello
- Coleção de Algas do Instituto de Botânica

## **X.7 Equipe**

- Vanderlei Perez Canhos (PhD, Fac. Engenharia de Alimentos, UNICAMP)
- Sidnei de Souza (Diretor de Informática, CRIA-BDT)
- Benedito A. Cruz (Coordenador de Suporte, CRIA-BDT)
- Alexandre Marino (Analista de Sistemas, CRIA-BDT)
- Dora Lange Canhos (Diretor de Projetos, CRIA-BDT)
- Marinez Ferreira de Siqueira (MSc, Gerente de Informação, CRIA-BDT)
- Luceli de Souza (MSc, Coordenador de Zoologia, CRIA-BDT)
- Érica Speglich (Bióloga, Pesquisador, CRIA-BDT)
- Paula Felício Drummond de Castro (Bióloga, Pesquisador, CRIA-BDT)
- A contratar: Ricardo Scachetti Pereira (MSc, Pesquisador, CRIA-BDT): coordenador da equipe de desenvolvimento
- A contratar: Especialista em redes de computadores;
- A contratar: 2 Analistas/Programadores de sistemas.

XXXXXXXXXXXXXXXXXXXXXXXXXXXX

Atividade	Semestres					
	1	2	3	4	5	6
Contratação dos membros da equipe	X					
Levantamento dos requisitos da rede de coleções	X					
Projeto e análise da rede de coleções biológicas	X					
Implementação de protótipo para avaliação e escolha da linguagem de programação e plataforma a serem usados na implementação da versão de produção	X	X				
Implementação do <i>software</i> da rede de coleções	X	X				
Conexão e integração dos nós da rede		X	X			
Integração com a rede do <i>Species Analyst</i>			X			
Revisão dos requisitos, projeto e análise do sistema distribuído e atualização dos componentes			X	X		
Manutenção da rede de coleções biológicas		X	X	X	X	X
Repatriação de coleções selecionadas no Brasil e no exterior		X	X	X	X	
Levantamento dos requisitos do <i>software</i> de modelagem de distribuição de espécies			X			
Projeto e análise dos modelos			X			
Implementação dos modelos			X	X	X	
Integração do <i>software</i> de modelagem ao sistema distribuído				X		
Levantamento dos requisitos dos possíveis aplicativos a serem desenvolvidos usando os servidores de dados e modelagem				X		
Seleção dos aplicativos a serem desenvolvidos (com base na viabilidade e relevância)				X		
Análise e projeto dos aplicativos selecionados				X	X	
Implementação dos aplicativos selecionados				X	X	X
Avaliação e acompanhamento dos aplicativos					X	X
Avaliação dos resultados					X	X

XXXXXXXXXXXXXXXXXXXXX

- ANSI/NISO, 1992. Information Retrieval (Z39.50): Application Service Definition and Protocol. Specification American National Standards Institute.
- Begon M., Townsend, C. R., Harper, J. L., 1998. Ecology: Individuals, Populations and Communities, Bk&Cd.
- Busby, J.R. 1991. BIOCLIM - A Bioclimatic Analysis and Prediction System. In: Margules, C.R. & M.P. Austin (eds.) *Nature Conservation: Cost Effective Biological Surveys and Data Analysis*. pp. 64-68. Canberra: CSIRO.
- Câmara, G., Casanova, M. A., Hemerly, A. S., Magalhães, G. C., Medeiros, C. M. B., 1996. Anatomia de Sistemas de Informação Geográfica. Instituto de Computação, Campinas, SP, Brasil.
- CRIA-BDT, IC/Unicamp, 1999. *SinBiota* - Sistema de Informação Ambiental do Programa Biota/FAPESP. <<http://www.biotasp.org.br/sia/>>. Centro de Referência em Informação Ambiental - BDT e Instituto de Computação – Unicamp.
- McMahon, J.P., Hutchinson, M.F., Nix, H.A. and Ord, K.D. (1996). ANUCLIM Version 1 User's Guide. Canberra: ANU, CRES.
- CRIA-BDT, 2000. Atlas Biota/São Paulo. <<http://www.biotasp.org.br/sia/atlas>>. Centro de Referência em Informação Ambiental - BDT e Instituto de Computação – Unicamp.
- Garrido, O. H., A. T. Peterson, and O. Komar. 1999. Geographic variation and taxonomy of the Cave Swallow (*Hirundo fulva*) complex, with the description of a new subspecies from Puerto Rico. *Bulletin of the British Ornithologists' Club* 119:80-91.
- Godown, M. E., and A. T. Peterson. 2000. Preliminary distributional analysis of U.S. endangered bird species. *Biodiversity and Conservation* 9:1-10.
- Nix, H.A., 1986. A biogeographic analysis of Australian Elapid snakes, in Longmore, R. (ed.) Atlas of Australian Elapid Snakes. *Australian Flora and Fauna Series* 8: 4-15.
- ONU, 1992. Convenção sobre a diversidade biológica. Organização das Nações Unidas. Rio de Janeiro, Brasil.
- Peterson, A. T., and I. L. Brisbin. 1998. Genetic endangerment of wild Red Junglefowl *Gallus gallus*? *Bird Conservation International* 8:387-394.
- Peterson, A. T. 1998. Atlas of the distribution of Mexican birds, Ibis.
- etc...