

Greice Cristina Mariano

**“Especificação e Implementação do Banco de Dados  
do Projeto e-phenology”**

CAMPINAS  
2013





Universidade Estadual de Campinas  
Instituto de Computação

Greice Cristina Mariano

## “Especificação e Implementação do Banco de Dados do Projeto e-phenology”

Orientador(a): **Ricardo da Silva Torres**

Co-Orientador(a): **Leonor Patrícia Cerdeira Morellato**

Dissertação de Mestrado apresentada ao Programa  
de Pós-Graduação em Ciência da Computação do Instituto de Computação da  
Universidade Estadual de Campinas para obtenção do título de Mestre em  
Ciência da Computação.

ESTE EXEMPLAR CORRESPONDE À  
VERSÃO FINAL DA DISSERTAÇÃO DEPENDI-  
DIDA POR GREICE CRISTINA MARIANO,  
SOB ORIENTAÇÃO DE RICARDO DA SILVA  
TORRES.

*Ricardo Torres*

Assinatura do Orientador(a)

CAMPINAS

2013

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca do Instituto de Matemática, Estatística e Computação Científica  
Maria Fabiana Bezerra Muller - CRB 8/6162

M337e Mariano, Greice Cristina, 1986-  
Especificação e implementação do banco de dados do projeto e-phenology /  
Greice Cristina Mariano. – Campinas, SP : [s.n.], 2013.

Orientador: Ricardo da Silva Torres.  
Coorientador: Leonor Patricia Cerdeira Morellato.  
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de  
Computação.

1. Banco de dados. 2. Fenologia. 3. Biodiversidade - Banco de dados. 4.  
Sistemas de informação. I. Torres, Ricardo da Silva, 1977-. II. Morellato, Leonor  
Patricia Cerdeira. III. Universidade Estadual de Campinas. Instituto de  
Computação. IV. Título.

Informações para Biblioteca Digital

**Título em outro idioma:** Specification and implementation of the database of the e-phenology project

**Palavras-chave em inglês:**

Database

Phenology

Biodiversity - Databases

Information systems

**Área de concentração:** Ciência da Computação

**Titulação:** Mestra em Ciência da Computação

**Banca examinadora:**

Ricardo da Silva Torres [Orientador]

Milton Cezar Ribeiro

Luiz Camolesi Júnior

**Data de defesa:** 09-08-2013

**Programa de Pós-Graduação:** Ciência da Computação

## TERMO DE APROVAÇÃO

Dissertação Defendida e Aprovada em 09 de Agosto de 2013, pela  
Banca examinadora composta pelos Professores Doutores:



---

**Prof. Dr. Milton Cezar Ribeiro**  
**IB / UNESP**



---

**Prof. Dr. Luiz Camolesi Júnior**  
**FT / UNICAMP**



---

**Prof. Dr. Ricardo da Silva Torres**  
**IC / UNICAMP**



# Especificação e Implementação do Banco de Dados do Projeto e-phenology

Greice Cristina Mariano<sup>1</sup>

09 de agosto de 2013

## Banca Examinadora:

- Ricardo da Silva Torres (*Orientador*)
- Prof. Dr. Milton Cezar Ribeiro  
Instituto de Biociências - UNESP
- Prof. Dr. Luiz Camolesi Júnior  
Faculdade de Tecnologia - UNICAMP
- Prof. Dr. André Santachè  
Instituto de Computação - UNICAMP (*Suplente*)
- Dr. Cleber Mira  
Scylla Bioinformatics (*Suplente*)

---

<sup>1</sup>Apoio financeiro de: Projeto FAPESP-MICROSOFT RESEARCH (processo 2010/52113-5) e bolsa vinculada de mestrado (processo 2011/51523-8) 2011–2013, CNPq e CAPES.



# Abstract

Environmental changes have become an important issue on the world. A representative example of these problems arises in context of studies of phenology. Recently, phenology has gained importance as the simplest and most reliable indicator of the effects of climate change on plants and animals. The shortage or lack of information and monitoring systems in tropical regions, in particularly in South America, has encouraged many centers to develop researches to fulfill this gap. One example includes the e-phenology project. The e-phenology is a multidisciplinary project that combines research in Computer Science and Phenology. The project's main goal is to attack the practical and theoretical problems involved in using new technologies for monitoring plant phenology remotely and integrating obtained data with on-the-ground observations.

In this context, this work presents the specification and implementation of a database to manage the information that should be handled by the e-phenology Project. The proposal allows the integration of phenology data collected from field observations, with climate data obtained from climate sensors and image data obtained by digital cameras. Both the modeling and the implementation of the database were based on studies on plant phenology conducted by biologists and ecologists of the Laboratory of phenology at UNESP of Rio Claro.



# Resumo

As mudanças ambientais tornaram-se uma questão importante na agenda global. Um exemplo representativo desses problemas surge no contexto dos estudos de fenologia. Recentemente, fenologia tem ganho importância como o indicador mais simples e confiável dos efeitos das mudanças climáticas sobre plantas e animais. A escassez ou falta de informações e sistemas de monitoramento em regiões tropicais, em particular, na América do Sul, vêm estimulando diversos centros de pesquisa a desenvolverem trabalhos visando preencher esta lacuna. Um exemplo é o Projeto e-phenology, que é multidisciplinar e combina pesquisas em Ciência da Computação e Fenologia. O principal objetivo do projeto é atacar os problemas práticos e teóricos envolvidos no uso de novas tecnologias para realizar a observação remota da fenologia de plantas e integrar estas informações com os dados de campo.

Neste contexto, este trabalho apresenta a especificação e implementação de um banco de dados para gerenciar as informações que devem ser manipuladas pelo Projeto e-phenology. A proposta apresentada permite a integração de dados de fenologia coletados a partir de observações no campo, com dados climáticos obtidos de sensores de clima e dados de imagens obtidas por câmeras digitais. Tanto a modelagem quanto a implementação do banco de dados tiveram como base os dados dos estudos de fenologia de plantas realizados pelos biólogos e ecólogos do grupo do Laboratório de Fenologia da UNESP de Rio Claro.



*Para os meus guerreiros, queridos e  
amados pais.*



# Agradecimentos

Agradeço, primeiramente a Deus, por todos os dias me guiar, me dar forças para lutar, buscar objetivos e vencer obstáculos, e por me permitir concluir mais esta etapa.

Aos meus pais, Édina Kirali Mariano e José Mariano Filho, pelo apoio sempre presente, dedicação, broncas, ensinamentos e principalmente pela motivação para sempre me dedicar aos estudos e ao aprendizado.

Ao professor Ricardo da S. Torres, por ter acreditado no meu potencial e ter me aceitado como sua orientanda. Obrigada pela oportunidade, pela paciência, pela compreensão e principalmente por sempre me motivar com suas novas idéias e pelos seus ensinamentos.

À professora Patrícia Morellato, da UNESP de Rio Claro, minha co-orientadora, pela oportunidade de participar do Projeto e-phenology e por estar sempre disposta a colaborar com meu desenvolvimento. Obrigada por tirar minhas dúvidas e me corrigir no decorrer deste trabalho.

Aos professores Marco Antonio, Celmar e André da Faculdade de Tecnologia da UNICAMP, ex-professores meus de graduação, que prontamente escreveram as cartas de recomendações para eu ingressar no Mestrado.

Aos colegas do projeto e-phenology, em especial, à Bruna, Gabriela, Natalia, Jurandy e Jefersson pela paciência e colaboração.

Aos colegas do laboratório RECOD do Instituto de Computação. Obrigada pelos momentos de risadas, pelos conselhos e pela ajuda com o latex e outras dúvidas.

Aos meus amigos de Piracicaba, que nas poucas oportunidades que tivemos nestes dois anos, pude descontraír em alguns dos momentos mais difíceis nesta jornada.

Ao pessoal que durante os seis primeiros meses me acompanharam na van de Piracicaba até Campinas, fazendo as longas horas de viagem menos cansativas e divertidas.

À Eliza Ribeiro por concordar em dividir casa comigo em Campinas e por aturar os momentos de tensão e crise, e por ter se mostrado sempre uma ótima amiga.

Ao Fernando Novello pelo companheirismo, compreensão e paciência.

E, por fim, agradeço o apoio financeiro recebido das agências de fomento CAPES, CNPq e FAPESP-MICROSOFT RESEARCH financiadora do Projeto e-phenology (processo 2010/52113-5) e bolsa de mestrado (processo 2011/51523-8). Muito obrigada!



# Sumário

<b>Abstract</b>	<b>ix</b>
<b>Resumo</b>	<b>xi</b>
<b>Dedicatória</b>	<b>xiii</b>
<b>Agradecimentos</b>	<b>xv</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Motivação . . . . .	3
1.2 Objetivos e Contribuições . . . . .	4
1.3 Organização do Trabalho . . . . .	4
<b>2 Trabalhos Relacionados e Conceitos Básicos</b>	<b>5</b>
2.1 Visão geral do projeto e-phenology . . . . .	5
2.2 Arquitetura . . . . .	8
2.3 Coleta de Dados . . . . .	10
2.4 Sistemas de Informação para dados de Biodiversidade . . . . .	11
2.5 Metadados em Biodiversidade . . . . .	14
<b>3 Modelo de Dados</b>	<b>17</b>
3.1 Visão Geral do Modelo . . . . .	17
3.2 Módulos de Dados . . . . .	21
3.2.1 Módulo Dados de Localização . . . . .	22
3.2.2 Módulo Dados Taxonômicos . . . . .	25
3.2.3 Módulo Dados Ecológicos . . . . .	27
3.2.4 Módulo Dados Climatológicos . . . . .	29
3.2.5 Módulo Dados Fenológicos . . . . .	30
3.2.6 Módulo Imagens e Metadados . . . . .	31
3.3 Exemplo de informações armazenadas . . . . .	34



3.4	Exemplo de consultas . . . . .	39
<b>4</b>	<b>Aspectos de Implementação</b>	<b>43</b>
4.1	Banco de Dados . . . . .	43
4.2	Inclusão dos Dados do Projeto e-phenology . . . . .	44
4.3	Consultas . . . . .	47
4.3.1	Consultas Simples . . . . .	47
4.3.2	Consultas Complexas . . . . .	50
4.4	Aplicação <i>Web</i> . . . . .	56
<b>5</b>	<b>Conclusões e extensões</b>	<b>67</b>
	<b>Referências Bibliográficas</b>	<b>71</b>
<b>A</b>	<b>Ficha de Coleta de Dados no Campo - Cerrado (Borda Leste)</b>	<b>77</b>
<b>B</b>	<b>Descrição de Dados e-phenology</b>	<b>79</b>
B.1	Módulo de Dados Localização . . . . .	79
B.2	Módulo de Dados Taxonômicos . . . . .	82
B.3	Módulo de Dados Ecológicos . . . . .	84
B.4	Módulo de Dados Climáticos . . . . .	89
B.5	Módulo de Dados Fenológicos . . . . .	90
B.6	Módulo de Imagens e Metadados . . . . .	91
<b>C</b>	<b>Planilha com Dados Climáticos</b>	<b>95</b>



# Lista de Tabelas

3.1	Informação de localização: Estados. . . . .	35
3.2	Informação de localização: Cidades. . . . .	35
3.3	Informação de localização: Áreas. . . . .	36
3.4	Informação de localização: Ambientes de Amostragem de coleta. . . . .	36
3.5	Informação de localização: Áreas de Ambientes de Amostragem de coleta. . . . .	36
3.6	Informação de localização: Locais (posição que o indivíduo se encontra efetivamente). . . . .	36
3.7	Informação de taxonomia: Espécies. . . . .	37
3.8	Informação de taxonomia: Indivíduos. . . . .	38
3.9	Informação de dados ecológicos: Síndromes de Polinização de Espécies. . . . .	38
3.10	Informação de dados ecológicos: Síndromes de Dispersão de Espécies. . . . .	38
3.11	Informação de dados fenológicos: Dados de Fenologia. . . . .	39
3.12	Informação de dados de imagens: Imagens. . . . .	39
3.13	Informação de dados de imagens: Imagens de Máscaras de Espécies. . . . .	39



# Lista de Figuras

2.1	Torre e câmara instaladas na área de Cerrado em Itirapina-SP. . . . .	6
2.2	Exemplo de imagem típica da câmara de Itirapina-SP. . . . .	7
2.3	Arquitetura básica do sistema e-phenology. . . . .	9
2.4	Diagrama de atividades para o processo de coleta de dados de fenologia. . .	12
3.1	Diagrama Entidade-Relacionamento proposto para o projeto e-phenology. .	20
3.2	Diagrama de pacotes – visão geral do banco de dados do projeto e-phenology.	21
3.3	Diagrama dos dados de Localização. . . . .	23
3.4	Visão da área de estudo de fenologia vegetal com seus respectivos ambientes, Itirapina-SP. . . . .	24
3.5	Diagrama do Módulo de Dados Taxonômicos. . . . .	26
3.6	Diagrama do Módulo de Dados Ecológicos. . . . .	28
3.7	Diagrama do Módulo de Dados Climatológicos. . . . .	29
3.8	Diagrama do Módulo de Dados Fenológicos. . . . .	31
3.9	Diagrama do Módulo de Imagens e Metadados. . . . .	32
3.10	Extração de características das imagens do Cerrado: (a) Visão da imagem tirada do Cerrado, (b) Máscara da espécie <i>Caryocar brasiliensis</i> aplicada à imagem, (c) Resultado do processamento da máscara da espécie <i>Caryocar Brasiliensis</i> . Apresenta a média dos canais R ( <i>Red</i> – Vermelho), G ( <i>Green</i> – Verde) e B ( <i>Blue</i> – Azul) no decorrer dos dias [4]. . . . .	34
4.1	Exemplo de planilha típica usada para armazenar dados fenológicos. . . . .	44
4.2	Exemplo de dados de indivíduos mortos representados por uma cruz e as colunas posteriores vazias, e a troca destes símbolos pela letra ‘M’ de ‘Morto’. . . . .	46
4.3	Exemplo de consulta simples: obter espécies cadastradas no banco de dados. SQL usando <i>alias</i> . . . . .	48
4.4	Exemplo de consulta simples: obter total de indivíduos cadastrados no banco de dados. SQL usando <i>count</i> . . . . .	48
4.5	Exemplo de consulta simples: obter observador responsável pela fenologia do indivíduo 5, na data de 25/09/2005. SQL usando <i>distinct</i> . . . . .	49



4.6	Exemplo de consulta simples: obter observador responsável pela fenologia do indivíduo 5, na data de 25/09/2005. SQL usando <i>distinct</i> . . . . .	50
4.7	Exemplo de consulta complexa: obter dados de localização e taxonômicos dos indivíduos. SQL usando <i>join</i> . . . . .	51
4.8	Exemplo de consulta complexa: obter dados de localização e taxonômicos dos indivíduos. SQL usando <i>join</i> . . . . .	52
4.9	Exemplo de consulta complexa: obter dados de localização e taxonômicos dos indivíduos. SQL usando <i>join</i> . . . . .	53
4.10	Tela Início da aplicação. . . . .	57
4.11	Interface de consulta de espécies. Relaciona todas as espécies cadastradas. . . . .	57
4.12	Interface de consulta de espécies com filtro. Aplicando a busca pelo nome binomial da espécie que iniciam com ‘g’. . . . .	58
4.13	Interface de cadastro de espécies. . . . .	58
4.14	Interface de cadastro de espécies. Exemplo para adicionar dados de flor da espécie. . . . .	59
4.15	Interface de consulta de indivíduos. Relaciona todos os indivíduos cadastrados. . . . .	60
4.16	Interface de consulta de indivíduos com filtro. Aplicando a busca para todos os indivíduos que pertencem à espécies com o nome binomial iniciado por ‘M’. . . . .	60
4.17	Interface de consulta de indivíduos que permite a visualização da fenologia de um indivíduo selecionado. . . . .	61
4.18	Interface de consulta de fenologia de um indivíduo selecionado. No exemplo, o indivíduo selecionado é o indivíduo que tem o número igual a 2 e pertence a espécie <i>Myrcia guianensis</i> . . . . .	61
4.19	Interface de resultado da consulta de fenologia de um indivíduo selecionado, apresentando apenas valores de fenologia. . . . .	62
4.20	Interface de consulta de fenologia com médias climáticas para um indivíduo selecionado. No exemplo, o indivíduo selecionado é o indivíduo que tem o número igual a 2 e pertence a espécie <i>Myrcia guianensis</i> . . . . .	62
4.21	Interface de consulta de fenologia com médias climáticas para um indivíduo selecionado. No exemplo, o indivíduo selecionado é o indivíduo que tem o número igual a 2 e pertence a espécie <i>Myrcia guianensis</i> , e apenas as fenofases botão e brotamento e o sensor temperatura foram selecionados . . . . .	63
4.22	Interface de resultado da consulta de fenologia com médias climáticas para um indivíduo selecionado. Apresenta apenas os valores das fenofases botão e brotamento e o sensor temperaturas selecionados. . . . .	64



4.23	Interface de filtros para realizar consultas de fenologia, por espécie, indivíduo ou localização . . . . .	65
A.1	Exemplo de ficha de fenologia em campo. . . . .	78
C.1	Exemplo de planilha com dados climáticos coletados pelo sensores da Torre de Itirapina - parte 1. Campos: Data e hora e Sensores com valores coletados.	96
C.2	Exemplo de planilha com dados climáticos coletados pelo sensores da Torre de Itirapina - parte 2. Campos: Data e hora e Sensores com valores coletados.	97



# Capítulo 1

## Introdução

As mudanças ambientais tornaram-se uma questão importante na agenda global [41]. A fim de apoiar a formulação de políticas para gestão ambiental e manter o equilíbrio de ecossistemas, é necessário ter uma visão precisa das condições existentes e compreender as complexas mudanças que ocorrem em todos os níveis do planeta [41].

Um passo essencial para a criação de cenários apropriados é coletar dados relevantes sobre o meio ambiente e desenvolver sistemas computacionais para gerenciar e descobrir conhecimento a partir desses dados. Esses sistemas devem, além disso, combinar os dados recém coletados com informações históricas e legadas (por exemplo, a partir de observações de indivíduos em campo ao longo dos anos), com uma gestão unificada. Por isso, cientistas envolvidos com questões ambientais usualmente buscam apoio de sistemas computacionais. Isso, naturalmente, apresenta vários desafios relacionados a problemas de interoperabilidade, devido à incompatibilidade de sistemas, diversidade e heterogeneidade de dados, e a uma variedade de perfis de usuário.

Um exemplo representativo desses problemas surge no contexto dos estudos de fenologia. Fenologia, o estudo de fenômenos naturais recorrentes e sua relação com o clima [46], é uma ciência tradicional de observar os ciclos das plantas e dos animais e relacioná-los com os dados meteorológicos locais, bem como interações bióticas e filogenéticas [42]. Recentemente, fenologia tem ganho importância como o indicador mais simples e confiável dos efeitos das mudanças climáticas sobre plantas e animais [39, 45, 50, 51].

No entanto, existe uma falta de informações e sistemas de monitoramento de regiões tropicais, em particular, na América do Sul, o que vêm estimulando diversos centros de pesquisa [33] a desenvolver trabalhos visando preencher esta lacuna. Dentre os vários trabalhos relacionados a fenologia, destaca-se o Projeto e-phenology<sup>1</sup>. O e-phenology visa incorporar resultados recentes de pesquisas correntes sobre fenologia tropical no Brasil [1] com as novas tecnologias para detectar mudanças ambientais e as novas técnicas de ob-

---

<sup>1</sup><http://www.recod.ic.unicamp.br/ephenology>. Data de acesso: 7 de janeiro de 2014.

servação remota da fenologia, como o uso de câmeras digitais para integrar os campos e compreender os efeitos do aquecimento global. Desta forma, o projeto propõe [34, 44] a longo prazo:

1. a integração da fenologia às novas tecnologias de monitoramento ambiental, sistemas de monitoramento remoto de fenologia e até de observação remota por câmeras digitais, aumentando a precisão em definir a relação entre a fenologia e o clima;
2. o estímulo de programas de acompanhamento de fenologia de mais longo prazo e a integração de diversas disciplinas, avançando o conhecimento de respostas sazonais dentro dos trópicos para mudanças climáticas de longo prazo;
3. a criação de modelos, métodos e algoritmos de apoio à gestão, integração e análise de dados de sistemas de monitoramento remoto de fenologia;
4. a implementação de pesquisas experimentais com foco na detecção dos efeitos da mudança climática sobre a vegetação;
5. a recuperação de séries de dados históricas de fenologia a partir de coleções de herbários e de outras fontes;
6. a construção de uma Rede Brasileira de Fenologia baseada na seleção de espécies distribuídas no âmbito nacional e na fenologia remota.

Para tanto, o projeto e-phenology envolve profissionais da computação e da biologia. Trata-se de colaboração estabelecida entre pesquisadores do laboratório *Reasoning for Complex Data* (RECOD) do Instituto de Computação (IC) da Universidade Estadual de Campinas (UNICAMP) e do Laboratório de Fenologia do Departamento de Botânica do Instituto de Biociências (IB) da Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP) de Rio Claro, São Paulo. A primeira etapa de desenvolvimento do projeto e-phenology foi financiada pelo convênio FAPESP-Microsoft (proc. 2010/52113-5) e terminou em julho de 2013. A segunda etapa, também financiada pelo convênio FAPESP-Microsoft (proc. 13/50155-0), teve início em agosto de 2013 e tem término previsto para julho de 2017.

Embora a computação seja uma grande aliada dos peritos de fenologia por fornecer-lhes ferramentas que auxiliam na análise, mineração e conclusão de relatórios sobre espécies e seus comportamentos ao longo dos anos, no âmbito dos estudos de fenologia, há ainda alguns desafios que precisam ser superados para possibilitar que os objetivos propostos pelo projeto e-phenology sejam alcançados. Estes desafios estão associados à especificação e implementação de sistemas de informação para apoiar estudos de fenologia e envolvem combinação de pesquisas em processamento de imagens, banco de dados e visualização de informação. Destacam-se:

1. lidar com grandes volumes de informação ao longo dos anos;
2. integrar informações de diferentes fontes e formatos (heterogeneidade);
3. manipular tipos especiais de dados como temporais e de imagem
4. manipular informação geoespacial de referência.

Neste sentido, este trabalho apresenta a especificação e implementação de um banco de dados capaz de armazenar e gerir dados no contexto do projeto e-phenology. A motivação, os objetivos e contribuições são apresentados nas seções a seguir.

## 1.1 Motivação

Os estudos relacionados à fenologia de plantas possibilitam que pesquisadores possam realizar inferências sobre as condições das plantas em relação ao seu ambiente e às variações climáticas [43]. Por isso, os estudos fenológicos são cada vez mais considerados essenciais para o estudo da ecologia, evolução e da compreensão da complexa dinâmica dos ecossistemas florestais [33]. As mudanças climáticas afetam a fenologia de diversas espécies, e isso provoca amplas consequências à biodiversidade [33, 34].

Para compreender a fenologia e sua influência nas alterações climáticas, esses estudos envolvem pesquisas periódicas, com dados temporais, que variam ao longo de dias, meses, anos e décadas. São estudos que dependem relativamente das observações feitas por diferentes pessoas, e por isso, enfrentam sérios problemas quanto aos métodos que são comumente utilizados na análise de seus resultados. Métodos científicos clássicos (métodos estatísticos), por exemplo, podem ser sensíveis para valores extremos, especialmente quando poucos anos são analisados [16]. Essas limitações rendem interpretações e comparações complicadas e, portanto, há a necessidade de se desenvolver e aperfeiçoar métodos que permitam a detecção de alterações em séries fenológicas, estabelecendo-se padrões.

O principal foco do projeto e-phenology é justamente criar um padrão de pesquisa a partir da observação remota, usando para isto câmeras digitais. Sua proposta é integrar as informações de pesquisas atuais com novas tecnologias para detectar alterações ambientais e usar câmeras digitais para realizar observações remotas, de maneira que facilite o processo de observar e coletar as mudanças da fenologia dos indivíduos.

Atualmente, os dados dos estudos de fenologia realizados em campo pelo laboratório de fenologia da UNESP são armazenados em planilhas eletrônicas. Essas planilhas possuem informações mensais de coletas realizadas desde dezembro de 2004. Devido ao grande volume de informações, processos simples como consultar ciclos de indivíduos de uma

dada espécie em uma determinada época, fazer levantamentos estatísticos de duração e ocorrência de ciclos, manter a consistência temporal dos dados e a qualidade destes dados, são trabalhosos e susceptíveis a erros.

## 1.2 Objetivos e Contribuições

Para compreender e especificar melhor o banco de dados no contexto do projeto e-phenology, os dados foram divididos em seis categorias: dados de localização; dados taxonômicos; dados ecológicos; dados climáticos; dados de fenologia e dados de imagens obtidas à partir de observações remotas com câmeras digitais. Com o conhecimento adquirido com a proposta do projeto e-phenology e com os requisitos do banco de dados destacam-se as principais contribuições deste trabalho:

- Levantamento e organização dos requisitos fornecidos pelos usuários-alvos (biólogos), do Laboratório de Fenologia da UNESP, responsáveis pelo projeto e-phenology;
- Especificação e implementação do banco de dados para apoiar estudos de fenologia, integrando informações de campo, e dados de imagens e dados de clima;
- Especificação de consultas de interesse dos biólogos;
- Implementação parcial das consultas em uma ferramenta web visando apoiar a manipulação do banco de dados com dados reais de pesquisas em fenologia realizadas por usuários-alvo.

## 1.3 Organização do Trabalho

Este documento está organizado como segue. O Capítulo 2 apresenta os trabalhos relacionados e os conceitos básicos sobre o projeto e-phenology e este trabalho. O Capítulo 3 descreve o modelo de dados desenvolvido para a criação do banco de dados de fenologia. O Capítulo 4 apresenta aspectos de implementação do banco de dados, com destaque para algumas das consultas implementadas e o Capítulo 5 apresenta as conclusões do trabalho e possíveis extensões.

# Capítulo 2

## Trabalhos Relacionados e Conceitos Básicos

Para desenvolver este trabalho, foi preciso efetuar pesquisa e obter conhecimentos em várias áreas. A seção 2.1 apresenta uma Visão Geral do Projeto e-phenology e conceitos relacionados. A seção 2.2 apresenta a Arquitetura do Projeto e-phenology. A seção 2.3 trata da Coleta de Dados no contexto do e-phenology. A seção 2.4 apresenta alguns Sistemas de Informação para dados de Biodiversidade. E a seção 2.5 apresenta alguns trabalhos relacionados ao armazenamento de metadados em biodiversidade.

### 2.1 Visão geral do projeto e-phenology

O e-phenology é um projeto multidisciplinar que combina pesquisas nas áreas de computação e fenologia. Seu foco principal é resolver os problemas teóricos e práticos envolvidos na utilização de novas tecnologias para observação fenológica remota, visando detectar mudanças ambientais locais e compreender os efeitos do aquecimento global nos trópicos. Para tanto, pretende:

- utilizar novas tecnologias de monitoramento ambiental com base em sistemas de monitoramento remoto de fenologia;
- criar um protocolo para um programa de monitoramento de fenologia de longo prazo no Brasil e para a integração entre diversas disciplinas, avançando o conhecimento atual de respostas sazonais dentro dos trópicos para mudanças climáticas;
- fornecer modelos, métodos e algoritmos de apoio à gestão, integração e análise de dados dos sistemas de fenologia remota.

O projeto pretende incorporar resultados recentes de pesquisas em campo (estudos de fenologia tropical) com as novas técnicas de observação fenológica remota, usando câmeras digitais.

As câmeras digitais para observação remota da fenologia são posicionadas em torres de observação em áreas de estudos de interesse dos pesquisadores de fenologia. O objetivo de instalar diversas câmeras e torres em diferentes áreas de estudo é tentar estabelecer uma rede de torres para integrar os dados de observações remotas de fenologia e inferir qual a relação das mudanças na fenologia com o clima em âmbitos locais e regionais. Para tanto, as câmeras são programadas para tirar fotos da área onde se encontram ao longo do dia em um intervalo de tempo programado, durante todo o ano.

A primeira torre, relativo ao primeiro caso de estudo do projeto, é apresentada na Figura 2.1. A torre está instalada em uma área de Cerrado, situada na cidade de Itirapina, Estado de São Paulo, Brasil. A torre tem 18 metros de altura e possui um braço posicionado para o lado nordeste, onde a câmera digital está instalada.



Figura 2.1: Torre e câmera instaladas na área de Cerrado em Itirapina-SP.

No caso de Itirapina, a câmera instalada foi programada para tirar cinco fotos em cada hora cheia do dia, do período das seis da manhã às seis da tarde. Estas fotos são

semelhantes à imagem apresentada na Figura 2.2 e, devido à variedade de fotos no decorrer do dia, tem sido possível processá-las para obter a variação temporal de características das plantas do Cerrado tais como cor e textura [3, 4]. É a partir do processamento destas imagens e obtenção destas características que espera-se conseguir comparar dados de coleta de fenologia do chão com os dados das imagens e do clima.

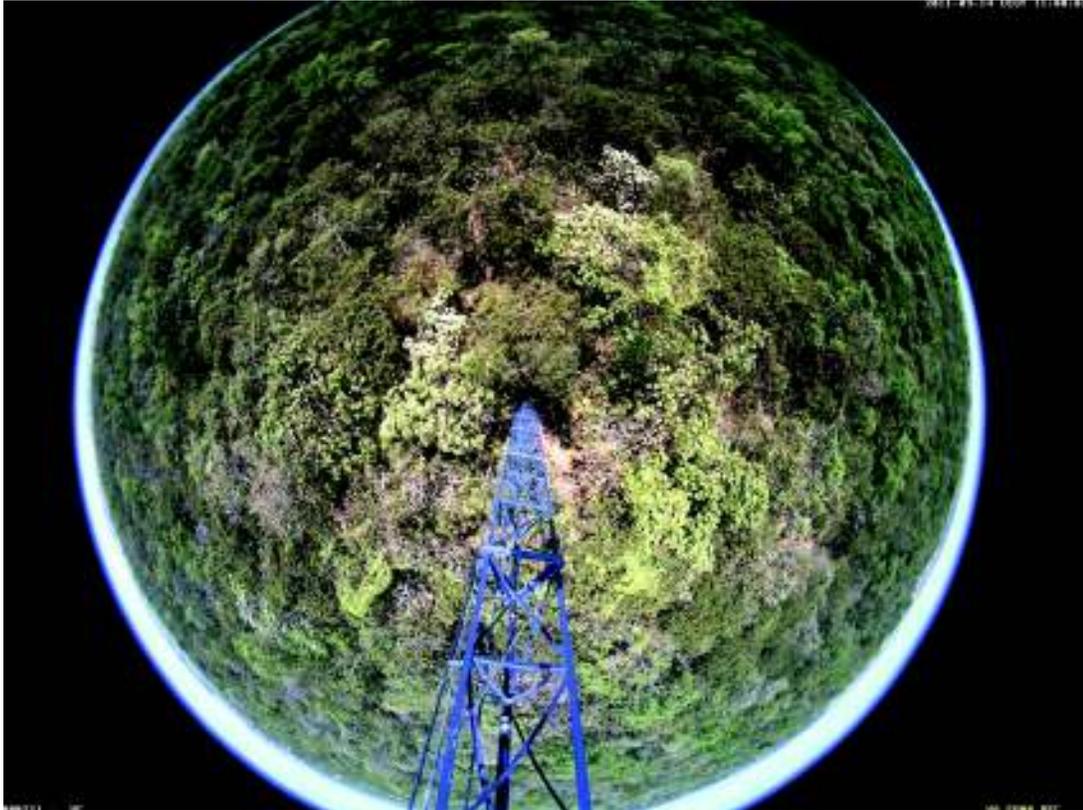


Figura 2.2: Exemplo de imagem típica da câmera de Itirapina-SP.

Além da câmera, uma estação climática local foi instalada e configurada na mesma torre. Na estação, foram instalados nove sensores ambientais com o objetivo de obter dados climáticos locais do Cerrado. Tais sensores medem, por exemplo, a umidade do solo, a quantidade de chuva, a velocidade do vento e a radiação solar. Os dados obtidos a partir destes sensores contribuirão tanto para melhor avaliar o impacto das mudanças climáticas sobre a fenologia das plantas quanto para comparar com as mudanças identificadas ao se processar as imagens na vegetação do Cerrado.

Assim, os resultados deste projeto dependem de atividades de desenvolvimento de modelos, ferramentas e técnicas relacionadas: a) ao processamento de imagens, como extração e indexação de descritores de conteúdo associados aos diferentes tipos de vegetação; b) à gestão dos dados científicos dos estudos de fenologia, incluindo mineração de dados,

processamento de séries temporais, e anotação de dados; c) à criação de modelos e metodologias para análise de mudanças climáticas com base na exploração de novos índices para avaliar as mudanças fenológicas.

## 2.2 Arquitetura

A Figura 2.3 mostra a arquitetura simplificada do sistema e-phenology e destaca, em vermelho, os módulos envolvidos no desenvolvimento deste trabalho. Três funcionalidades devem ser apoiadas por essa arquitetura: a inserção de dados, o processamento de consultas e a visualização dos resultados.

A inserção de dados considera três fontes distintas: a) os usuários podem incluir apenas dados textuais, como por exemplo, dados que localizam, definem ou caracterizam espécies e indivíduos ou ainda dados de observações de fenologia realizados em campo que estão anotados em papel a partir de um sistema com interface *Web*; b) dados textuais coletados por meio de sensores climáticos e anotados em arquivos eletrônicos incluídos por mecanismos automatizados e c) dados obtidos a partir do processamento de imagens.

No caso de inclusão de dados a partir do processamento de imagens, a inserção de dados envolve o módulo de processamento de consultas, onde as características textuais e visuais apropriadas das imagens são extraídas para serem armazenadas no banco de dados. Esse processo é normalmente realizado *offline* [11].

Comparativamente, o processamento de consultas pode ser feito sobre duas perspectivas: a) consultas textuais e b) consultas a partir do processamento das imagens ou dos dados relacionados. Quaisquer que sejam os interesses dos usuários em realizar as consultas, a organização do processamento de consultas é a mesma: primeiro o usuário interage com uma interface, a partir da qual ele especifica um padrão de consulta desejado e, conforme o padrão escolhido (texto ou imagem) o módulo de processamento de consultas irá realizar as tarefas necessárias para buscar as informações no banco de dados. O resultado gerado pela consulta, então, pode ser visualizado.

De modo parecido ao que foi descrito para incluir dados relacionados às imagens, o processo de consultas envolve o módulo de consultas que extrai um vetor de características do padrão definido pelo usuário e aplica uma função de distância para avaliar a sua similaridade em relação aos dados do banco. A seguir, ordenam-se esses resultados em ordem decrescente de similaridade em relação ao padrão de consulta especificado e enviam-se os resultados mais relevantes para o módulo de interface [11] onde pode ser visualizado.

A Figura 2.3 destaca também os fluxos que se espera que ferramentas computacionais realizem para apoiar estudos de fenologia usando a arquitetura apresentada neste trabalho. As setas tracejadas azuis e identificadas pela letra A representam o fluxo para a inclusão, processamento das consultas e visualização dos resultados de dados textuais.

Neste caso, o módulo de processamento é responsável apenas por fazer que a consulta seja executada no banco e os resultados retornem para a visualização do usuário. Por sua vez, as setas tracejadas pretas e identificadas pela letra B referem-se ao fluxo para a inclusão, processamento de consultas e visualização dos resultados de dados de imagens.

O banco de dados especificado e implementado neste trabalho foi projetado inicialmente para atender duas das três funcionalidades: a inclusão de dados e o processamento de consultas. A etapa de visualização será melhor trabalhada e delineada em trabalhos futuros, nos quais novas técnicas de visualização de informação serão aplicadas e utilizadas.

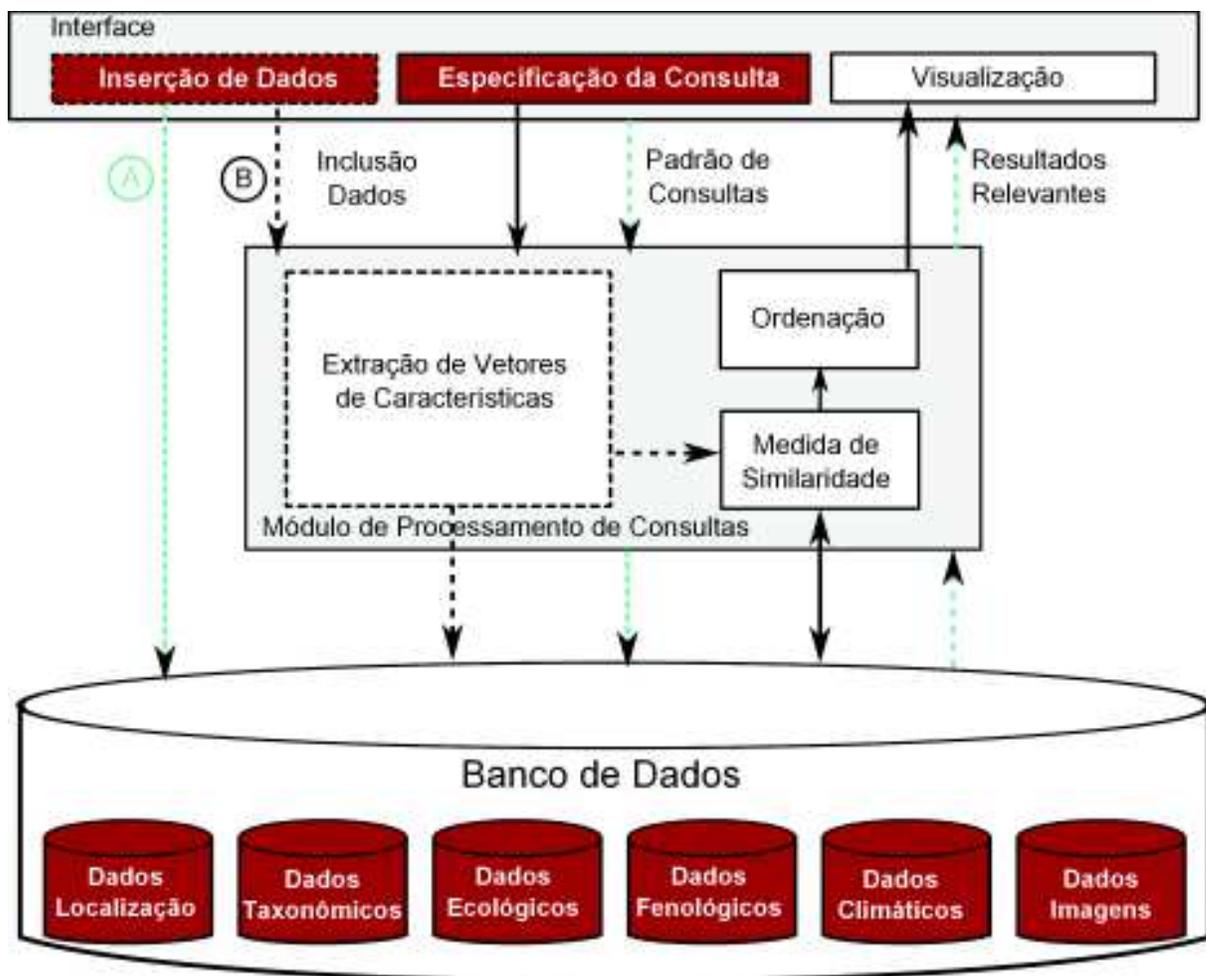


Figura 2.3: Arquitetura básica do sistema e-phenology.

## 2.3 Coleta de Dados

Como explicado na seção 2.1, o primeiro caso de estudo do projeto e-phenology, refere-se a uma área de Cerrado, situada na região de Itirapina, interior do Estado de São Paulo. Os dados de pesquisas de fenologia utilizados para a construção do modelo do banco de dados, assim como as imagens, são oriundos desta área. A coleta de fenologia em Itirapina é realizada pelos biólogos do Laboratório de Fenologia da UNESP desde setembro de 2004 e trata-se da observação da ocorrência de fenômenos biológicos do ciclo de vida da planta, chamados de fenofases.

As fenofases observadas dizem respeito a floração, frutificação, brotamento e queda foliar. A produção de flores é dividida em botões e floração propriamente dita ou período de antese (abertura de flores) e a produção de frutos em períodos de frutos maduros ou frutos verdes [48]. A Figura 2.4 apresenta as etapas para o processo de coleta desses dados.

Todo mês, um grupo do laboratório – com no mínimo dois observadores – vai a campo para realizar a observação dessas fenofases. Cada observador do grupo leva consigo uma prancheta com fichas impressas semelhantes à apresentada pelo Apêndice A. Cada uma dessas fichas possuem campos que relacionam a localização do indivíduo (campos *Local* e *Transecto*), o número de identificação do indivíduo naquela localização (campo *Indivíduo*), a família (campo *Família*), a espécie (campo *Espécie*) e as fenofases botão, antese, fruto verde, fruto maduro, brotamento e queda foliar (campos *Bot*, *Ante*, *FV*, *FM*, *Brot*, *Qued*, respectivamente).

Para cada indivíduo listado na ficha, os observadores utilizam um método de avaliação para quantificar a fenofase. No caso, um método de avaliação direto quantitativo – em que os observadores podem mensurar em uma escala ordinal pré-definida a ocorrência da fenofase [15] – é o utilizado nos estudos de Itirapina. As fenofases são quantificadas, então, de acordo com três classes de intensidade, em que: (i) 0 representa que não há indícios da fenofase para aquele indivíduo (ausência da fenofase), (ii) 1 representa que a fenofase está presente em menor intensidade (intensidade moderada – até 50%) e (iii) 2 representa que a fenofase está presente em maior intensidade (de 51% a 100%) [36, 37, 38].

Terminadas as observações em toda a área do Cerrado e feitas as anotações nas fichas, a próxima etapa consiste em passar essas informações para planilhas eletrônicas semelhantes às fichas do campo. Esta etapa é realizada no Laboratório de Fenologia (UNESP, Rio Claro). Completada essa tarefa, os dados então são divididos em planilhas eletrônicas por fenofase, na qual serão feitas a interpretação dos ciclos e as análises estatísticas desses ciclos.

Além de facilitar a interpretação e análises dos ciclos, um outro fator importante para que essa divisão dos dados em planilhas por fenofase seja feita é que uma fenofase depende

da outra, por exemplo, não pode haver antese (flor) sem ter ocorrido a fenofase botão no mês anterior. Assim, se uma observação no mês corrente indica que o indivíduo está na fenofase antese (classe 1 de intensidade), ele deve ter tido botão no mês anterior (classes 1 ou 2 de intensidade), caso a fenofase botão esteja 0, considera-se que há um problema de observação, que pode ter sido do observador ou por ele não ter encontrado o indivíduo, ou por ele não ter conseguido visualizar estes botões; ou ainda por não ter sido detectada no intervalo de observação das coletas. Com base na observação atual, o valor do mês anterior é corrigido e a planilha é armazenada.

Cada uma dessas planilhas contém cerca de 2122 indivíduos relacionados, na mesma ordem e com a mesma identificação – um número sequencial utilizado para facilitar o controle, a manipulação e a comparação dos valores. Esse processo de coleta se iniciou em 2004, e desde então, a planilha é atualizada mensalmente com os novos dados de observações, e contém atualmente cerca de um milhão de registros.

## 2.4 Sistemas de Informação para dados de Biodiversidade

Sistemas de Informação em Biodiversidade (SIB) são sistemas computacionais desenvolvidos para gerenciar, manipular e correlacionar diversos conjuntos de informações ambientais e envolvem banco de dados volumosos que armazenam registros a respeito de espécies, coleta de observações de campo, dados experimentais, entre outros [41]. Com tais sistemas, os pesquisadores em biodiversidade podem realizar análises, inferências e planejar ações sobre os impactos causados pelas mudanças ambientais nos mais diferentes aspectos da biodiversidade, como mudanças no ecossistema, ciclo de vida das plantas, dos animais, dentre outros.

Dentre os diversos projetos que envolvem SIBs, existem vários destinados a desenvolver meios para gerenciar dados que estão relacionados à pesquisa em biodiversidade. Um destes projetos é o speciesLink <sup>1</sup>. O speciesLink foi criado em outubro de 2001 e tem como objetivo integrar a informação catalogada sobre coleções biológicas e observações documentadas de organismos vivos na natureza, disponíveis em museus, herbários e coleções microbiológicas. Para tanto, o speciesLink dispõe de protocolos de comunicação para obter dados de diferentes tipos e fontes de banco de dados, além de contar com uma aplicação *Web* para acessar estes dados.

Outro exemplo, é o Specify <sup>2</sup>, cujo objetivo é fornecer uma plataforma computacional para o gerenciamento de diversas coleções de dados, incluindo descrição geográfica da

---

<sup>1</sup><http://splink.cria.org.br/>. Data de acesso: 16/02/2012.

<sup>2</sup><http://www.specifysoftware.org/>. Data de acesso: 16/02/2012.

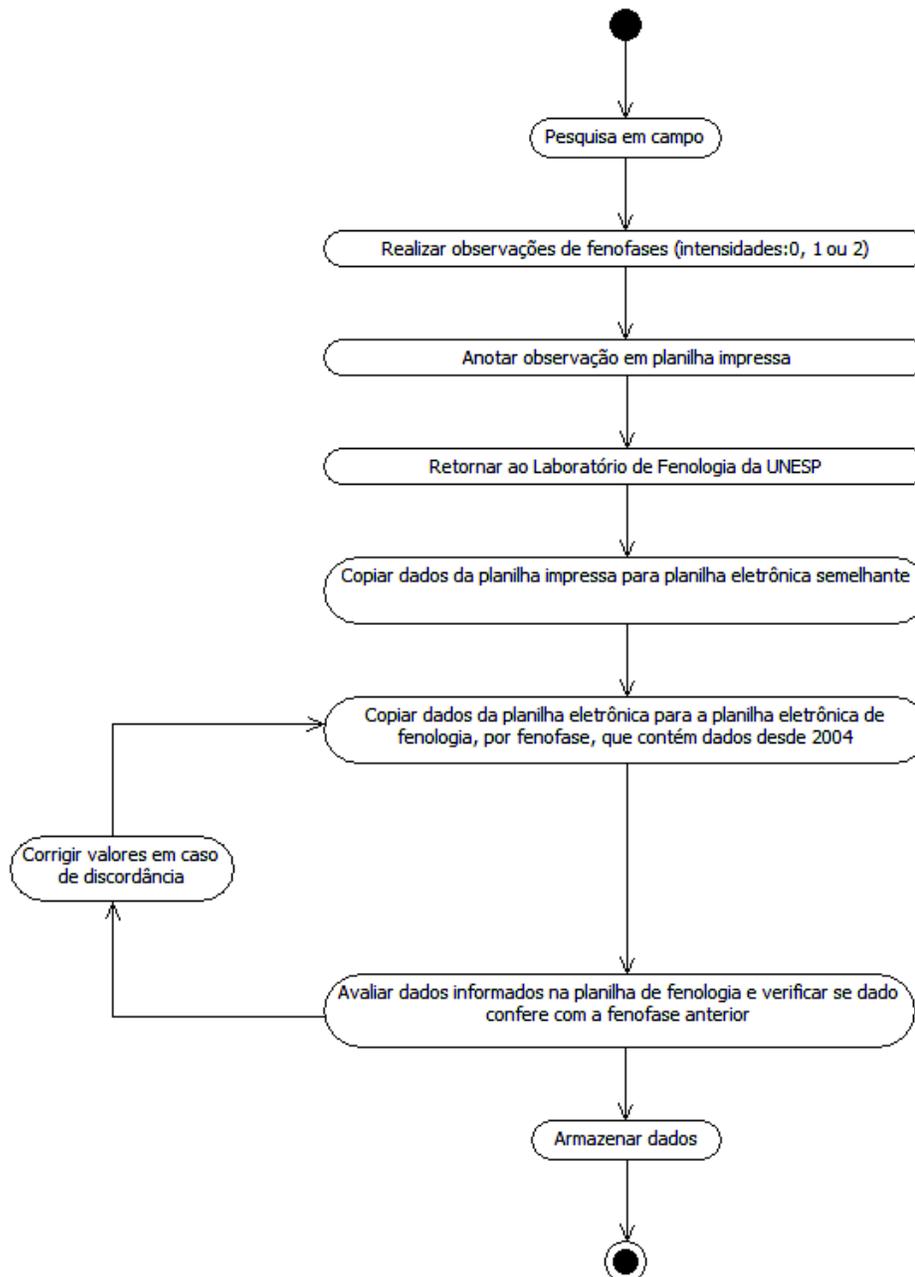


Figura 2.4: Diagrama de atividades para o processo de coleta de dados de fenologia.

coleta, dados dos coletores e algumas operações que devem ser realizadas sobre o acervo, como empréstimos, intercâmbios, adesões e doações. O banco de dados do Specify foi desenvolvido usando a ferramenta MySQL e a aplicação de acesso aos dados é do tipo

desktop.

Outros tipos de projetos relacionados ao desenvolvimento de aplicações de sistemas de informação de biodiversidade são aqueles destinados a gerenciar dados de coletas realizadas em campo. Um exemplo é o projeto Biota [6] que foi o primeiro projeto desenvolvido para registrar ocorrências realizadas pelos biólogos no campo e a propor um banco de dados para gerenciar inventários sobre biodiversidade para o projeto ALAS (*Artropodos de La Selva*). Um outro exemplo desse tipo de sistema é o SinBiota<sup>3</sup> que gerencia registros de observações de campo realizadas por grupos de pesquisa financiados pela FAPESP, no estado de São Paulo. Atualmente, o SinBiota encontra-se em reformulação e expansão para atender novas demandas da comunidade científica [32].

Há ainda projetos destinados a lidar com o gerenciamento das hierarquias taxonômicas das espécies, que são a base e caracterizam as aplicações em biodiversidade, pois as árvores taxonômicas (como estas hierarquias são também conhecidas) são utilizadas pelos pesquisadores para classificar as formas de vida [35]. Além disso, essas aplicações lidam com a com a distribuição geográfica das espécies. É o caso dos projetos *The Tree of Life* [26], *Catalogue of Life*<sup>4</sup> e TaiBIF [47] que disponibilizam estes dados na *Web* para consultas.

Outros sistemas como o OBIS (*Ocean Biogeographic Information System*) [10] que trata de informações referentes à biodiversidade marinha; o GBIF (*Global Biodiversity Information Facility*)<sup>5</sup> que cobre aspectos de biodiversidade em um contexto global, com banco de dados distribuídos e interoperáveis; o SEEK (*Science Environment for Ecological Knowledge*) [29] e o CNBIQS (*China National Biodiversity Information Query System*) [53] que auxiliam os usuários no processo de consulta as diversas fontes de dados sobre biodiversidade, são tipos de aplicações de sistemas que atuam na área de biodiversidade encontrados no meio acadêmico.

Além dos projetos apresentados, há ainda outros que tentam definir aplicações e padrões para a integração e a interoperabilidade de dados das coleções biológicas, como é o caso do ITIS (*Integrated Taxonomic Information System*)<sup>6</sup>, Species 2000<sup>7</sup>; TDWG (*Taxonomic Database Working Group*)<sup>8</sup> e NBII (*National Biological Information Infrastructure*)<sup>9</sup>.

Ainda há outros trabalhos relacionados, que são aqueles desenvolvidos no Instituto de Computação da UNICAMP. O WebBios<sup>10</sup>, por exemplo, foi desenvolvido em parceria com o Instituto de Biologia da UNICAMP. Este projeto contém trabalhos relacionados a pes-

<sup>3</sup><http://sinbiota1.biota.org.br/>. Data de acesso: 7 de janeiro de 2014.

<sup>4</sup><http://www.catalogueoflife.org/>. Data de acesso: 16/02/2012.

<sup>5</sup><http://data.gbif.org/>. Data de acesso: 16/02/2012.

<sup>6</sup><http://www.itis.gov/>. Data de acesso: 16/02/2012.

<sup>7</sup><http://www.sp2000.org/>. Data de acesso: 16/02/2012.

<sup>8</sup><http://www.tdwg.org/>. Data de acesso: 16/02/2012.

<sup>9</sup><http://www.nbii.gov/>. Data de acesso: 16/02/2012.

<sup>10</sup><http://www.lis.ic.unicamp.br/projects/webios>. Data de acesso: 16/03/2012.

quisas em consultas a dados de biodiversidade e questões ambientais. A primeira pesquisa relacionada a este projeto que pode ser citada, é o projeto Aondê [13]. Aondê é um Serviço Web de Ontologias, cujo objetivo é prover acesso, manipulação, análise e integração de ontologias auxiliando na solução de problemas de heterogeneidade de dados e associando semântica às operações das aplicações. Outra pesquisa trata do desenvolvimento de uma arquitetura para o processamento de consultas a dados de biodiversidade armazenados em repositórios na Web com o objetivo de permitir que pesquisadores possam especificar consultas que envolvam predicados ecológicos, taxonômicos e geográficos [24].

O BIOCORE <sup>11</sup>, assim como o WebBios, é um projeto desenvolvido no Instituto de Computação, que possui como um de seus módulos o Serviço de Expansão de Consultas [49] relacionado. Esse serviço tem como objetivo permitir que biólogos consultem informações de repositórios variados para encontrar os registros procurados, detectar relações que não estão diretamente representadas e também correlacionar os registros.

Outros projetos são o Simimbu [20] e a criação do Repositório de Coletas [27]. O primeiro trata-se de uma ferramenta criada para apoiar consultas de coletas de espécies combinando buscas textuais com buscas que usam o conteúdo de imagens. O segundo projeto refere-se a criação de um repositório de coletas que unifica em um único modelo registros de acervos e de observações de espécies.

## 2.5 Metadados em Biodiversidade

Metadados são definidos de maneira geral como dados sobre dados que descrevem o conteúdo, o contexto, qualidade, estrutura e os propósitos de um conjunto de dados específico [30]. Em ecologia, metadados são as informações que identificam *quem, o que, onde, quando, por que e como* de um conjunto de dados ecológicos coletado [21]. Paratanto, são utilizados alguns padrões de metadados ambientais, que isoladamente, descrevem um conjunto específico de dados.

Um exemplo é o padrão de metadados criado pelo comitê *Federal Geographic Data Committee* (FGDC) que coordena o compartilhamento e os padrões dos metadados geográficos nos Estados Unidos [7]. O padrão FGDC é muito utilizado para descrever dados geográficos em sistemas de informação em biodiversidade.

Outro exemplo é o padrão *Dublin Core* desenvolvido para buscar e descrever recursos eletrônicos [52]. Este é um padrão muito utilizado para gerenciar dados textuais e realizar a indexação e descoberta de documentos digitais.

Alguns outros trabalhos relacionados são aqueles que definem novos padrões como a hierarquia de classes proposta por Gonçalves [22] e o desenvolvimento de um banco de

---

<sup>11</sup><http://www.lis.ic.unicamp.br/projects/biocore>. Data de Acesso em: 16/03/2012.

dados específico para armazenar e gerenciar metadados em sistemas ambientais [19].

Embora o projeto e-phenology seja um sistema de informação em biodiversidade, a atual modelagem feita para o banco de dados apresentada neste trabalho considera e apresenta dados e metadados combinados nas tabelas. Investigações das vantagens e desvantagens ficam como sugestão para trabalhos futuros.



# Capítulo 3

## Modelo de Dados

Este capítulo descreve o modelo de banco de dados criado para o gerenciamento e armazenamento dos dados de fenologia do Projeto e-phenology. A seção 3.1 apresenta uma visão geral desse modelo. A seção 3.2 define as principais características e relacionamentos de cada módulo apresentado anteriormente na arquitetura do projeto (Figura 2.3). Em seguida, a seção 3.3 mostra alguns exemplos dos dados armazenados, respeitando o modelo apresentado. Finalmente, a seção 3.4 apresenta alguns exemplos de consultas que podem ser realizados a partir do modelo proposto.

### 3.1 Visão Geral do Modelo

Para realizar a modelagem e implementação do banco de dados, foi feito um levantamento de requisitos juntos aos biólogos (usuários-alvo do banco de dados). Estes requisitos foram elicitados a partir de reuniões realizadas periodicamente e com revisões de modelos e discussões com o grupo do laboratório de fenologia da UNESP.

O levantamento de requisitos foi dividido em duas etapas: a primeira ocorreu no período de agosto de 2011 à julho de 2012, envolvendo reuniões periódicas na UNESP e na UNICAMP para especificar o modelo do banco de dados; a segunda etapa foi realizada no período de agosto de 2012 à agosto de 2013 e consistiu em reuniões para especificar e implementar consultas de interesse dos especialistas em fenologia para validar o banco de dados.

As duas fases que envolveram o levantamento de requisitos permitiram dividi-los de acordo com: (i) as dificuldades encontradas no processo tradicional de coleta de dados de fenologia; (ii) dados legados de observações realizadas em campo; (iii) uso de novas tecnologias e inclusão de novos dados; e (iv) consultas e análises dos dados, o que facilitou no processo de modelagem e construção do banco de dados.

O que se observou neste processo de levantamento de requisitos é que uma das princi-

As dificuldades encontradas pelo grupo de fenologia da UNESP na coleta de dados está no processo de observar a fenologia no campo que envolve desde a observação e anotação das informações sobre as fenofases no campo até a disponibilização destas informações para todo o grupo, em meio eletrônico. O problema é o tempo que o dado leva para estar disponível para análise e a alta taxa de erro que este processo pode ocasionar na disponibilização do mesmo, o que faz necessário rever as etapas e modificar novamente (muitas vezes isso significa alterar) seu valor no laboratório.

Outra dificuldade são as pesquisas independentes, isto é, dentro de um mesmo grupo de estudo em fenologia de plantas, os pesquisadores tentam responder a diferentes questões utilizando este estudo e associando-o com outros fatores por eles pesquisados. Por exemplo, um ou mais pesquisadores do grupo podem estar interessados em saber e analisar apenas qual a frequência da ocorrência da fenofase de frutificação em determinados indivíduos, porém outros pesquisadores podem estar interessados em relacionar esta ocorrência da frutificação com outros fatores, como por exemplo, com a contagem de frutos. A diferença de interesses faz com que cada pesquisador registre seus próprios dados em planilhas eletrônicas distintas. Um problema notado com esta ação é que o número de indivíduos nas planilhas geralmente diferem (embora sejam da mesma área) o que indica que informações se perderam ao longo do tempo.

Como explicado na seção 2.3, estudos em fenologia geram dados que variam ao longo de dias, meses ou anos, por isso, normalmente estes estudos geram um volume de dados muito grande ao longo do tempo, e no caso do Laboratório de Fenologia da UNESP, a quantidade deste tipo de dados já ultrapassa um milhão de registros distribuídos em seis planilhas eletrônicas.

Quando análises são feitas com estes dados, dispende-se muito tempo comparando-se as planilhas e também buscando informações complementares sobre os indivíduos em outras fontes. A busca de dados em outras fontes se deve principalmente porque apenas os dados de localização, identificação da espécie e valor da fenofase em uma determinada data estão registrados nas planilhas. Assim, quando há a necessidade de se saber características específicas de uma espécie (como sinônimos, autores, entre outros) é necessário recorrer a estas outras fontes, como a *Web*.

Outro ponto importante sobre as análises de fenologia é que normalmente elas estão associadas a dados climatológicos. Até a concepção do projeto e-phenology, esses dados eram obtidos em estações meteorológicas mais próximas. No caso do Cerrado em Itirapina, a estação mais próxima está a cerca de 4 km de distância [5] e os dados obtidos são regionais e não locais. Com o objetivo de obter estes dados locais, nove sensores de microclima foram instalados no Cerrado por volta de setembro de 2011 e desde então, as coletas são feitas de cinco em cinco minutos durante o dia.

Além dos novos dados de clima e dos dados legados de fenologia, há ainda a pesquisa

pelo monitoramento remoto da fenologia de plantas. O monitoramento remoto depende de uma câmera digital instalada em uma torre mais alta que as copas das árvores da área de estudo. Com as imagens obtidas é possível detectar, por exemplo, mudanças no padrão foliar e validar os dados digitais com as informações coletadas no campo [2]. Com o tempo, mais câmeras serão instaladas em torres localizadas em áreas com vegetações distintas, possibilitando a formação de uma rede de monitoramento remoto de fenologia e análise mais profunda das mudanças e comportamentos das vegetações.

Com o intuito de atender a todos os requisitos apresentados, o banco de dados foi modelado e implementado tanto para agilizar o processo de consultas e análise dos dados temporais de fenologia, quanto possibilitar que novas questões de biodiversidade sejam descobertas e respondidas. A Figura 3.1 apresenta o Diagrama de Entidade-Relacionamento do banco de dados. Atributos foram omitidos para melhorar a visualidade da figura. A seção 3.2 descreve melhor e detalha este modelo.



## 3.2 Módulos de Dados

Considerando a complexidade em atender todos os requisitos apresentados na seção anterior, as entidades e os relacionamentos foram divididos em módulos, conforme é mostrado no diagrama de pacotes da Figura 3.2. As dependências existentes entre as entidades de módulos distintos estão representadas no diagrama pelas setas pontilhadas.

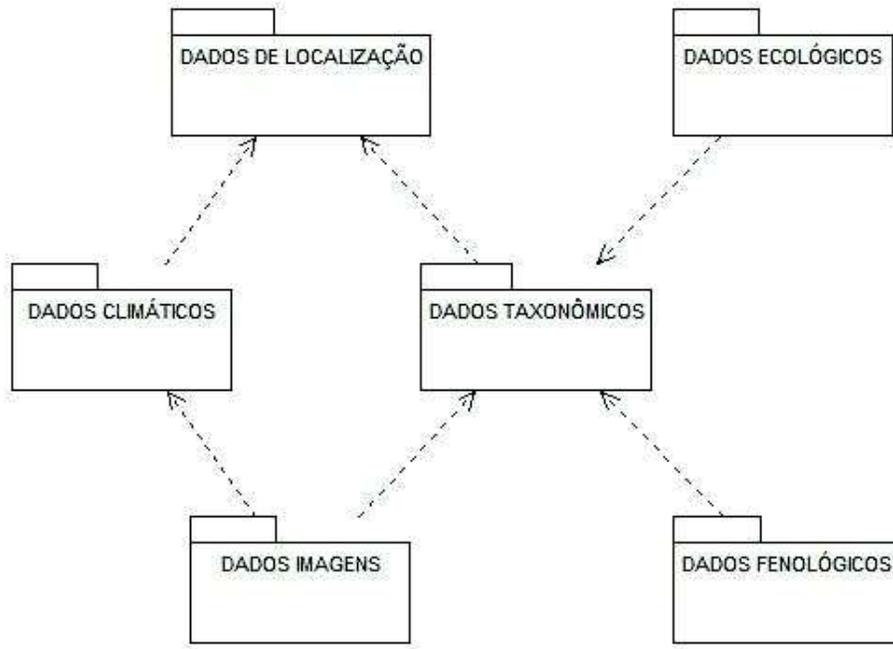


Figura 3.2: Diagrama de pacotes – visão geral do banco de dados do projeto e-phenology.

A maneira como as entidades foram divididas e os relacionamentos existentes entre eles são explicados a seguir:

- **Dados de Localização:** este módulo agrupa as entidades que definem a localização dos indivíduos no campo. Principais relacionamentos: (a) a entidade que identifica a posição de um indivíduo (*Site*) relaciona-se com a entidade de indivíduo (*Individual*) do módulo de dados taxonômicos; (b) a entidade de cidades (*City*) relaciona-se com a entidade de torres (*Tower*) do módulo de dados climáticos.
- **Dados Taxonômicos:** este módulo inclui as entidades que definem a classificação taxonômica dos indivíduos. Principais relacionamentos: (a) a entidade de indivíduos (*Individual*) relaciona-se com a entidade de dados fenológicos (*DataPhenology*) do módulo de dados fenológicos e com as entidades de visitantes dispersores *Disperser* e

polinizadores (*Pollinator*) do módulo de dados ecológicos; (b) a entidade de espécies (*Specie*) relaciona-se com a entidade de imagens (*Images*) do módulo de dados de imagens e com todas as entidades do módulo de dados ecológicos.

- Dados Ecológicos: este módulo inclui as entidades que definem as características ecológicas das espécies. Principais relacionamentos: todas as entidades se relacionam ou com a entidade de indivíduos (*Individual*) ou com a entidade de espécies (*Specie*), ambas do módulo de dados taxonômicos.
- Dados Climatológicos: este módulo considera as entidades que guardam as informações de sensores e leituras de dados climáticos. Principais relacionamentos: a) a entidade de torres (*Tower*) relaciona-se com a entidade de cidades (*City*) do módulo de dados de localização e (b) a mesma entidade de torres (*Tower*) relaciona-se com a entidade de cameras (*Camera*) do módulo de dados de imagens.
- Dados Fenológicos: este módulo contempla as entidades que guardam as pesquisas realizadas no campo quanto a fenologia das plantas, e séries temporais. Principais relacionamentos: a entidade de dados fenológicos (*DataPhenology*) relaciona-se com a entidade de indivíduos (*Individual*) do módulo de dados taxonômicos.
- Imagens e Metadados: este módulo inclui as entidades que armazenam diretórios de imagens e seus metadados. Principais relacionamentos: a entidade de câmeras (*Camera*) relaciona-se com a entidade de torres (*Tower*) do módulo de dados climáticos para especificar a localização da câmera.

As próximas subseções deste capítulo detalham os módulos, com as respectivas entidades mapeadas em tabelas e seus atributos. A descrição destas tabelas e seus atributos é apresentada no Apêndice B.

### 3.2.1 Módulo Dados de Localização

O módulo de dados de localização refere-se às tabelas que permitem a identificação e localização de um indivíduo no campo, ou seja, ao local onde as coletas de dados e observações são realizadas. A Figura 3.3 mostra o Diagrama de Entidade-Relacionamento proposto para esse módulo. As principais tabelas desse diagrama são *Site* e *SamplingMethod*, que determinam a posição do indivíduo com relação a área que ele se encontra e possibilitam a análise do comportamento de indivíduos iguais em ambientes distintos dentro de uma mesma área.

Para métodos de amostragem diferentes (*SamplingMethod*), as informações a serem armazenadas para encontrar a posição final do indivíduo muda. Desta forma, para pos-

sibilitar que outros métodos de amostragem pudessem também ser utilizados, as tabelas foram modeladas conforme a divisão que os fenólogos estipulam para encontrar o indivíduo no campo: primeiro divide-se a área em ambientes (*SamplingEnvironment*); depois definem-se sub-regiões nesses ambientes de acordo com sua localização (*SamplingAreaEnvironment*); e por fim, demarca-se um perímetro que define de fato a posição (*Site*) de um indivíduo.

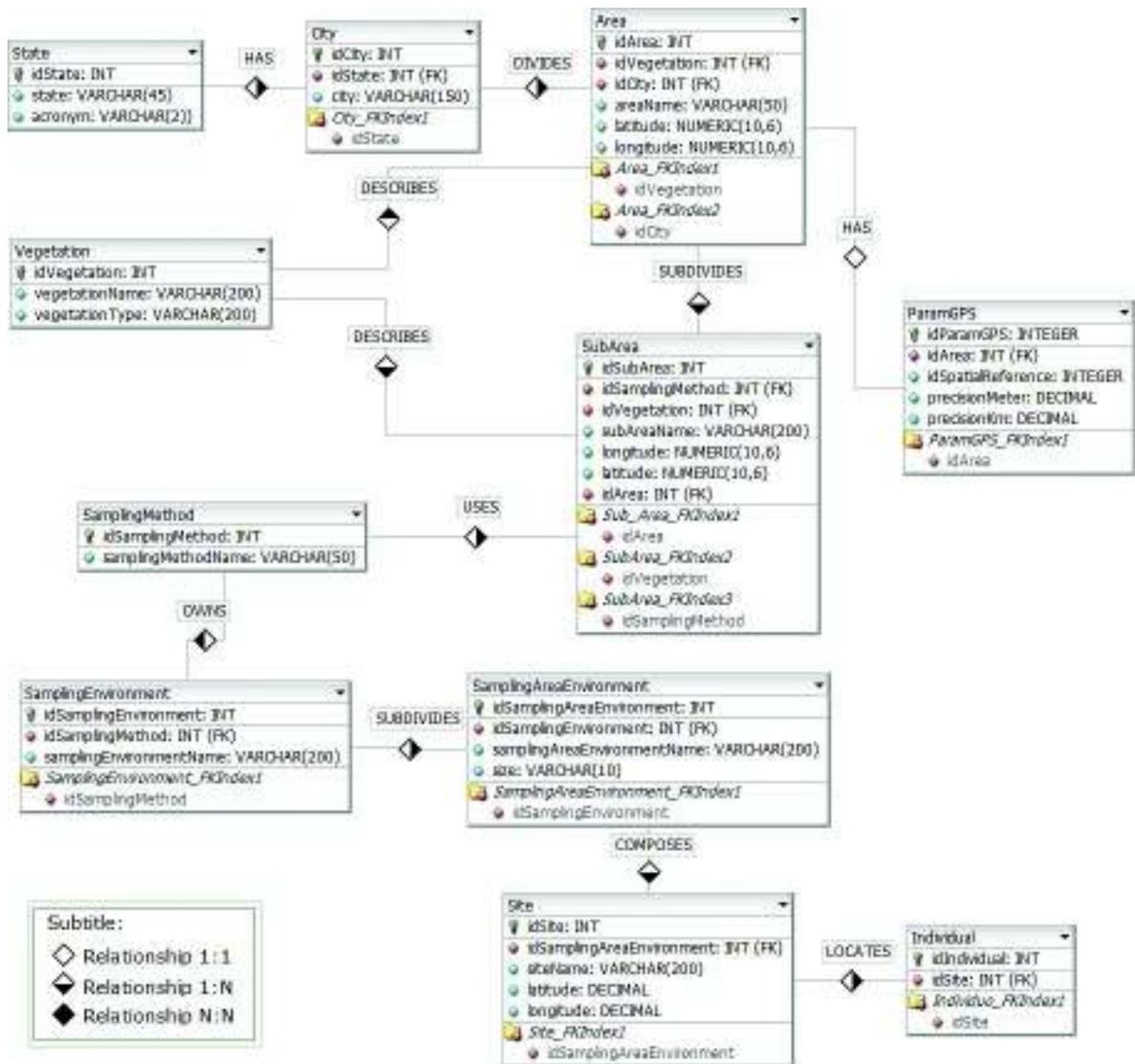


Figura 3.3: Diagrama dos dados de Localização.

Exemplificando, os dados de coleta de fenologia realizados no Cerrado pelo grupo do Laboratório de Fenologia da UNESP são de indivíduos (*Individual*) que estão em uma área (*Area*) do município (*City*) de Itirapina do Estado (*State*) de São Paulo. Essa área,

também conhecida como Botelho, possui um único tipo de vegetação (*Vegetation*) que a caracteriza e por isso, possui apenas uma sub-área (*SubArea*) também chamada de Botelho. Cada sub-área usa um método de amostragem (*SamplingMethod*), que para o exemplo dado chama-se Transecto (representado pelos fragmentos na Figura 3.4).

Este método de amostragem determina ambientes (*SamplingEnvironment*) que identificam regiões onde os indivíduos estão; no exemplo, são Bordas ou Interiores. Esses ambientes podem, por sua vez, serem sub-divididos de acordo com sua localização (*SamplingAreaEnvironment*), como Borda Leste e Borda Sul, Interior Leste e Interior Sul. Cada uma dessas áreas dentro de cada ambiente podem, ainda, serem divididos em pedaços que refinam ainda mais a posição (*Site*) que o indivíduo está localizado. Para o exemplo, as Bordas Sul e Leste podem ser divididas em perímetros numerados de 1 a 10 cada uma, já os interiores Sul e Leste podem ser divididos em perímetros numerados de 1 a 8 cada um, o que totaliza um conjunto de 36 perímetros onde os indivíduos são, assim, facilmente encontrados.



Figura 3.4: Visão da área de estudo de fenologia vegetal com seus respectivos ambientes, Itirapina-SP.

Outra informação importante na coleta de dados no campo é a localização referenciada pelo uso de aparelhos de geoprocessamento. Os dados gerenciados por estes aparelhos serão armazenados na tabela de parâmetros de geoprocessamento (*ParamGPS*) e se referem a uma área de estudo.

### 3.2.2 Módulo Dados Taxonômicos

O módulo de dados taxonômicos refere-se às tabelas que descrevem e classificam os indivíduos que fazem parte da pesquisa. A Figura 3.5 mostra o diagrama de Entidade-Relacionamento para esse módulo. As principais tabelas desse diagrama são *Individual* e *Specie*. A primeira tabela refere-se a um indivíduo que representa uma planta estudada, enquanto que a segunda refere-se ao agrupamento de indivíduos (espécimes) com características semelhantes e que, por serem extremamente parecidos, pertencem a uma espécie.

Uma espécie possui dados estatísticos associados a ela, isto é, dados que indicam a ocorrência de características específicas das espécies e a sua variação no espaço que são denominados dados fitossociológicos (*Phytossociologic*). Além disso, as espécies podem sofrer alterações taxonômicas ao longo do tempo, o que dificulta o trabalho dos biólogos em classificar seus indivíduos, tendo que recorrer a catálogos que registram as árvores taxonômicas com seus sinônimos e autores. É cada vez mais comum os esforços para tentar resolver as inconsistências geradas por estas alterações [31].

Para o projeto e-phenology, as alterações que ocorrem com uma espécie também são muito importantes. Os pesquisadores precisam sempre do nome mais recente, bem como seus sinônimos. Isto gerou um requisito, em que, a cada alteração que uma espécie sofresse referente à sua classificação taxonômica (em família, gênero ou epíteto específico) o banco deveria guardar a informação original e referenciá-la. Para atender ao requisito apresentado, optou-se por acrescentar datas que definem a validade do registro e mantêm a referência para o estado anterior, dentro da própria tabela espécie, conforme o trabalho [17] resume.

Um indivíduo é caracterizado tanto por um hábito (*Habit*) quanto por um estrato (*CanopyStatus*). Um hábito define o tipo da planta, como por exemplo, árvore ou arbusto. Um estrato representa a altura da planta quando da análise vertical do ambiente que ela se encontra. Por exemplo, o estrato de uma planta rasteira chama-se herbácea (fica próxima ao solo); o estrato de uma planta alta (com cerca de mais de vinte metros de altura) chama-se dossel. Ainda para o indivíduo, há a relação com a tabela método (*Method*), que identifica as diferentes maneiras que uma observação no campo pode ser feita. Por exemplo, os dados podem ser coletados usando um método qualitativo que avalia a presença ou ausência da fenofase ou usando um método quantitativo que quantifica a fenofase [15].

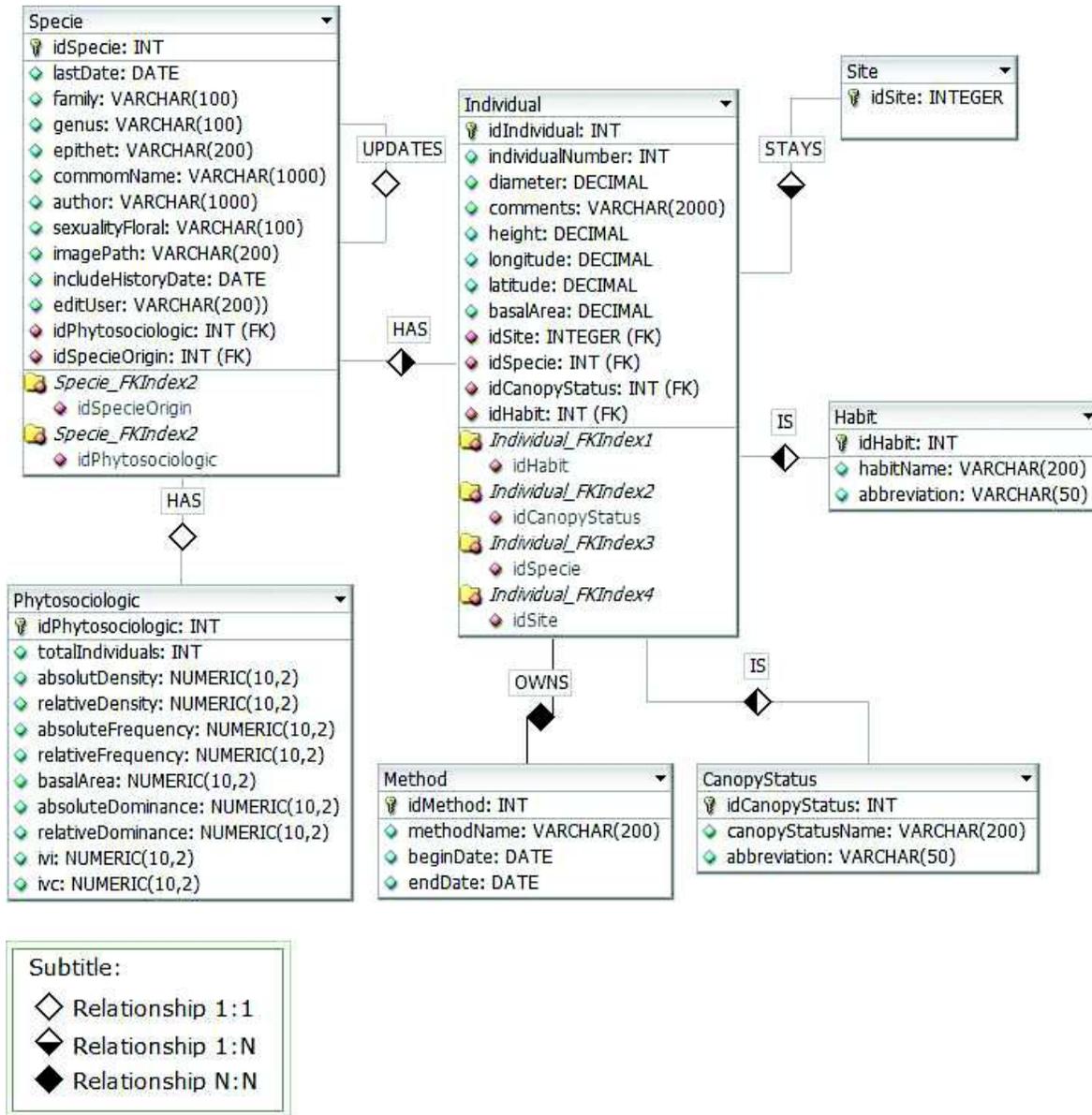


Figura 3.5: Diagrama do Módulo de Dados Taxonômicos.

### 3.2.3 Módulo Dados Ecológicos

O módulo de dados ecológicos refere-se às tabelas que descrevem as características dos indivíduos que fazem parte da pesquisa. A Figura 3.6 mostra o Diagrama de Entidade-Relacionamento para esse módulo. As principais tabelas desse diagrama são aquelas que permitem definir e caracterizar as espécies de acordo com seu tipo de flor (*Flower*), folha (*Leaf*) e fruto (*Fruit*).

Outras tabelas importantes deste diagrama são aquelas que relacionam os eventos de polinização e de dispersão de cada espécie. Para que o processo de fecundação das espécies vegetais – dita polinização – aconteça, as plantas precisam de algum tipo de agente polinizador (*Pollinator*) que seja capaz de transmitir o pólen da antera de uma flor até o estigma de outra flor. Este processo pode ser ocasionado por diferentes fatores ditos agentes polinizadores: animais, água, vento, entre outros. Para cada tipo de agente polinizador, os cientistas esquematizam as síndromes de polinização (*PollinationModelSpecie*) de cada espécie, como por exemplo: Anemofilia (para os animais) e Hidrofilia (para a água)<sup>1</sup>.

Tal como na polinização, a dispersão precisa de agentes dispersores (*Disperser*) capazes de espalhar as sementes dos frutos das espécies para que ela se reproduza. Normalmente, estes agentes são os mesmos da polinização. Assim, para cada tipo de agente dispersor, tem-se uma síndrome de dispersão (*DisperserModelSpecie*), como por exemplo: Zoocoria (para os animais) e Anemocoria (para o vento).

Outra tabela que também está representada neste módulo é a deciduidade (*Deciduous*), que refere-se à frequência de perdas de folhas e a níveis de indivíduo e de espécies.

---

<sup>1</sup><http://www.infoescola.com/plantas/polinizacao/>. Acesso em: 11/03/2013

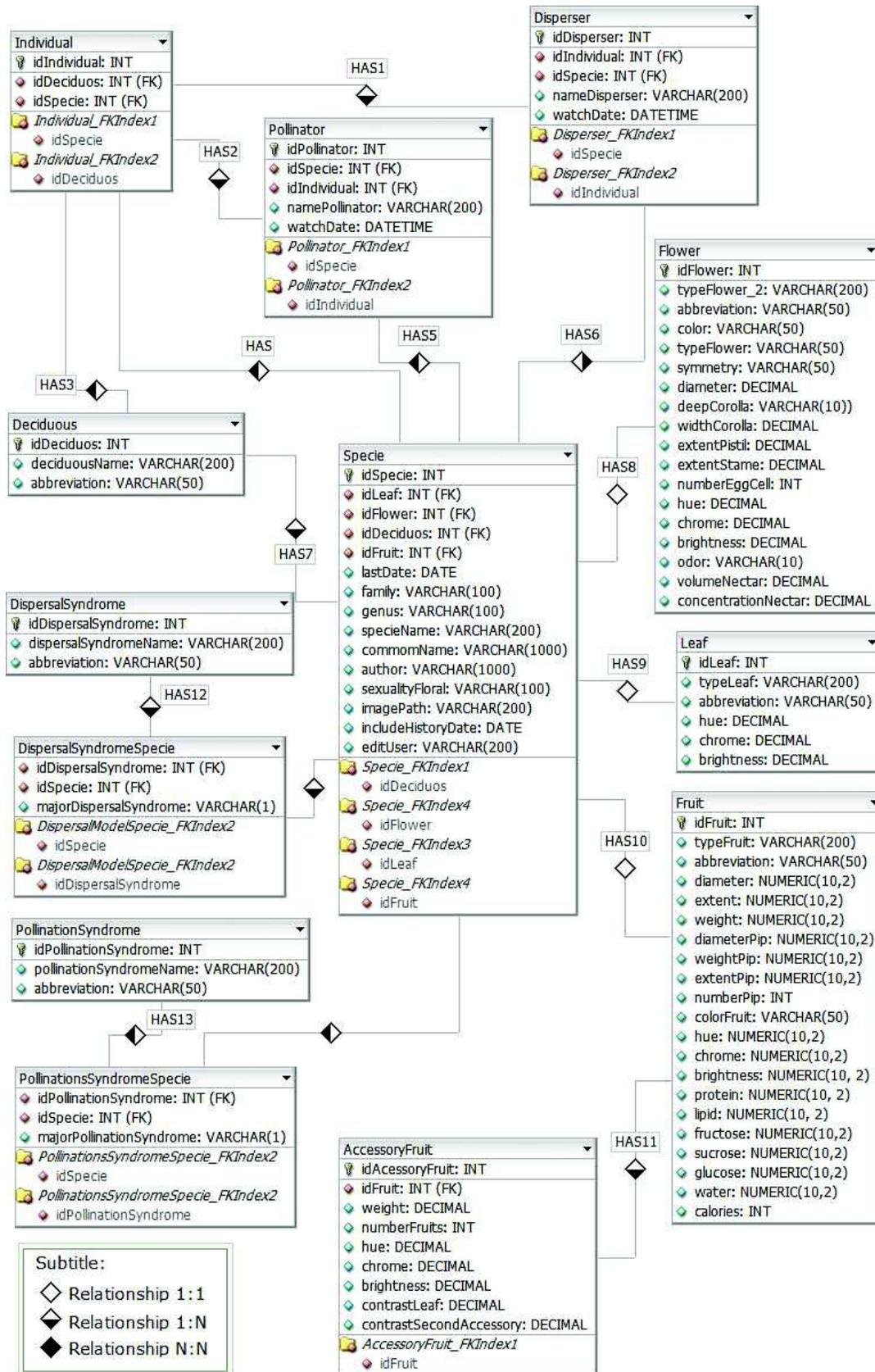


Figura 3.6: Diagrama do Módulo de Dados Ecológicos.

### 3.2.4 Módulo Dados Climatológicos

O módulo de dados climatológicos refere-se às tabelas relacionadas ao monitoramento de dados temporais do clima. A Figura 3.7 mostra o Diagrama de Entidade-Relacionamento para esse módulo. A principal tabela desse diagrama é a *DateClimate*, responsável por armazenar os dados climáticos coletados por diversos sensores instalados em uma área de observação de fenologia.

No campo de pesquisa, por exemplo, em Botelho, foi instalada uma torre (*Tower*) para observação remota de fenologia por câmeras digitais. Além da câmera, nove sensores (*Sensor*) de clima, entre eles de umidade de solo, pluviômetro e temperatura do solo (*SensorType*), também foram instalados. Estes são os responsáveis pela leitura dos dados climáticos. Os dados lidos atualizam, por meio de conexão com a internet, um arquivo no formato *Comma-separated values* (CSV) de dez em dez minutos. Um exemplo dos dados deste arquivo são apresentados pelo Apêndice C, em que as colunas representam os campos data e hora e nomes dos sensores em que os valores coletados são guardados.

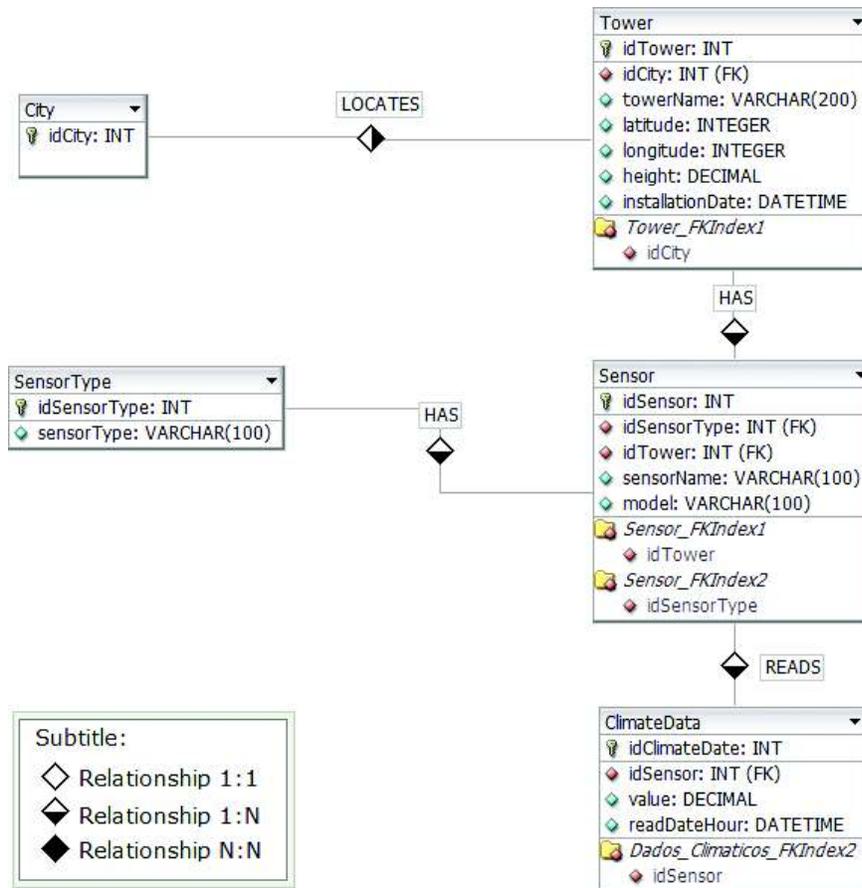


Figura 3.7: Diagrama do Módulo de Dados Climatológicos.

### 3.2.5 Módulo Dados Fenológicos

O módulo de dados fenológicos refere-se às tabelas relacionadas à coleta de dados de observação de fenologia. A Figura 3.8 mostra o Diagrama de Entidade-Relacionamento para esse módulo. A principal tabela desse diagrama é *DataPhenology*, que armazena os dados de fenologia dos indivíduos.

Os dados de fenologia são valores de observação que mensuram a intensidade de ocorrência de uma fenofase (*Phenophase*) de um indivíduo, em uma determinada data. São dados coletados mensalmente a partir das observações realizadas no campo por especialistas em fenologia.

A partir destes valores de intensidade coletados para os indivíduos, alguns dados importantes são calculados para estimar, por exemplo, quando há maior probabilidade de uma determinada espécie ter flores ou frutos maduros. Este tipo de levantamento é feito observando os chamados ciclos de fenofase na tabela de fenologia, isto é, quando um determinado indivíduo apresenta intensidade diferente de 0 para uma fenofase.

Por exemplo, um indivíduo de uma espécie pode ter anotado na planilha de fenofase fruto maduro a intensidade diferente de 0 para os períodos de março à maio de 2012 e outubro à novembro de 2012. Desta observação, alguns dados estatísticos serão considerados e armazenados na tabela *DataCycle*:

- a duração de cada ciclo para a fenofase fruto maduro (*durationCycle*);
- a data que iniciou cada ciclo (*beginCycle*);
- o total de meses no ano em que foi observada a ocorrência do ciclo (*durationYear*);
- o mês no qual houve registro de maior intensidade da ocorrência do ciclo no período de um ano (*peakYear*);
- o mês em que houve maior intensidade de ocorrência da fenofase durante o ciclo (*peakCycle*).

No caso do pico do ciclo, quando há um ciclo de fenofase que ultrapassa um ano (quando o início da fenofase é observada em um ano e seu término em outro ano), é considerado como maior pico o mês onde encontra-se o maior valor de intensidade dentro do período do ano observado, no caso, onde houver a intensidade 2.

Portanto, a tabela *DataCycle* foi mapeada para armazenar estas consultas que levantam estes dados estatísticos, visto que há a necessidade de se fazer diferentes consultas a nível de indivíduos e de espécies. A avaliação também pode ser feitas usando-se diferentes períodos de interesse, o que pode gerar diferentes dados para os picos de acordo com o período escolhido.

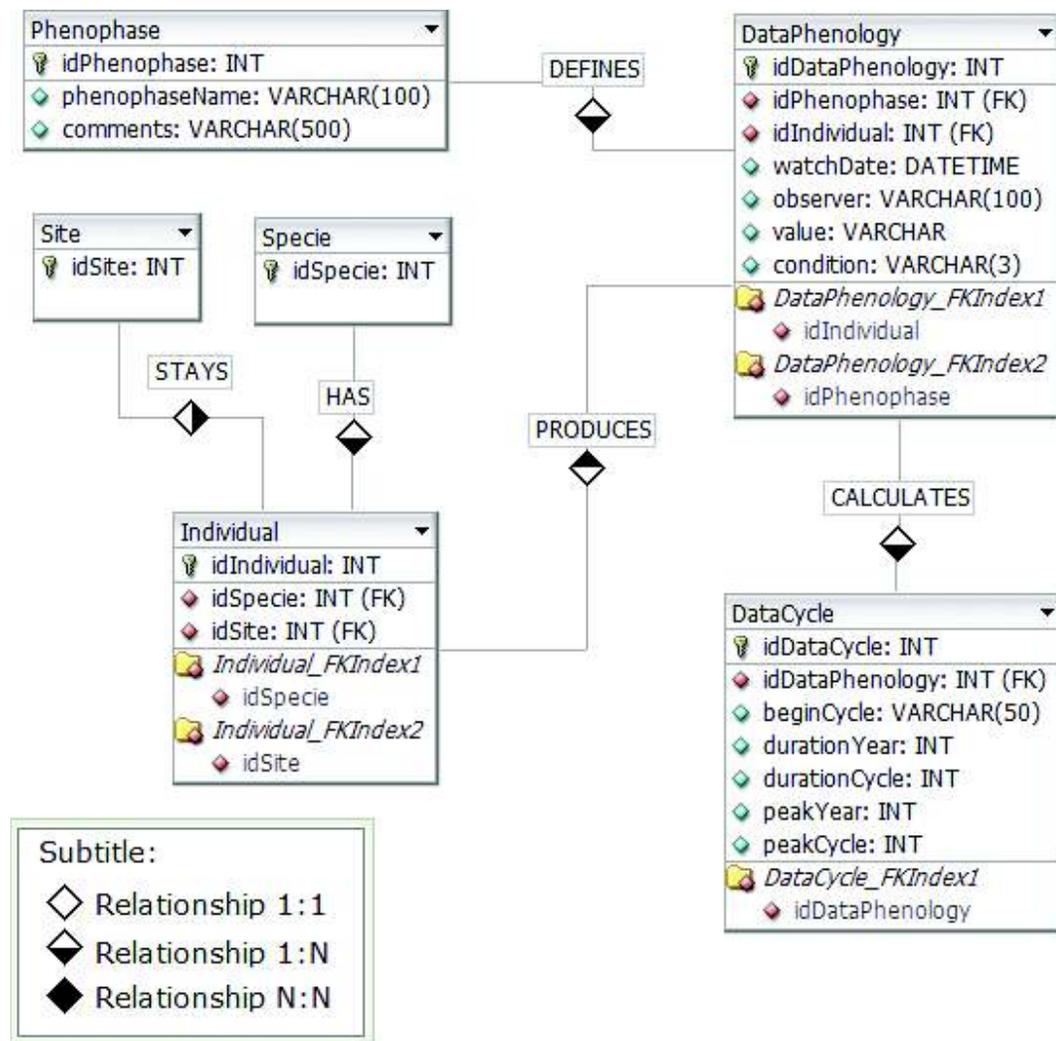


Figura 3.8: Diagrama do Módulo de Dados Fenológicos.

### 3.2.6 Módulo Imagens e Metadados

O módulo de imagens e metadados refere-se ao armazenamento das informações relacionadas aos dados e imagens provenientes da câmera digital instalada na torre na área de coleta de dados. Essas imagens são fotos de áreas de interesse de pesquisa e a partir das quais se dará a observação remota da fenologia.

A Figura 3.9 mostra o Diagrama de Entidade-Relacionamento para esse módulo. Uma tabela em destaque desse diagrama é a tabela *Image* que, além de armazenar qual câmera (*Camera*) produziu a imagem, também guarda a data e hora que a foto foi tirada e o

caminho (diretório) da imagem no sistema de arquivos.

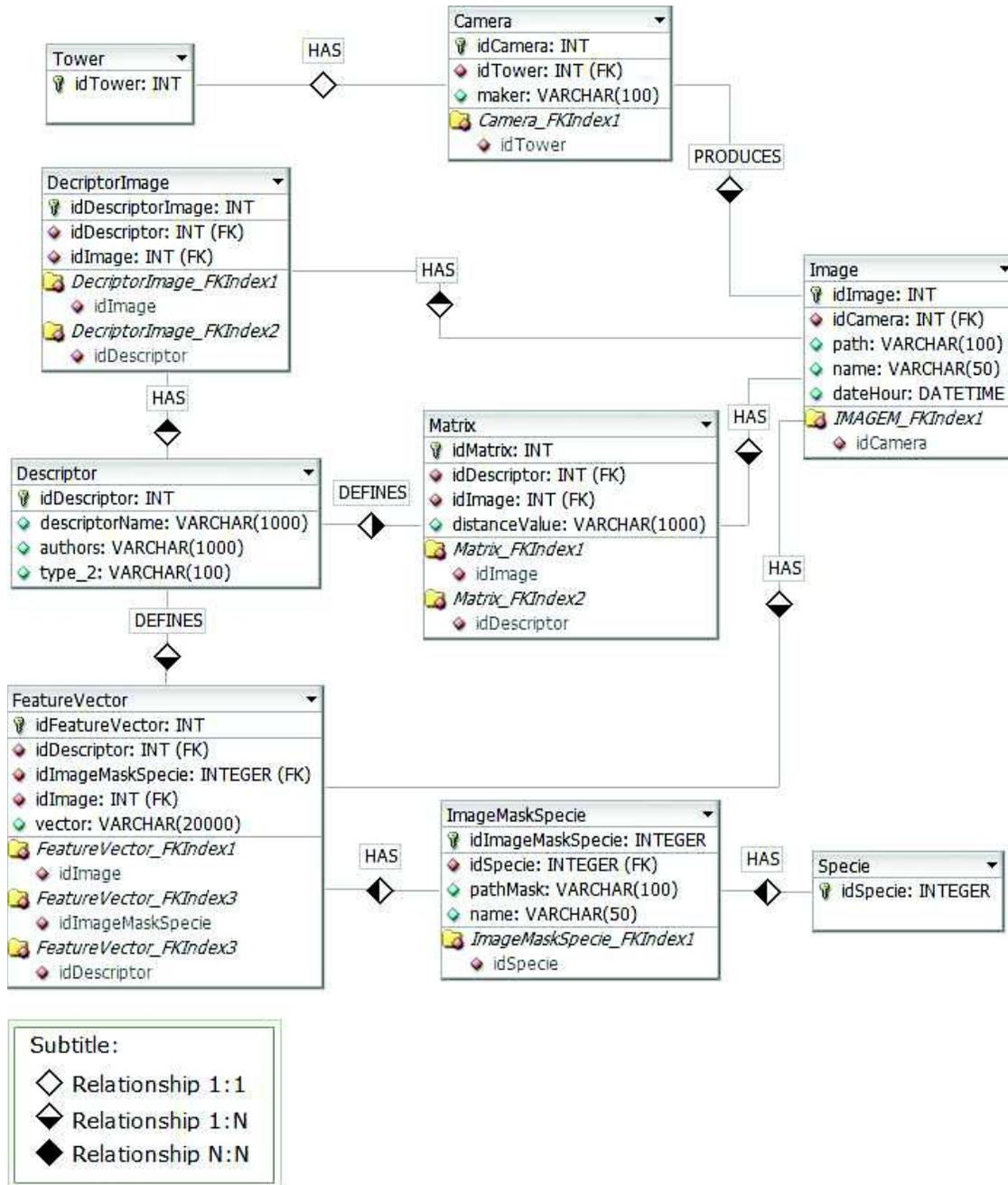


Figura 3.9: Diagrama do Módulo de Imagens e Metadados.

Para a proposição dessa modelagem, tomou-se como base um trabalho de mestrado realizado no Instituto de Computação da UNICAMP [40], no qual foi realizado um estudo

comparativo de descritores de imagens considerando a Web como cenário de uso. Da mesma forma, descritores (*Descriptor*) serão utilizados neste trabalho para caracterizar imagens a partir do seu conteúdo [12] e possibilitar a análise fenológica remotamente.

Um descritor de imagem pode ser definido como um par composto de um vetor de características (*Feature Vector*) e uma função de distância [11]. Um vetor de características representa um conjunto de propriedades de uma determinada imagem [11], como por exemplo, cor, textura ou forma. A propriedade de um vetor é determinada pelo tipo do descritor. Uma função de distância calcula a similaridade entre duas imagens por meio de seus vetores de características [12]. Por ora, os valores de similaridade obtidos a partir dos cálculos pela aplicação da função de distância serão armazenados como um valor de distância (*Matrix*) de uma imagem em relação a todas as outras.

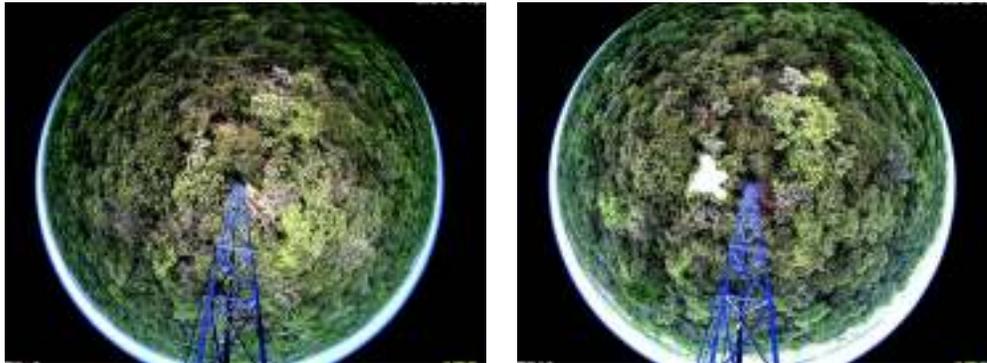
O e-phenology conta, atualmente, com uma câmera digital (*Camera*) instalada em uma Torre (*Tower*) na cidade de Itirapina, São Paulo. Esta câmera foi configurada para obter imagens de hora em hora, em um período de doze horas (das seis horas da manhã até às seis horas da tarde). Dada uma hora cheia, por exemplo, meio-dia, a câmera liga e gera cinco imagens (*Image*), tal qual como exemplifica a Figura 3.10-(a), automaticamente da área para a qual ela está direcionada (focalizada).

Cada imagem possui, entre os seus metadados, os atributos câmera e data e hora que foi fotografada. As imagens ficam armazenadas no cartão de memória da câmera e também ficam disponíveis na internet, por meio da qual é possível obtê-las e armazená-las no servidor de imagens do projeto, em um diretório específico para isso. O diretório fica armazenado com a imagem (na tabela *Image*, assim como seus descritores (*DescriptorImage*), seus vetores de características (*Feature Vector*) e sua matriz de distância (*Matrix*), informações importantes para análises das mesmas.

Um ponto importante em usar o monitoramento remoto para a fenologia está em relacionar os fenômenos observados na variação das imagens com os dados do campo. Para tanto, regiões de interesse, chamadas de máscaras, foram definidas em campo pelos especialistas de fenologia responsáveis por estes estudos. As copas das árvores foram mapeadas nas fotos e destas fotos foram geradas imagens binárias para que o processamento individual de cada espécie pudesse ser realizado. Desta forma, os dados destas imagens são armazenados e relacionados à espécie que ela se refere (*ImageMaskSpecie*) e às características extraídas (*Feature Vector*).

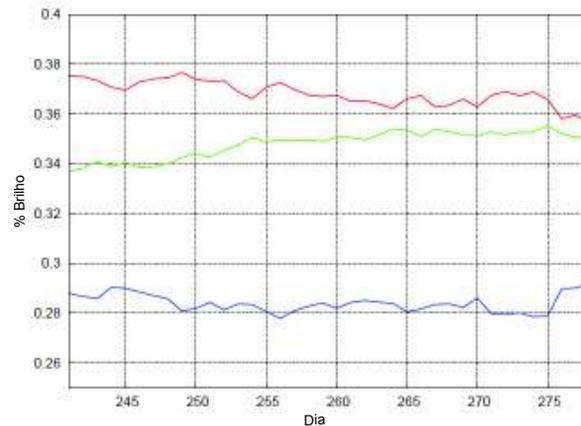
Um exemplo de mapeamento de máscara e os resultados gerados pelo processamento da máscara sobre as imagens obtidas no Cerrado são apresentadas pela Figura 3.10-(b) e (c), respectivamente. A Figura 3.10-(b) apresenta a máscara da espécie *Caryocar Brasiliensis* mapeada sobre o cerrado. A Figura 3.10-(c) apresenta o vetor de característica extraído que consiste nas médias dos canais vermelho, verde e azul (RGB) para o processamento realizado da máscara *Caryocar Brasiliensis* sobre um conjunto de imagens, de vários dias,

do Cerrado.



(a) *Visão Cerrado da Câmera*

(b) *Caryocar brasiliensis*



(c) *Caryocar brasiliensis*

Figura 3.10: Extração de características das imagens do Cerrado: (a) Visão da imagem tirada do Cerrado, (b) Máscara da espécie *Caryocar brasiliensis* aplicada à imagem, (c) Resultado do processamento da máscara da espécie *Caryocar Brasiliensis*. Apresenta a média dos canais R (*Red* – Vermelho), G (*Green* – Verde) e B (*Blue* – Azul) no decorrer dos dias [4].

### 3.3 Exemplo de informações armazenadas

Esta seção apresenta alguns exemplos das informações que os usuários-alvo desejam armazenar com a construção do banco de dados. Os exemplos aqui apresentados referem-se aos dados existentes que popularão o banco de dados. Isto significa que os dados são apresentados em conformidade com o que foi apresentado nas seções anteriores. Logo, para cada módulo, são apresentadas apenas as tabelas cujos dados já são conhecidos, isto

é, fazem parte do dia-a-dia das pesquisas em fenologia de plantas realizadas pelos biólogos do Laboratório de Fenologia da UNESP.

Para melhor compreensão das tabelas, as chaves primárias foram omitidas e as chaves estrangeiras foram substituídas pelo valor que elas representam em uma consulta geral, sendo indentificadas por \* na coluna campo.

Todas as tabelas de exemplos são representadas por dados existentes para a área de estudo do Laboratório de Fenologia da UNESP. Portanto, as primeiras tabelas apresentadas são as do Módulo de Dados de Localização, que define a localização da área de Cerrado situada em Itirapina no Estado de São Paulo. Assim, estas informações são armazenadas no banco de dados conforme explicado a seguir.

A tabela *State* (Tabela 3.1) apresenta os atributos para o Estado de São Paulo. A tabela *City* (Tabela 3.2) apresenta os atributos para a cidade de Itirapina. A tabela *Area* (Tabela 3.3) apresenta os atributos para a área do Cerrado, denominada Botelho. A área Botelho não possui sub-área, por isso, a tabela *SubArea* será semelhante a de *Area* e não é exibida neste exemplo.

Tabela 3.1: Informação de localização: Estados.

<i>State</i>	
<b>Campo</b>	<b>Valor</b>
state	São Paulo
acronym	SP

Tabela 3.2: Informação de localização: Cidades.

<i>City</i>	
<b>Campo</b>	<b>Valor</b>
city	Itirapina
*state	São Paulo

Como foi explicado anteriormente, em Botelho, o método de amostragem chama-se Transecção que define marcações sistemáticas no campo de comprimentos e direções diferentes, por isso, as tabelas *SamplingEnvironment* (Tabela 3.4) e *SamplingAreaEnvironment* (Tabela 3.5) apresentam as respectivas demarcações para esta metodologia, sendo a definição da região e as sub-divisões que são feitas, respectivamente.

Por fim, a tabela *Site* (Tabela 3.6) apresenta as demarcações finais feitas no Cerrado para localizar os indivíduos.

Tabela 3.3: Informação de localização: Áreas.

<i>Area</i>	
<b>Campo</b>	<b>Valor</b>
areaName	Botelho
latitude	
longitude	
*city	Itirapina
*vegetation	Cerrado <i>sensu strictu</i>

Tabela 3.4: Informação de localização: Ambientes de Amostragem de coleta.

<i>SamplingEnvironment</i>	
<b>Campo</b>	<b>Valor</b>
samplingEnvironment	Borda
samplingMethod	Transecto

Tabela 3.5: Informação de localização: Áreas de Ambientes de Amostragem de coleta.

<i>SamplingAreaEnvironment</i>	
<b>Campo</b>	<b>Valor</b>
samplingAreaEnvironment	Borda Leste
*samplingEnvironment	Borda
size	25x50m

Tabela 3.6: Informação de localização: Locais (posição que o indivíduo se encontra efetivamente).

<i>Site</i>	
<b>Campo</b>	<b>Valor</b>
siteName	BL1
latitude	
longitude	
*samplingAreaEnvironment	Borda Leste

Para o Módulo de Dados Taxonômicos foram selecionadas as principais tabelas citadas e cujos valores estão disponíveis. Desta forma, as tabelas *Method*, *Habit*, *CanopyStatus* e *Phytosociologic* não são exibidas. A tabela *Specie* (Tabela 3.7) apresenta um exemplo de

informações da espécie *Rapanea guianensis* e seus atributos. Os campos *deciduous*, *leaf*, *flower* e *fruit* presentes na tabela são campos referentes às tabelas que caracterizam a espécie ecologicamente e não são exemplificadas.

A tabela *Individual* (Tabela 3.8) apresenta os atributos de um indivíduo que pertence à espécie *Rapanea guianensis*. Os campos *canopyStatus* e *Habit* se referem a duas tabelas que caracterizam o indivíduo quanto ao seu ambiente (estrato e habito). Detalhes são omitidos.

Tabela 3.7: Informação de taxonomia: Espécies.

<i>Specie</i>	
<b>Campo</b>	<b>Valor</b>
family	Myrsinaceae
genus	Rapanea
specieName	guianensis
author	Aubl
commumName	Capororoca Branca
sexualityFloral	BL1
imagePath	C:\Imagem
lastDate	10/10/2012
includeHistoryDate	
editUser	Ana
*deciduous	
*leaf	dispersas
*flower	cachos de botões florais
*fruit	drupa

Do Módulo de Dados Ecológicos, são exemplificados os modos de dispersão e polinização de sementes das espécies e indivíduos. Assim, a tabela *PollinationModelSpecie* (Tabela 3.9) relaciona a síndrome de polinização melitofilia (polinização por abelhas) como principal para a espécie *Rapanea guianensis*. A tabela *DispersalModelSpecie* (Tabela 3.10) apresenta a síndrome de dispersão zoocórica (dispersão por animais) como principal para a mesma espécie.

O Módulo de Dados Fenológicos é representado por duas tabelas: a de coleta de fenologia e a que armazenará os dados estatísticos gerados a partir de consultas na tabela de coleta.

Com o objetivo de separar as fenofases observadas para o mesmo indivíduo, a tabela *DataPhenology* (Tabela 3.13) apresenta como seria um registro para coleta observada para o indivíduo 1 exemplificado, o valor da fenofase botão no dia 10 de outubro de 2012 coletado pela bióloga Ana. Similarmente, as demais fenofases são registradas.

Tabela 3.8: Informação de taxonomia: Indivíduos.

<i>Individual</i>	
<b>Campo</b>	<b>Valor</b>
idIndividual	1
comments	sem placa, atrás da 5
diameter	15cm
height	20m
latitude	
longitude	
*siteName	BL1
*canopyStatus	arboreo
*habit	arvore
*specie	Rapanea guianensis

Tabela 3.9: Informação de dados ecológicos: Síndromes de Polinização de Espécies.

<i>PollinationSyndromeSpecie</i>	
<b>Campo</b>	<b>Valor</b>
pollinationSyndromeName	melitofilia
specie	Rapanea guianensis
majorPollinationSyndrome	S

Tabela 3.10: Informação de dados ecológicos: Síndromes de Dispersão de Espécies.

<i>DispersalSyndromeSpecie</i>	
<b>Campo</b>	<b>Valor</b>
dispersalSyndrome	Zoocórica
specie	Rapanea guianensis
majorDispersalSyndrome	S

O Módulo de Dados Climáticos contém uma tabela que é a *ClimateData* em que os registros são valores numéricos registrados de acordo com o sensor que coletou, e por conter apenas números, é omitida. E, finalmente, o Módulo de Dados de Imagens é representado por sua principal tabela *Image* e a tabela *ImageMaskSpecie*. Em ambas as tabelas guarda-se a informação do local onde a imagem está armazenada no servidor. Na tabela *ImageMaskSpecie* o exemplo supõe a máscara da espécie *Rapanea guianensis*.

Tabela 3.11: Informação de dados fenológicos: Dados de Fenologia.

<i>DataPhenology</i>	
<b>Campo</b>	<b>Valor</b>
watchDate	10/10/2012
observer	Ana
*phenophaseName	Botão
idIndividual	1
valor	2

Tabela 3.12: Informação de dados de imagens: Imagens.

<i>Image</i>	
<b>Campo</b>	<b>Valor</b>
*camera	Botelho
path	C:/server/photos
name	2011_230_16_1.jpg
dateHour	2012-11-13 06:00:00

Tabela 3.13: Informação de dados de imagens: Imagens de Máscaras de Espécies.

<i>ImageMaskSpecie</i>	
<b>Campo</b>	<b>Valor</b>
*specie	Rapanea guianensis
pathMask	C:/server/photosMask
name	mask_R.guianensis.jpg

### 3.4 Exemplo de consultas

Esta seção apresenta alguns exemplos de consultas que podem ser efetuadas a partir do modelo proposto e que foram apontadas como importantes para os usuários-alvo. Seguindo a mesma metodologia adotada nos trabalhos desenvolvidos em [25] e [27], esta dissertação divide as consultas em simples e complexas.

Consultas simples são aquelas cujos resultados são obtidos a partir do uso de predicados simples envolvendo uma única tabela. Em geral, considera-se nesse caso o uso das funções padrões da Álgebra Relacional [18]. Um exemplo de consulta simples é “Mostrar quais são as espécies cadastradas”. Esta consulta exige apenas uma operação de seleção que pesquisa na tabela espécie (*Specie*) e retorna todos os registros presentes nela. Outro exemplo deste tipo de consulta seria “Mostrar todas as espécies cujo autor é King”. Esta

consulta, assim como a primeira, exige que seja realizada apenas uma operação de seleção na tabela espécie (*Specie*), porém ela tem um diferencial que é o filtro pelo autor, e portanto, deve retornar apenas os registros de acordo com o autor que foi especificado.

Todas as demais consultas, que envolvem a junção de uma ou mais tabelas e/ou cujos predicados exigem processamento de funções específicas, são consideradas complexas. Um exemplo deste tipo de consulta é “Mostrar as características de flor, folha e fruto das espécies”. Esta consulta requer uma operação de junção que combina informações das tabelas de espécie (*Specie*), flor (*Flower*), fruto (*Fruit*) e folha (*Leaf*). Em Álgebra Relacional, esta combinação é feita com junções, o que exige a busca em um conjunto maior de dados para retornar o resultado desejado. Embora o uso de filtro – como apresentado no exemplo de consultas simples - seja possível, esta consulta continuará sendo complexa pelo fato de combinar dados de tabelas diferentes.

Consultas complexas podem, também, envolver relacionamentos espaciais ou temporais [25]. No contexto deste trabalho, muitas consultas temporais serão realizadas para, principalmente, correlacionar informações de fenofases, indivíduos e espécies com dados que variam ao longo do tempo – dados fenológicos e climáticos. Um exemplo de consulta complexa com predicados temporais é: “Quais as características dos indivíduos que tiveram dados de fenologia coletados acima de 0, para o período de setembro de 2005 à setembro de 2007 e cuja fenofase é brotamento?”. Esta consulta exige uma operação de busca que utiliza junções entre as tabelas indivíduo (*Individual*), espécie (*Specie*), fenofase (*Phenophase*) e de dados fenológicos (*DataPhenology*) em que o tempo e o valor da coleta de fenologia são as informações importantes para o filtro e a obtenção do resultado desejado.

Outras consultas de interesse do projeto são:

1. Qual a fenologia de um indivíduo para uma fenofase específica em um determinado período?
2. Qual a fenologia de um indivíduo para todas as fenofases coletadas em um determinado período?
3. Qual a fenologia de indivíduos de mesma espécie para um determinado período e/ou por fenofase?
4. Quais os dados de ciclo de fenologia para um indivíduo/espécie por fenofase (ou todas)?
5. Quais os indivíduos cujos métodos de coleta são ou foram do tipo Presença/Ausência?

6. Quais os indivíduos que já tiveram coleta de fenologia pelo método Presença/Ausência em determinado período?
7. Qual a fenologia dos indivíduos que foram coletados com método de Presença/Ausência em um determinado período?
8. Qual a correlação dos dados de fenologia observados com os dados de clima coletados pelos sensores?
9. Qual a correlação dos dados de fenologia observados com os dados extraídos das imagens?
10. Qual são os vetores de características e respectivos descritores de uma imagem em determinada data e hora?
11. Quais são os vetores de características das imagens relacionadas a uma determinada espécie?
12. Qual as médias dos canais R, G, B das imagens para uma determinada espécie?
13. Qual as médias dos canais R, G, B das imagens para uma determinada espécie, em data especificada?
14. Quais as médias dos canais R, G, B das imagens de todas as espécies ao longo do tempo?



# Capítulo 4

## Aspectos de Implementação

Este capítulo trata de aspectos de implementação do banco de dados proposto. Na seção 4.1, são apresentadas as etapas envolvidas no processo de construção do banco de dados para o Projeto e-phenology e quais foram as tecnologias utilizadas. A seção 4.2 explica como foram feitas as inclusões dos dados e quais dificuldades foram encontradas neste processo. A seção 4.3 apresenta algumas das consultas, definidas pelos usuários-alvos, implementadas para validar o uso do banco de dados. E, por fim, a seção 4.4 apresenta um protótipo de sistema *Web* que tem por objetivo permitir a manipulação dos dados do projeto e-phenology por interface gráfica.

### 4.1 Banco de Dados

Para gerenciar, organizar e manipular dados deve-se usar um Sistema Gerenciador de Banco de Dados (SGBD). Um SGBD é um programa computadorizado responsável por administrar um ou mais banco de dados e por tratar requisições de usuários referentes às manutenções dos dados presentes nestes bancos [14].

Existem diversos tipos de SGBDs, dentre os mais conhecidos e utilizados nos mais diferentes setores estão o PostGreSQL [23], Oracle [9] e SQL Server [8]. O banco de dados neste trabalho foi implementado usando o SGBD SQL Server 2012.

O SQL Server foi desenvolvido e é mantido pelo grupo *Microsoft Corporation*. A opção por este SGBD, para desenvolver o banco de dados deste trabalho, se deve à robustez da plataforma que o SQL Server provê, bem como o amplo conjunto de ferramentas que ele possui para facilitar a manipulação, gerenciamento, desenvolvimento e configuração de um ou mais banco de dados. Além disto, levou-se em consideração também o acesso facilitado às ferramentas da *Microsoft* por meio do convênio estabelecido entre a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) e a *Microsoft Research*, financiadores do Projeto e-phenology e deste trabalho.

Após escolher o SGBD que seria utilizado, o processo de criação do banco de dados para estudos de fenologia de plantas foi realizado em três etapas: (a) especificação e modelagem do banco de dados; (b) inclusão parcial dos dados das pesquisas realizadas pelos especialistas em fenologia e (c) especificação de consultas.

A primeira etapa foi descrita no Capítulo 3. A segunda e terceira etapa são detalhadas nas próximas seções.

## 4.2 Inclusão dos Dados do Projeto e-phenology

O primeiro passo para realizar a inclusão dos dados dos estudos de fenologia de plantas pertencentes ao Projeto e-phenology foi obter as planilhas eletrônicas utilizadas pelos pesquisadores nas quais constavam os dados que deveriam integrar o banco de dados.

As principais planilhas são aquelas relacionadas propriamente ao estudo da fenologia, onde constam os dados coletados em campo para cada fenofase analisada. Nelas estão presentes os dados básicos que os pesquisadores precisam para localizar e identificar o indivíduo e sua fenologia. Um exemplo pode ser visto na Figura 4.1 que apresenta a estrutura geral dessas planilhas por meio de uma representação parcial da planilha cuja fenofase é botão.

	A	B	C	E	F	G	H	I	J	K
1							2004			
2	Index ▼	Local ▼	Transecto ▼	Indivíduo ▼	Familia	Espécie	set	out	nov	dez
3	1	BL	11	1	Vochysiaceae	<i>Qualea grandiflora</i>	0	0	0	0
4	2	BL	11	2	Myrtaceae	<i>Myrcia guianensis</i>	1	1	0	0
5	3	BL	11	3	Myrtaceae	<i>Myrcia bella</i>	1	1	0	0
6	4	BL	11	4	Rubiaceae	<i>Tocoyena formosa</i>	0	1	1	1
7	5	BL	11	5	Rubiaceae	<i>Tocoyena formosa</i>	0	0	0	0
8	6	BL	11	7	Annonaceae	<i>Xylopia aromatica</i>	0	0	0	0
9	7	BL	11	8	Verbenaceae	<i>Aegiphila lhotskyana</i>	0	0	0	0
10	8	BL	11	9	Annonaceae	<i>Xylopia aromatica</i>	0	1	0	0
11	9	BL	11	10	Annonaceae	<i>Xylopia aromatica</i>	0	1	0	0
12	10	BL	11	11	Fabaceae	<i>Anadenanthera falcata</i>	0	0	0	0
13	11	BL	11	12	Annonaceae	<i>Xylopia aromatica</i>	0	0	0	0
14	12	BL	11	13	Malpighiaceae	<i>Banisteriopsis campestris</i>	0	0	0	0
15	13	BL	11	14	Araliaceae	<i>Schefflera vinosa</i>	0	0	0	0
16	14	BL	11	15	Fabaceae	<i>Machaerium acutifolium</i>	0	0	0	0
17	15	BL	11	16	Ochnaceae	<i>Ouratea spectabilis</i>	0	0	0	0
18	16	BL	11	18	Annonaceae	<i>Xylopia aromatica</i>	0	0	0	0
19	17	BL	11	22	Fabaceae	<i>Bauhinia rufa</i>	0	0	0	0
20	18	BL	11	23	Sapotaceae	<i>Pouteria ramiflora</i>	0	0	0	0
21	19	BL	11	24	Sapotaceae	<i>Pouteria ramiflora</i>	0	0	0	0
22	20	BL	11	25	Melastomataceae	<i>Miconia rubiginosa</i>	0	0	1	2
23	21	BL	11	26	Bignoniaceae	<i>Tabebuia ochracea</i>	0	0	0	0
24	22	BL	11	27	Erythroxylaceae	<i>Erythroxylum suberosum</i>	1	0	0	0

Figura 4.1: Exemplo de planilha típica usada para armazenar dados fenológicos.

A coluna indicada na figura pelo título *Index* refere-se a um número único para cada indivíduo, com o objetivo de identificá-lo nas demais planilhas de fenofases. As colunas *Local*, *Transecto* e *Indivíduo* referem-se às informações que permitem localizar o indivíduo no campo. No caso, a figura apresenta um exemplo de indivíduos que estão localizados na Borda Leste (BL) no Transecto 11. As colunas *Família* e *Espécie*, por sua vez, identificam os indivíduos quanto ao seu nível taxonômico. As demais colunas representam os valores coletados em campo no mês e ano indicados, no caso, os dados apresentados referem-se às coletas realizadas no período de setembro à outubro de 2004 para a fenofase Botão. A linha inteira destacada significa que aquele indivíduo em algum momento se perdeu, ou morreu ou não tem dados coletados (por erro do observador ou porque está perdido).

São nestas planilhas que estão identificados os 2122 indivíduos e as 114 espécies que fazem parte dos estudos de fenologia do Cerrado de Itirapina. Estes dados se repetem aos longos das planilhas, e devido à falta de interoperabilidade entre elas, a integração destes dados precisam ser feitas manualmente, o que quase sempre implica gerar novas planilhas para os novos estudos.

Com o objetivo de facilitar a inserção destes dados, eles foram divididos de acordo com a arquitetura proposta para o banco de dados. Assim, foram separados os dados de localização, dados taxonômicos e dados de fenologia em outras planilhas. Como, por enquanto, o banco envolve apenas uma área de estudo, todas as tabelas do módulo de dados de localização foram incluídas manualmente, isto é, por códigos gerados usando a linguagem de banco de dados denominada *Structured Query Language* (SQL)<sup>1</sup>.

As demais planilhas, cujos dados se referiam às espécies, aos indivíduos e à fenologia dos indivíduos, foram processadas respectivamente por algoritmos escritos usando a linguagem Java, a biblioteca *jxl.jar* (para ler as planilhas e obter os dados) e a biblioteca *sqldb4.jar* (para realizar a conexão e a inclusão no banco de dados no SQL Server).

Mesmo utilizando automatização para incluir estes dados no banco, algumas dificuldades foram enfrentadas para viabilizar a inclusão dos dados. Os dados referentes à fenologia dos indivíduos tinham duas inconsistências que precisaram ser tratadas.

A primeira diz respeito à presença de quatro colunas extras que apareciam ao final de cada ano com dados de análises estatísticas para cada indivíduo e espécie. Com a criação do banco de dados, estes dados deverão ser gerados a partir de consultas e processamento de algoritmos de análise de séries temporais e armazenados no banco na tabela de dados de ciclo (*DataCycle*) do módulo de fenologia. Por isso, estas colunas foram desconsideradas na inclusão.

O segundo problema refere-se à presença de símbolos nas planilhas para identificar indivíduos mortos. O símbolo utilizado para identificar estes indivíduos é uma cruz. Desta forma, todos os indivíduos que possuem esta cruz em alguma data, não possui

---

<sup>1</sup><http://www.w3schools.com/sql/>. Data de acesso: 7 de janeiro de 2014.

dados de coleta posteriores, ficando as colunas de fenologia vazias. Contudo, colunas vazias também podem representar indivíduos perdidos ou algum erro na observação no campo. Isto poderia gerar um erro no algoritmo de inclusão destes dados. Definiu-se, por isso, junto aos biólogos que todos os indivíduos que estivessem mortos passassem a possuir o status de ‘Morto’, e ao invés da cruz, fossem substituídos por ‘M’. O mesmo foi feito para indivíduos conhecidos como perdidos, cujo status ficou definido como ‘Perdido’ por ‘P’. A Figura 4.2 mostra um exemplo da planilha original de dados, e o mesmo dado substituído por M.

0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	†		
0	0	0	0	0	0	0	0	0	†			
0	0	0	0	0	0	0	0	0	0	1	1	1

0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	M	M	M
0	0	0	0	0	0	0	0	0	0	M	M	M
0	0	0	0	0	0	0	0	0	0	1	1	1

Figura 4.2: Exemplo de dados de indivíduos mortos representados por uma cruz e as colunas posteriores vazias, e a troca destes símbolos pela letra ‘M’ de ‘Morto’.

O período selecionado para a inclusão dos dados de fenologia foi de setembro de 2004 à dezembro de 2007, restando assim os dados do período de janeiro de 2008 à presente data para serem incluídos. Com base em outras pesquisas realizadas pelo grupo de fenologia da UNESP, outras informações foram incluídas no banco, como os dados de síndrome de dispersão das espécies.

Há ainda uma outra forma que o grupo usa para coletar a fenologia de frutos, que é feita por contagem. Os dados desta pesquisa – que também estavam em planilhas, mas separadas da coleta de fenologia por intensidade – incluíam espécies e indivíduos de mesma localidade da fenologia por intensidade, porém, ao incluir estes dados, notou-se que alguns indivíduos tinham nomes diferentes – devido à atualização na literatura na espécie ao qual ele pertence – e alguns já não existem mais, e outros novos que são cobertos apenas na pesquisa de contagem. Logo, o número de espécies passou para 118 e de indivíduos para 2166. Isto reforça ainda mais o problema da falta de interoperabilidade das planilhas e consequentemente dos dados.

## 4.3 Consultas

Após a inclusão dos dados, iniciou-se a etapa de validação e exploração do banco de dados implementado. Para tanto, várias consultas foram definidas em conjunto com os biólogos do Laboratório de Fenologia da UNESP, como os exemplos anteriormente citados na seção 3.4. O objetivo desta seção é apresentar a implementação, em SQL, das consultas consideradas mais importantes, seguindo a mesma divisão explicada naquela seção que divide as consultas em simples e complexas.

Um ponto importante a ser ressaltado é que do ponto de vista dos biólogos a complexidade das consultas diverge da definição dada neste trabalho. Aqui as consultas complexas foram definidas do ponto de vista da implementação das consultas, isto é, da maneira como elas devem ser criadas para obter os resultados desejados no banco de dados. Para os biólogos, a complexidade se dá de acordo com as dificuldades enfrentadas para integrar e relacionar os dados, assim como fazer a análise dos ciclos das fenofases.

Uma consulta simples para um biólogo, por exemplo, é verificar qual o valor da fenofase botão no mês de setembro de 2005. Uma consulta complexa é analisar os ciclos e obter informações de pico no ano ou dentro do ciclo, como por exemplo, verificar quando (em que mês) um determinado indivíduo teve maior produção de flor dentro de seu ciclo ou no período de um ano. Muitas vezes, as consultas complexas para eles, são processamentos que devem ser realizados em cima dos dados que são fornecidos pelo banco de dados.

Desta forma, o banco apresentado também dá suporte para que as consultas complexas do ponto de vista dos biólogos sejam realizadas. Exemplos de consultas implementadas para este banco, de acordo com o modelo apresentado, são mostradas nas seções seguintes.

### 4.3.1 Consultas Simples

Um exemplo de consulta simples, muito importante para os pesquisadores em fenologia, é saber todas as espécies com as quais trabalham. Uma forma de mostrar este resultado em SQL seria:

```
SELECT * FROM Specie
```

Esta consulta exibe todos os atributos das espécies cadastradas. Outra forma seria informar no SQL os atributos desejados, como por exemplo:

```
SELECT idSpecie, family, genus, specieName FROM Specie
```

Pode-se ainda refinar a consulta usando *alias* para os nomes das tabelas (representados no SQL por 'AS'):

```
SELECT idSpecie AS Id, family AS 'Família',  
(genus + ' ' + specieName) AS 'Nome Espécie', author FROM specie
```

O resultado desta consulta é mostrado na Figura 4.3.

	Id	Família	Nome Espécie	author
1	1	Annonaceae	Annona coriacea	NULL
2	2	Annonaceae	Duguetia furfuracea	NULL
3	3	Annonaceae	Xylopia aromatica	NULL
4	4	Apocynaceae	Aspidosperma sp.	NULL
5	5	Apocynaceae	Aspidosperma tomentosum	NULL
6	6	Araliaceae	Schefflera vinosa	NULL
7	7	Arecaceae	Syagrus flexuosa	NULL
8	8	Asteraceae	Bacharis dracunculifolia	NULL
9	9	Asteraceae	Chromalaena laevigata	NULL
10	10	Asteraceae	Gochnatia barrosii	NULL
11	11	Asteraceae	Gochnatia pulchra	NULL
12	12	Asteraceae	Piptocarpha rotundifolia	NULL

Figura 4.3: Exemplo de consulta simples: obter espécies cadastradas no banco de dados. SQL usando *alias*.

Outro exemplo de consulta simples é saber quantos indivíduos estão cadastrados no banco:

```
SELECT COUNT(idIndividual) FROM Individual
```

Esta consulta utiliza uma função do SQL (*count*), e o resultado é apresentado pela Figura 4.4.

	(Nenhum nome de coluna)
1	2166

Figura 4.4: Exemplo de consulta simples: obter total de indivíduos cadastrados no banco de dados. SQL usando *count*.

Mais um exemplo, desta vez relacionado a aspectos fenológicos, é consultar o observador responsável pela coleta dos dados de fenologia (todas as fenofases) do indivíduo 5 na data de 25 de setembro de 2005:

```
SELECT observer FROM dataphenology WHERE idIndividual = 5
and CAST(watchDate as Date) = '2005-09-25'
```

Esta consulta retorna para cada fenofase o observador responsável. Neste exemplo, são retornados seis vezes o nome da bióloga Bruna, que está como observadora responsável

neste dia por este indivíduo. Para resolver este problema, pode-se usar a palavra chave DISTINCT:

```
SELECT DISTINCT observer FROM dataphenology WHERE idIndividual = 5
and CAST(watchDate as Date) = '2005-09-25'
```

O uso da palavra DISTINCT na consulta faz com que o resultado contenha apenas um nome, como é exibido pela Figura 4.5.

	observer
1	Bruna

Figura 4.5: Exemplo de consulta simples: obter observador responsável pela fenologia do indivíduo 5, na data de 25/09/2005. SQL usando *distinct*.

Outro exemplo, ainda em fenologia, e de interesse dos biólogos envolvidos no projeto, e que do ponto de vista deles pode ser considerado complexo, é saber quantos indivíduos floresceram (tiveram fenofases de botão e antese diferente de 0) no período de dados cadastrados no banco. Supondo que se saiba que a fenofase botão tem seu identificador único igual a 1 e a fenofase antese igual a 2 (na tabela *Phenofase*), esta consulta seria especificada da seguinte forma:

```
SELECT idPhenofase, COUNT(DISTINCT idIndividual) AS 'qtd'
FROM DataPhenology
WHERE (idPhenofase = 1 OR idPhenofase = 2)
AND value IS NOT NULL AND value != '0'
GROUP BY idPhenofase
```

O resultado para esta consulta seria igual ao da Figura 4.6-(a), que significa que 1473 indivíduos tiveram botão, e 1220 indivíduos tiveram flor. Para saber quantos desses indivíduos floresceram (tiveram botão e flor), então:

```
SELECT COUNT(DISTINCT idIndividual) AS 'qtd' FROM DataPhenology
WHERE (idPhenofase = 1 OR idPhenofase = 2)
AND value IS NOT NULL AND value != '0'
```

O resultado é apresentado pela Figura 4.6-(b). Este resultado sugere que há uma inconsistência que deve ser verificada pelos biólogos (devido à divergência dos valores).

Muitas outras consultas simples podem ser definidas. Dentro do contexto apresentado, qualquer consulta para obter os atributos de qualquer tabela do banco de dados, desde que esta seja única na busca pelo dado, é considerada uma consulta simples.

	idPhenofase	qtd
1	1	1473
2	2	1220

(a)

	qtd
1	1510

(b)

Figura 4.6: Exemplo de consulta simples: obter observador responsável pela fenologia do indivíduo 5, na data de 25/09/2005. SQL usando *distinct*.

### 4.3.2 Consultas Complexas

Um exemplo de consulta complexa, muito utilizada e importante, é a identificação de cada indivíduo tanto taxonomicamente quanto à sua localização. Assim, se os biólogos querem, por exemplo, saber qual a espécie o indivíduo pertence e todos os seus dados de localização, a consulta fica:

```
SELECT i.idIndividual, i.individualNumber AS 'Número Indivíduo',
spe.family AS Família, spe.genus+ ' ' + spe.specieName as 'Espécie',
s.siteName AS Local, st.state as Estado, c.city AS Cidade,
a.areaName AS 'Área', sba.subAreaName AS SubArea,
sm.samplingMethodName AS 'Método Amostragem',
se.samplingEnvironmentName AS 'Ambiente',
sae.samplingAreaEnvironmentName AS 'Transecto'
FROM individual i
JOIN specie spe ON spe.idSpecie = i.idSpecie
JOIN site s ON s.idSite = i.idSite
JOIN SamplingAreaEnvironment sae ON
sae.idSamplingAreaEnvironment = s.idSamplingAreaEnvironment
JOIN SamplingEnvironment se ON
se.idSamplingEnvironment = sae.idSamplingEnvironment
JOIN SamplingMethod sm ON sm.idSamplingMethod = se.idSamplingMethod
JOIN SubArea sba ON sba.idSamplingMethod = se.idSamplingMethod
JOIN Area a ON a.idArea = sba.idArea
JOIN City c ON c.idCity = a.idCity
JOIN State st ON st.idState = c.idState
```

O resultado desta consulta é similar ao exibido na Figura 4.7. Esta consulta exige junções entre todas as tabelas do módulo de localização, com espécie e indivíduo para obter os dados mostrados.

idIndividual	Número Individual	family	Espécie	Local	Estado	Cidade	Área	Sub-Área	Método Amostragem	Ambiente	Transecto
1	1	Vochysiaceae	Qualea grandiflora	BL11	São Paulo	Itapina	Boteho	Boteho	Transecto	Borda	Borda Leste
2	2	Myrtaceae	Myrcia guianensis	BL11	São Paulo	Itapina	Boteho	Boteho	Transecto	Borda	Borda Leste
3	3	Myrtaceae	Myrcia bala	BL11	São Paulo	Itapina	Boteho	Boteho	Transecto	Borda	Borda Leste
4	4	Rubiaceae	Tocoyena formosa	BL11	São Paulo	Itapina	Boteho	Boteho	Transecto	Borda	Borda Leste
5	5	Rubiaceae	Tocoyena formosa	BL11	São Paulo	Itapina	Boteho	Boteho	Transecto	Borda	Borda Leste
6	6	Annonaceae	Xylopa aromatica	BL11	São Paulo	Itapina	Boteho	Boteho	Transecto	Borda	Borda Leste
7	7	Verbenaceae	Agpphila fruticosa	BL11	São Paulo	Itapina	Boteho	Boteho	Transecto	Borda	Borda Leste
8	8	Annonaceae	Xylopa aromatica	BL11	São Paulo	Itapina	Boteho	Boteho	Transecto	Borda	Borda Leste
9	9	Annonaceae	Xylopa aromatica	BL11	São Paulo	Itapina	Boteho	Boteho	Transecto	Borda	Borda Leste
10	10	Fabaceae	Andenanthus foliata	BL11	São Paulo	Itapina	Boteho	Boteho	Transecto	Borda	Borda Leste
11	11	Annonaceae	Xylopa aromatica	BL11	São Paulo	Itapina	Boteho	Boteho	Transecto	Borda	Borda Leste
12	12	Malvaceae	Boraginaceae campestris	BL11	São Paulo	Itapina	Boteho	Boteho	Transecto	Borda	Borda Leste

Figura 4.7: Exemplo de consulta complexa: obter dados de localização e taxonômicos dos indivíduos. SQL usando *join*.

Com relação às consultas complexas com dados temporais, um exemplo é consultar todos os valores coletados para a fenofase brotamento, do período de setembro de 2004 à dezembro de 2007 para indivíduos da espécie *Qualea grandiflora*:

```
SELECT dp.idIndividual, individualNumber,
CONVERT(varchar(10), watchDate, 103), value
FROM dataPhenology dp
JOIN individual i ON i.idIndividual = dp.idIndividual
JOIN specie s ON s.idSpecie = i.idSpecie
JOIN phenofase ph ON ph.idPhenofase = dp.idPhenofase
WHERE
ph.idPhenofase = (SELECT idPhenofase
FROM phenofase WHERE phenofaseName = 'Brotamento')
AND CAST(watchDate as Date) >= '2004-09-25'
AND CAST(watchDate as Date) <= '2007-12-25'
AND s.genus='Qualea' and s.specieName='grandiflora'
```

Esta consulta gera o resultado apresentado pela Figura 4.8.

Outro exemplo é consultar todas as fenofases de um indivíduo em uma data determinada. Por exemplo, consultar a fenologia do indivíduo *Myrcia guianensis*, cujo identificador único é 2, para o período que se inicia em 13 de setembro de 2004, fica:

```
SELECT CONVERT(varchar(10), watchDate, 103) AS Data, dp.idPhenofase,
ph.phenofaseName AS Fenofase, value as 'Valor Fenologia'
FROM dataPhenology dp
```

	idIndividual	individualNumber	(Nenhum nome de coluna)	value
1	1	1	25/09/2004	1
2	132	151	25/09/2004	1
3	174	200	25/09/2004	1
4	194	224	25/09/2004	1
5	196	226	25/09/2004	1
6	275	317	25/09/2004	1
7	282	327	25/09/2004	1
8	415	499	25/09/2004	1
9	426	511	25/09/2004	1
10	436	524	25/09/2004	1
11	453	542	25/09/2004	1
12	474	567	25/09/2004	1

Figura 4.8: Exemplo de consulta complexa: obter dados de localização e taxonômicos dos indivíduos. SQL usando *join*.

```

JOIN Phenofase ph ON ph.idphenofase = dp.idphenofase
WHERE idIndividual = 2 AND
(dp.idPhenofase = (SELECT idPhenofase
FROM phenofase where replace(phenofaseName, 'ã', 'a') = 'Brotamento')
OR dp.idPhenofase = (SELECT idPhenofase
FROM phenofase where replace(phenofaseName, 'ã', 'a') = 'Botao')
OR dp.idPhenofase = (SELECT idPhenofase
FROM phenofase where replace(phenofaseName, 'ã', 'a') = 'Antese')
OR dp.idPhenofase = (SELECT idPhenofase
FROM phenofase where replace(phenofaseName, 'ã', 'a') = 'Fruto Maduro')
OR dp.idPhenofase = (SELECT idPhenofase
FROM phenofase where replace(phenofaseName, 'ã', 'a') = 'Fruto Imaturo')
OR dp.idPhenofase = (SELECT idPhenofase
FROM phenofase where replace(phenofaseName, 'ã', 'a') = 'Queda Foliar')
)
AND (CAST(watchDate as Date) >= '2004-09-13' )
ORDER BY watchDate, ph.phenofaseName

```

O resultado é apresentado parcialmente pela Figura 4.9.

Há ainda os exemplos que envolvem os dados climáticos e as imagens, que são obtidos com uma frequência diferente dos dados de fenologia: diariamente. A correlação destes dados permite responder a questões importantes na área de biodiversidade do tipo:

	Data	idPhenofase	Fenofase	Valor Fenologia
1	25/09/2004	2	Antese	0
2	25/09/2004	1	Botão	1
3	25/09/2004	5	Brotamento	1
4	25/09/2004	4	Fruto Imaturo	0
5	25/09/2004	3	Fruto Maduro	0
6	25/09/2004	6	Queda Foliar	0
7	25/10/2004	2	Antese	1
8	25/10/2004	1	Botão	1
9	25/10/2004	5	Brotamento	1
10	25/10/2004	4	Fruto Imaturo	0
11	25/10/2004	3	Fruto Maduro	0
12	25/10/2004	6	Queda Foliar	0

Figura 4.9: Exemplo de consulta complexa: obter dados de localização e taxonômicos dos indivíduos. SQL usando *join*.

1. ‘Se há variação do canal verde (extraído do padrão RGB) nas imagens, esta pode estar relacionada ao período de brotamento da espécie?’
2. ou o contrário: ‘Se há variação do canal vermelho (extraído do padrão RGB), esta pode estar relacionada ao período de queda foliar da espécie?’
3. ‘Será que quando inicia a fenofase de brotamento para uma espécie, também há aumento da precipitação? E quanto a variação de temperatura?’
4. ‘O pico de brotamento coincide com a maior média do canal verde (do RGB)?’
5. ‘Qual a época de pico das fenofases para indivíduo/espécie dentro de um ano? Qual a época de pico das fenofases para indivíduos/espécies dentro do seu ciclo?’
6. ‘Qual a duração de um ciclo para indivíduos/espécies?’

É neste ponto que a visão de consultas complexas dos especialistas de fenologia faz sentido. Para responder estas questões, não são necessários apenas os dados armazenados no banco, mas sim de um conjunto de funções que processam as imagens e correlacionam os dados obtidos deste processamento com os dados de fenologia e clima obtidos por consultas no banco de dados.

Por exemplo, para a primeira questão, alguns passos são necessários para que os especialistas possam fazer a avaliação. Primeiro o banco precisa retornar a imagem de máscara

da espécie determinada e as imagens do período informado, supondo que a espécie seja *Myrcia guianensis*, a consulta é simples:

```
SELECT * FROM imageMaskSpecie WHERE idSpecie = 89
```

Esta consulta retornará o local da máscara desta espécie. Também o banco deverá retornar todas as imagens para um período, supondo que seja o período de 09 de janeiro de 2012 à 09 de janeiro de 2013, a consulta também será simples:

```
SELECT * FROM image
WHERE CONVERT(VARCHAR(19), dateHour, 120)
BETWEEN '2012-09-01' and '2013-09-01'
```

Esta consulta requer que a data seja convertida para o formato yyyy-mm-dd hh:mi:ss, pois as imagens são obtidas várias vezes durante o dia. Para tanto, a função CONVERT do SQL Server é utilizada passando o parâmetro que representa o formato desejado, no caso, 120. O resultado listará todas as imagens obtidas no período solicitado. Uma vez listadas estas imagens, elas deverão ser obtidas por um algoritmo e processadas em conjunto com a máscara obtida anteriormente para se obter as médias dos canais RGB. Isto, atualmente, é feito por um aplicativo desenvolvido na linguagem Java, plataforma *desktop*. As médias geradas (representando o mês), deverão ser comparadas com os valores obtidos da fenologia dos indivíduos pertencentes à espécie *Myrcia guianensis*. A consulta será:

```
SELECT dp.idindividual, CONVERT(varchar(10), watchDate, 103) AS Data,
value as 'Valor Fenofase Brotamento'
FROM dataPhenology dp
JOIN Phenofase ph ON ph.idphenofase = dp.idphenofase
JOIN individual i ON i.idIndividual = dp.idIndividual
WHERE i.idspecie = 89
AND dp.idPhenofase = (SELECT idPhenofase
FROM phenofase where replace(phenofaseName, 'ã', 'a') = 'Brotamento')
AND (CAST(watchDate as Date) between '2012-09-01' and '2013-09-01')
ORDER BY dp.idindividual, watchDate
```

O resultado conterà, para todos os indivíduos da espécie, a data e o respectivo valor da fenofase brotamento para aquela data. Dado que a observação da fenologia é feita uma vez por mês, este dado deverá ser comparado com as médias mensais obtidas anteriormente dos canais verde das imagens, e a questão então poderá ser analisada pelos especialistas.

A questão dois é similar à primeira questão, e pode ser tratada da mesma forma, exceto que o valor da fenologia, para todos os indivíduos, a ser consultado é o da fenofase queda foliar.

Para a terceira questão, não são necessárias as etapas relacionadas às imagens. Neste caso, é preciso buscar os dados de fenologia para todos os indivíduos da espécie referente à fenofase brotamento para o período desejado. Supondo que o exemplo ainda seja a espécie *Myrcia guianensis* e o período 09 de janeiro de 2012 à 09 de janeiro de 2013, a consulta da fenologia será a mesma apresentada anteriormente. O que muda agora são as consultas referentes ao clima, que deve retornar a média mensal dos valores de temperatura e precipitação. A consulta para este exemplo, fica:

```
SELECT DATEPART(MONTH, CAST(readDateHour AS date)) as month,
DATEPART(YEAR, CAST(readDateHour AS date)) as year,
AVG(temperature) as temperature, AVG(rain) as rain
FROM (SELECT readDateHour, 'temperature' = case
st.sensorType when 'Temperatura' then cd.value end,
'rain' = case st.sensorType when 'pluviômetro' then
cd.value end FROM climateDate cd JOIN sensor s ON
s.idsensor = cd.idSensor JOIN SensorType st ON
st.idSensorType = s.idSensorType
WHERE (CONVERT(VARCHAR(19), cd.readDateHour, 120) > '2012-09-25 12:00:00'
AND CONVERT(VARCHAR(19), cd.readDateHour, 120) < '2013-09-25 12:00:00')
AND ((st.idSensorType = (SELECT idSensorType FROM sensorType
WHERE SensorType = 'Temperatura'))
OR (st.idSensorType = (SELECT idSensorType FROM sensorType
WHERE replace(sensorType, 'ô', 'o') = 'Pluviometro')))) as climate
GROUP BY DATEPART(MONTH, cast(readDateHour as date)),
DATEPART(YEAR, cast(readDateHour as date))
ORDER BY DATEPART(MONTH, cast(readDateHour as date)),
DATEPART(YEAR, cast(readDateHour as date))
```

Esta consulta usa função DATEPART permite obter da data apenas o mês ou ano. Assim, ela permite que a média seja retornada por mês/ano. Similar a esta consulta, as médias de precipitação podem ser calculadas. Com estes valores, os especialistas em fenologia poderão usar as funções de correlações de variáveis usuais para analisar a questão dada.

Por fim, as últimas três questões são exemplos em que os dados fenológicos deverão ser avaliados a partir de algoritmos que analisam séries temporais, sendo a função do banco fornecer estas séries para que a busca pelos picos possam ser realizadas. Atualmente, fazer o levantamento das informações de pico é um dos pontos mais críticos das pesquisas para os especialistas, pois são feitos manualmente e demandam muito tempo.

## 4.4 Aplicação *Web*

Uma aplicação *Web* foi desenvolvida na ferramenta de desenvolvimento Netbeans<sup>2</sup>, usando a linguagem Java<sup>3</sup>, com a biblioteca de componentes primefaces-3-4-2.jar<sup>4</sup> do Java Server Faces (JSF), a biblioteca de conexão com o banco sqldb4.jar e a biblioteca de tema do primefaces redmond.jar.

Esta aplicação foi desenvolvida com o objetivo de prover um protótipo inicial para manipular o banco por meio de uma interface gráfica. Entretanto, não foram utilizadas técnicas de Interface-Humano-Computador (IHC) na especificação e implementação deste protótipo. Atividades de pesquisa com esta finalidade são deixadas para trabalhos futuros.

O intuito foi implementar algumas das funcionalidades solicitadas ao mesmo tempo que as consultas de interesse foram identificadas. O objetivo é destacar a diferença entre o que o sistema pode processar e o que o banco de dados pode fornecer.

As interfaces foram desenvolvidas usando um tema padrão presente na biblioteca que implementa os componentes JSF. Existem outros padrões e a escolha foi realizada apenas para mostrar a possibilidade de se utilizar cores para definir a identidade do e-phenology.

A Figura 4.10 apresenta a interface inicial da aplicação. Esta interface simula a conexão de um usuário ao sistema, já que funcionalidades relacionadas ao gerenciamento do controle de acesso dos usuários ainda não foi definida para o projeto. Dado que o projeto e-phenology pretende abranger outras equipes e áreas de estudos, ainda não está claro como o controle de usuários deverá ser realizado. Por isso, a interface apresenta uma mensagem fixa.

As Figuras 4.11 e 4.12 apresentam a interface a partir da qual se pode consultar todas as espécies cadastradas no banco de dados (como apresentado em consultas simples). Podem ser feitos filtros a partir dos cabeçalhos das colunas, facilitando a busca por parte dos especialistas de fenologia. Assim como os dados da espécie selecionada podem ser visualizados pelos botões e fazer a alteração da mesma.

As Figuras 4.13 e 4.14 apresentam as interfaces de cadastro de espécies. Ainda não há tratamento para famílias, gêneros ou dados informados de forma incorreta. A interface permite a associação de uma imagem a espécie, que terá que ser armazenada no servidor de dados e seu diretório associado aos atributos da espécie. Além disso, na mesma interface, podem ser incluídos os dados ecológicos referentes à espécie: dados de deciduidade (frequência da queda de folhas), folha, flor e frutos podem ser associados à espécie, e estes dados só podem existir se estiverem relacionados a uma espécie.

As Figuras 4.15 e 4.16 apresentam a interface a partir da qual se pode consultar todos

---

<sup>2</sup><https://netbeans.org/>. Data de acesso: 22/06/2013.

<sup>3</sup><http://www.oracle.com/br/technologies/java/overview/index.html>. Data de acesso: 22/06/2013.

<sup>4</sup><http://primefaces.org/>. Data de acesso: 22/06/2013.



Figura 4.10: Tela Início da aplicação.

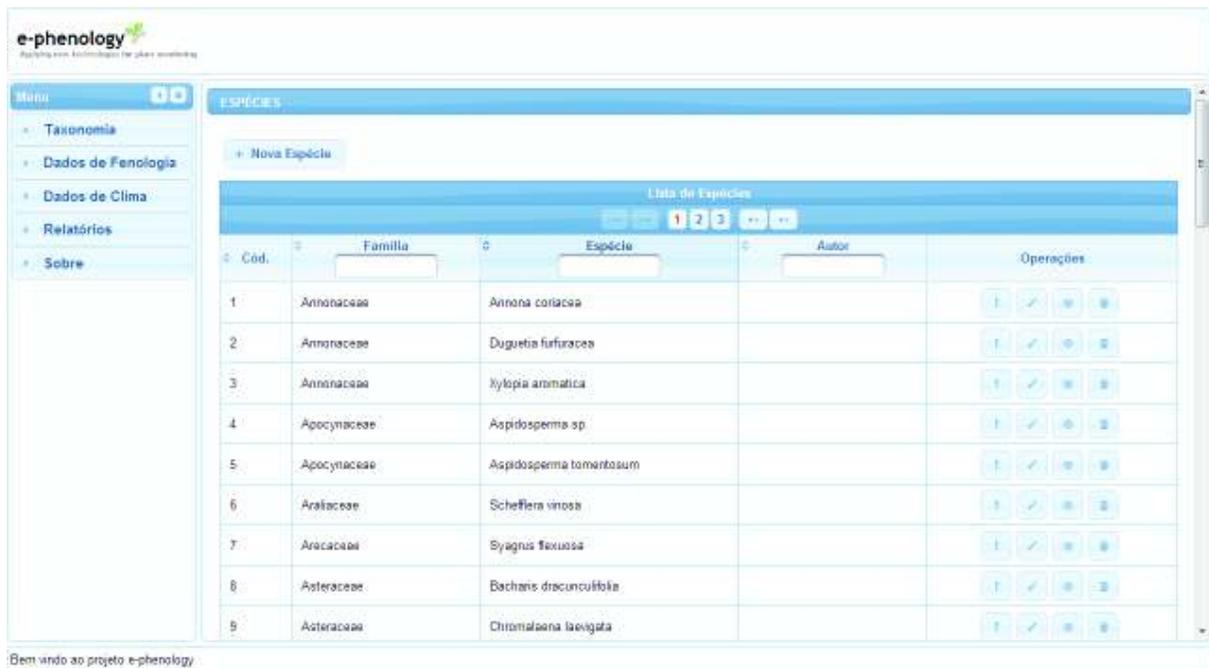


Figura 4.11: Interface de consulta de espécies. Relaciona todas as espécies cadastradas.



Figura 4.12: Interface de consulta de espécies com filtro. Aplicando a busca pelo nome binomial da espécie que iniciam com ‘g’.

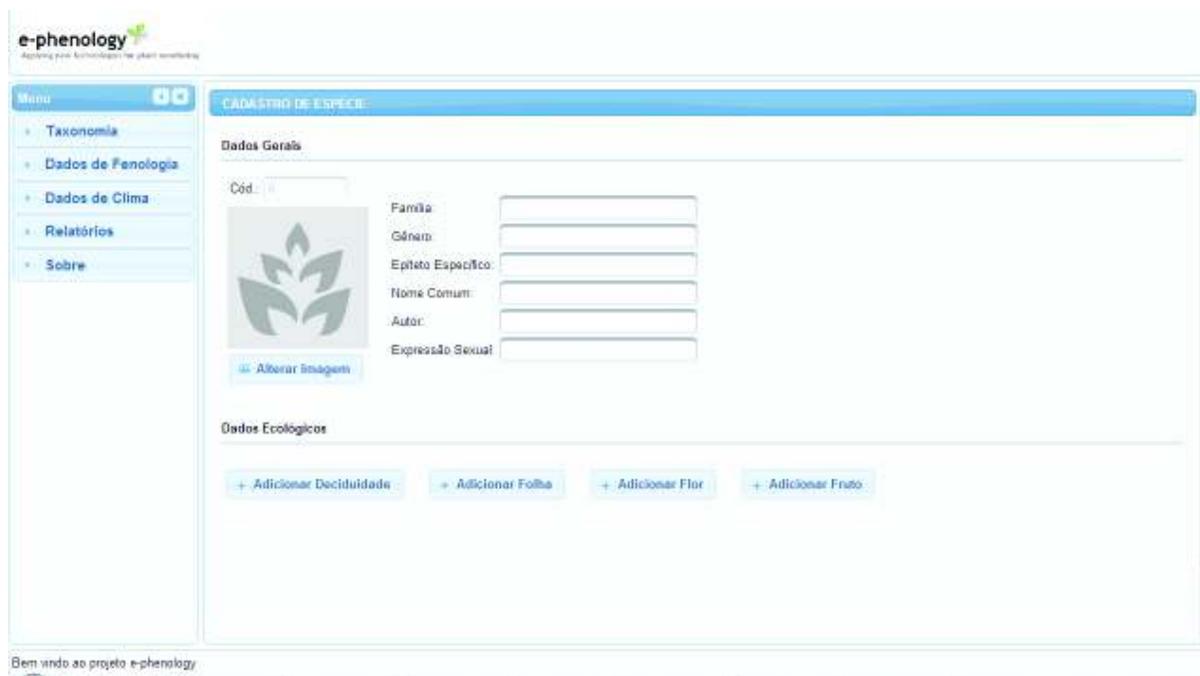


Figura 4.13: Interface de cadastro de espécies.

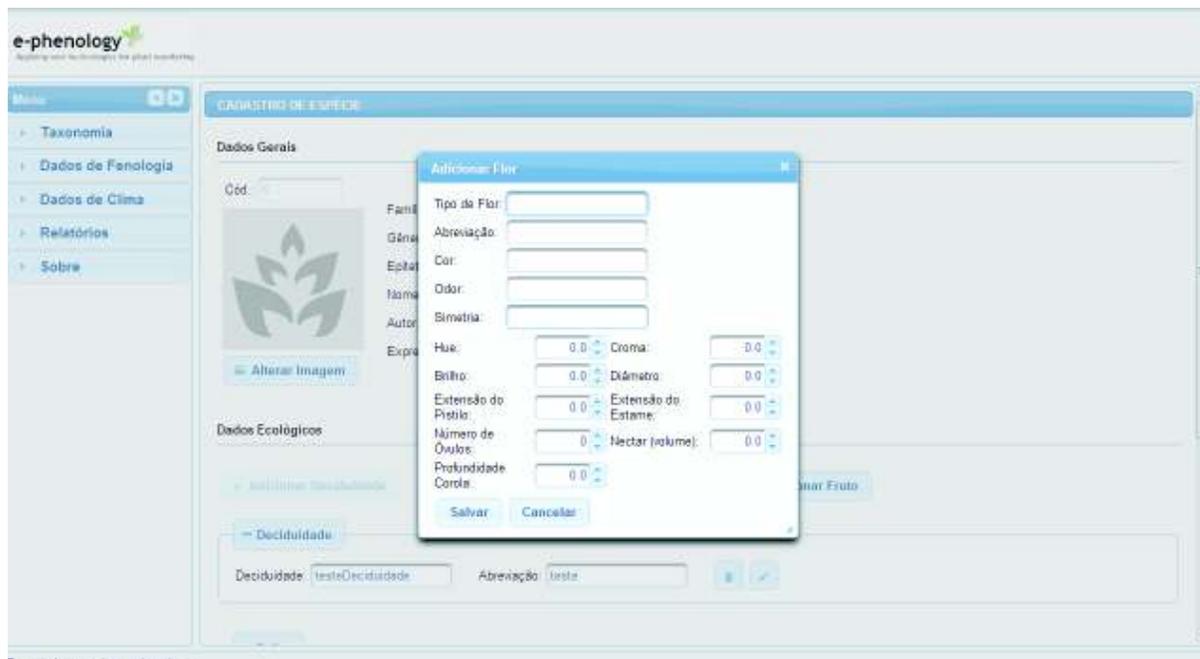


Figura 4.14: Interface de cadastro de espécies. Exemplo para adicionar dados de flor da espécie.

os indivíduos cadastrados no banco, e similarmente à interface de espécies, existem nos cabeçalhos filtros para que os indivíduos sejam mais facilmente encontrados. Assim, se há algum indivíduo específico que se quer consultar a fenologia, por exemplo, o nome da espécie pode ser digitado no respectivo campo e o filtro será aplicado.

A Figura 4.17 mostra que, a partir da seleção de um indivíduo, existe uma opção para consultar a fenologia daquele indivíduo. Ao escolher esta opção de visualizar a fenologia do indivíduo selecionado, a interface de filtros para consultar a fenologia é aberta. A Figura 4.18 mostra os campos desta interface de consulta, que no caso, foi aberta para o indivíduo de número 2 da espécie *Myrcia guianensis*.

É possível selecionar um período (informar a data de início e/ou fim), assim como quais fenofases devem ser exibidas na consulta. No exemplo, ao selecionar a opção de consulta de fenologia na linha do indivíduo de número 2 e pertencente a espécie *Myrcia guianensis*, foi escolhida como data de início 07/09/04, que define a partir de quando os dados das fenofases selecionadas deverão ser buscados e apresentados.

O resultado desta consulta é exibido na Figura 4.19. Vale notar que a forma como os dados são visualizados independem da forma como ele é consultado no banco. Esta consulta é semelhante ao terceiro exemplo de consulta SQL mostrado na seção 4.3.2 e cujo resultado foi exibido na Figura 4.9. Neste caso, os dados foram manipulados para que todas as fenofases fossem visualizadas juntas na mesma data.

Table data for Figure 4.15:

Index	Local	Número	Espécie	Habito	Estrato	Operações
1	BL11	1	Qualea grandiflora			[Icons]
2	BL11	2	Myrcia guianensis			[Icons]
3	BL11	3	Myrcia bella			[Icons]
4	BL11	4	Tocoyena formosa			[Icons]
5	BL11	5	Tocoyena formosa			[Icons]
6	BL11	7	Xylopia aromatica			[Icons]
7	BL11	8	Aegiphila ilotzkiana			[Icons]
8	BL11	9	Xylopia aromatica			[Icons]
9	BL11	10	Xylopia aromatica			[Icons]

Figura 4.15: Interface de consulta de indivíduos. Relaciona todos os indivíduos cadastrados.

Table data for Figure 4.16:

Index	Local	Número	Espécie	Habito	Estrato	Operações
2	BL11	2	Myrcia guianensis			[Icons]
3	BL11	3	Myrcia bella			[Icons]
14	BL11	15	Machaerium acutifolium			[Icons]
20	BL11	25	Miconia rubiginosa			[Icons]
26	BL11	31	Myrcia bella			[Icons]
27	BL11	32	Myrcia guianensis			[Icons]
28	BL11	34	Myrcia guianensis			[Icons]
29	BL11	35	Myrcia bella			[Icons]
31	BL11	36	Myrcia bella			[Icons]

Figura 4.16: Interface de consulta de indivíduos com filtro. Aplicando a busca para todos os indivíduos que pertencem à espécies com o nome binomial iniciado por 'M'.



Figura 4.17: Interface de consulta de indivíduos que permite a visualização da fenologia de um indivíduo selecionado.

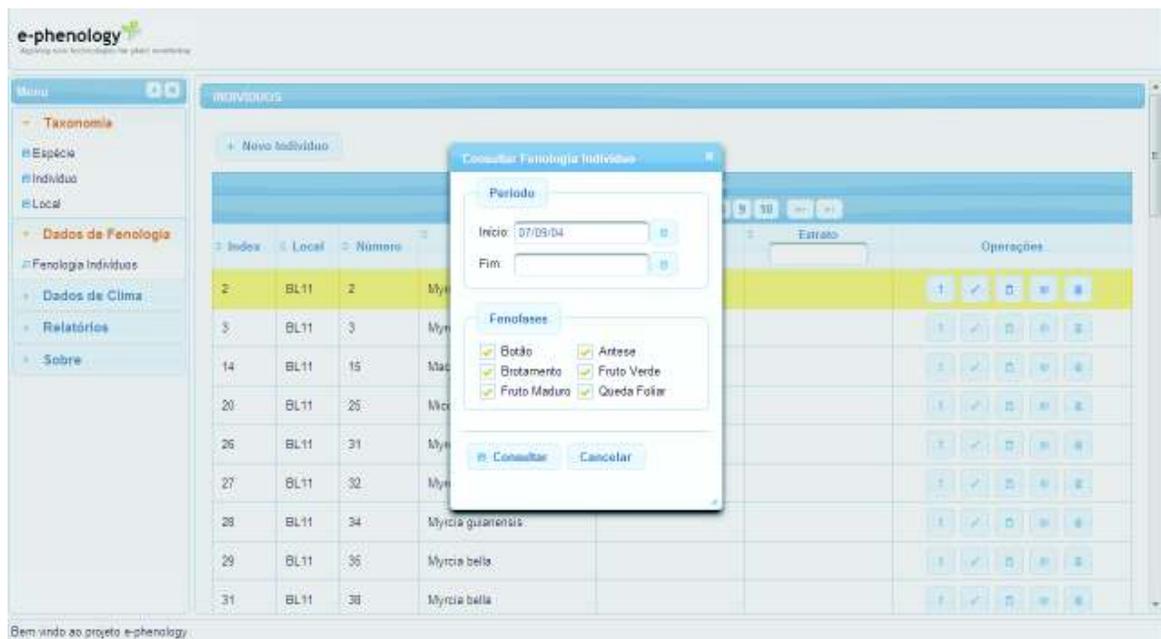


Figura 4.18: Interface de consulta de fenologia de um indivíduo selecionado. No exemplo, o indivíduo selecionado é o indivíduo que tem o número igual a 2 e pertence a espécie *Myrcia guianensis*.

Phenologia de indivíduos 2 - Myrcia guianensis

Data Observação	Botão	Brotamento	Antese	Fruto Imaturo	Fruto Maduro	Queda Foliar
25/09/2004	1	1	0	0	0	0
25/10/2004	1	1	0	0	0	0
25/11/2004	0	0	0	0	0	1
25/12/2004	0	0	0	0	0	1
25/01/2005	0	0	0	0	0	0
25/02/2005	0	0	0	0	0	0
25/03/2005	0	1	0	0	0	1
25/04/2005	0	0	0	0	0	0
25/05/2005	0	1	0	0	0	0
25/06/2005	0	1	0	0	0	0
25/07/2005	1	2	0	0	0	0
25/08/2005	2	1	0	0	0	0
25/09/2005	2	1	0	0	0	0
25/10/2005	0	0	0	2	1	0
25/11/2005	0	0	0	1	1	0
25/12/2005	0	0	0	0	1	0
25/01/2006	0	0	0	0	0	0
25/02/2006	0	0	0	0	0	1
25/03/2006	0	1	0	0	0	0
25/04/2006	0	0	0	0	0	0

Figura 4.19: Interface de resultado da consulta de fenologia de um indivíduo selecionado, apresentando apenas valores de fenologia.

INDIVÍDUOS

+ Novo Indivíduo

Lista de indivíduos

Índice	Local	Número	Espécie	Habito	Estrato	Operações
2	BL11	2	Myrcia guianensis			[1] [✓] [✕] [↶] [↷]
3	BL11	3	Myrcia bella			[1] [✓] [✕] [↶] [↷] [Consultar Fenologia e Clima]
14	BL11	15	Machaerium acutifolium			[1] [✓] [✕] [↶] [↷]
20	BL11	25	Niconia rubiginosa			[1] [✓] [✕] [↶] [↷]
26	BL11	31	Myrcia bella			[1] [✓] [✕] [↶] [↷]
27	BL11	32	Myrcia guianensis			[1] [✓] [✕] [↶] [↷]
28	BL11	34	Myrcia guianensis			[1] [✓] [✕] [↶] [↷]
29	BL11	35	Myrcia bella			[1] [✓] [✕] [↶] [↷]
31	BL11	38	Myrcia bella			[1] [✓] [✕] [↶] [↷]

Figura 4.20: Interface de consulta de fenologia com médias climáticas para um indivíduo selecionado. No exemplo, o indivíduo selecionado é o indivíduo que tem o número igual a 2 e pertence a espécie *Myrcia guianensis*.

De modo similar, a Figura 4.20 mostra que, a partir da seleção de um indivíduo, há uma opção para consultar a fenologia com o clima para aquele indivíduo. Esta opção é similar a de consulta de fenologia, porém permite que sejam escolhidos sensores climáticos cujas médias devem ser mostradas em conjunto com a fenologia.

A Figura 4.21 mostra os campos desta interface de consulta com os sensores. No caso, a interface foi aberta para o indivíduo de número 2 da espécie *Myrcia guianensis*. Para a consulta foram informados: a data inicial 17/08/2004, as fenofases botão e brotamento e o sensor de temperatura.

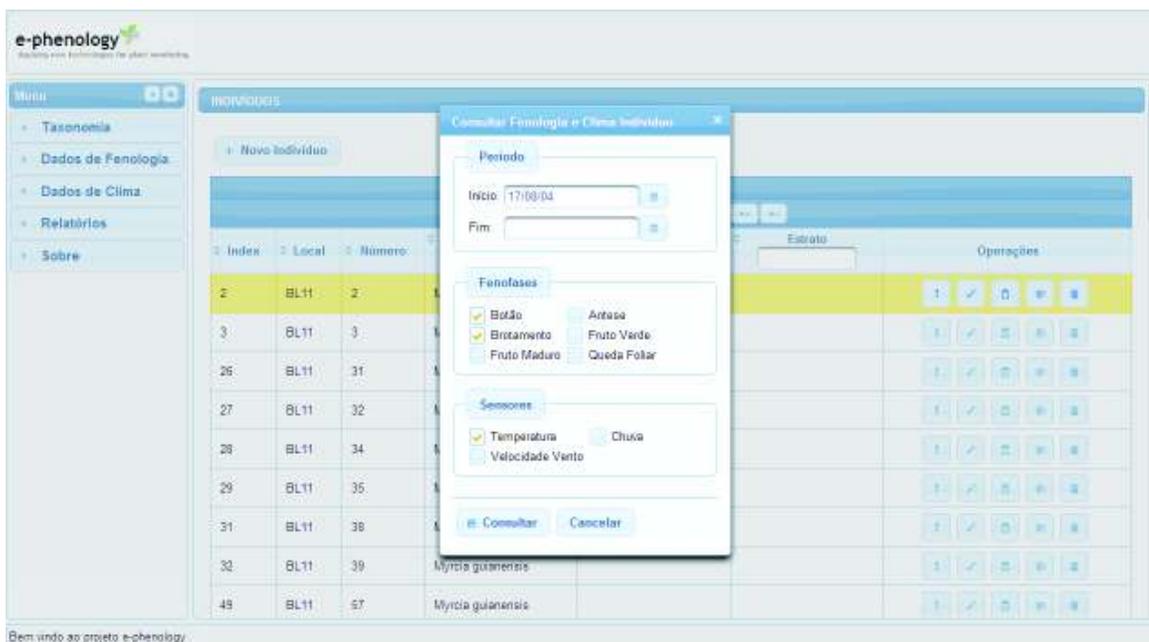
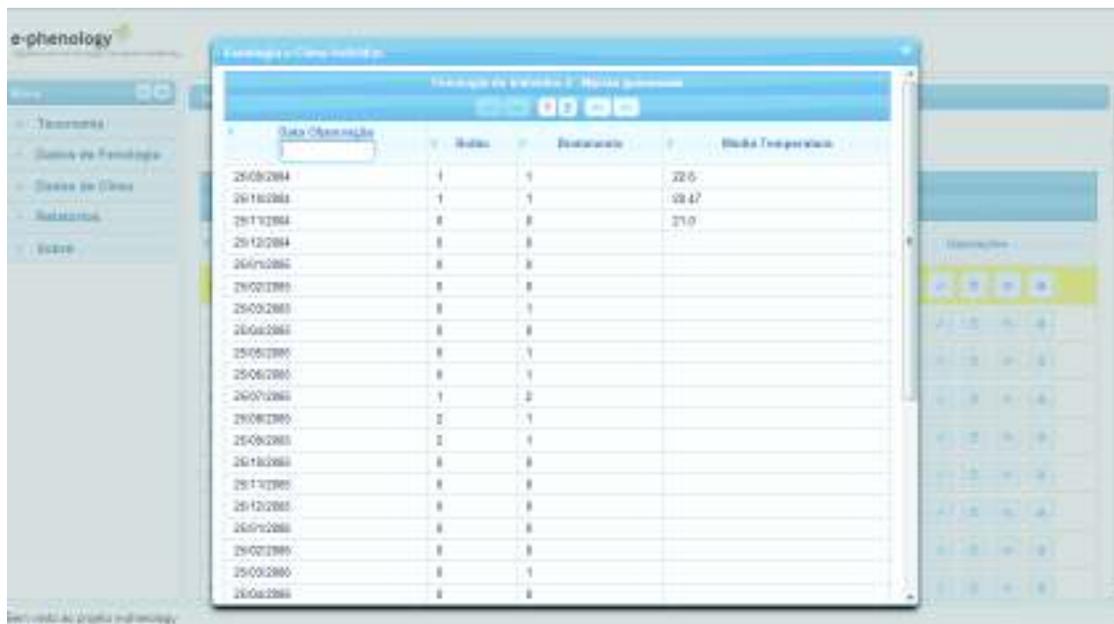


Figura 4.21: Interface de consulta de fenologia com médias climáticas para um indivíduo selecionado. No exemplo, o indivíduo selecionado é o indivíduo que tem o número igual a 2 e pertence a espécie *Myrcia guianensis*, e apenas as fenofases botão e brotamento e o sensor temperatura foram selecionados

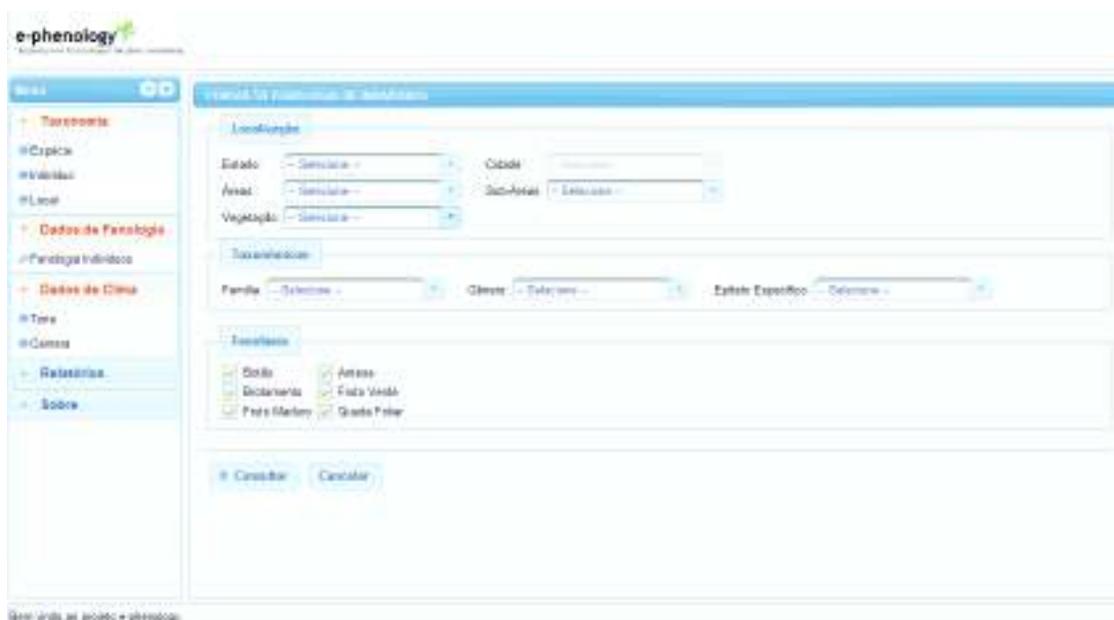
O resultado da consulta é apresentado na Figura 4.22, que exhibe apenas as fenofases selecionadas, no exemplo botão e brotamento, e as médias existentes no banco de dados para o sensor de temperatura, de acordo com o período informado. Os dados de clima foram simulados apenas para realização de testes da aplicação, pois estes só passaram a ser coletados no Cerrado a partir de agosto de 2011.

Por fim, a Figura 4.23 apresenta uma interface de filtros solicitados para realizar a consulta de fenologia a partir de diferentes dados. Por exemplo, pode-se querer consultar a fenologia de todos os indivíduos de uma espécie, ou de uma mesma família, ou da mesma



Data Observação	Botões	Brotamento	Média Temperatura
25/05/2004	1	1	22,5
26/10/2004	1	1	23,47
26/11/2004	0	0	
26/12/2004	0	0	21,0
26/01/2005	0	0	
26/02/2005	0	0	
26/03/2005	0	1	
26/04/2005	0	1	
26/05/2005	1	2	
26/06/2005	2	1	
26/07/2005	0	1	
26/08/2005	0	0	
26/09/2005	0	0	
26/10/2005	0	0	
26/11/2005	0	0	
26/12/2005	0	0	
26/01/2006	0	0	
26/02/2006	0	0	
26/03/2006	0	1	
26/04/2006	0	0	

Figura 4.22: Interface de resultado da consulta de fenologia com médias climáticas para um indivíduo selecionado. Apresenta apenas os valores das fenofases botão e brotamento e o sensor temperaturas selecionados.



e-phenology  
 Ferramenta de coleta e análise de dados fenológicos

Estado:  Cidade:   
 Área:  Sub-Área:   
 Vegetação:

Família:  Gênero:  Espécie Específica:

Fenofases:  
 Botão  Antese  Fruto Verde  Fruto Maduro  Queda Folha

Figura 4.23: Interface de filtros para realizar consultas de fenologia, por espécie, indivíduo ou localização

área, enfim, há várias possibilidades. Esta interface também poderá ser usada para gerar relatórios (com execução de macros a partir do banco).

Paralelamente a esta aplicação, foi desenvolvida uma aplicação *desktop* exclusivamente para processar as imagens. Para essa aplicação, apenas os dados de imagens devem ser fornecidos. Esta funcionalidade, no entanto, ainda não foi integrada ao banco de dados e à aplicação *Web*.



# Capítulo 5

## Conclusões e extensões

Este trabalho apresentou a especificação e a implementação do banco de dados para estudos realizados na área de fenologia de plantas, que deverá apoiar as atividades de pesquisa no âmbito do projeto e-phenology. O projeto e-phenology é uma colaboração envolvendo pesquisadores do Instituto de Computação da UNICAMP e do Instituto de Biosciências da UNESP de Rio Claro, e tem como objetivo criar ferramentas que possibilitem o monitoramento remoto da fenologia nos trópicos brasileiros, o que ainda é algo novo no Brasil.

O banco de dados foi desenvolvido conforme os requisitos e expectativas apontadas pelos especialistas em fenologia de plantas do Laboratório de Fenologia da UNESP. Foram utilizados os dados já existentes para mapear e construir uma base de dados de acordo com estas expectativas, integrando os dados de clima – que até então precisavam ser buscados em outras fontes meteorológicas, dados de fenologia – que por anos foram armazenados em planilhas que dificultam a interoperabilidade dos dados, e dados de imagens – uma novidade para se analisar a fenologia de plantas.

Para melhor compreensão do banco, os dados foram classificados dentro de seis categorias: dados de localização, dados ecológicos, dados taxonômicos, dados fenológicos, dados climáticos e dados de imagens. Estas divisões facilitaram as discussões sobre a modelagem do banco de dados entre os pesquisadores da área de computação e biologia. A implementação do banco e a inclusão dos dados também fez uso da mesma categorização.

Uma das principais dificuldades encontradas durante o processo de implementação diz respeito à inclusão dos dados no banco. As planilhas comumente utilizadas apresentaram alguns problemas como a incompatibilidade de nomes de indivíduos, uso de símbolos para identificar status do indivíduo no dia da coleta (morto ou perdido) e dados de coletas faltantes. Tais problemas precisaram ser tratados e revisados para que fosse possível realizar a inclusão dos dados de forma automática por algoritmos desenvolvidos usando a linguagem Java.

Além da especificação do banco de dados, também foram definidas consultas de maior interesse e importantes para os biólogos, as quais foram apresentadas neste trabalho considerando sua complexidade de implementação. Com a implementação das consultas e o decorrer do projeto e-phenology, novas questões – do ponto de vista fenológico – de biodiversidade surgiram, e um dos objetivos e contribuições deste trabalho é possibilitar que tais questões sejam mais facilmente respondidas.

Alguns exemplos de questões de biodiversidade investigadas dizem respeito à validação do monitoramento remoto por imagens digitais. As imagens obtidas permitem fazer a verificação da alteração da cor das folhas no decorrer dos dias. A variação observada a partir do processamento de imagens pode ser verificada e validada com os dados obtidos no campo, por meio das análises da coleta de dados sobre as fenofases dos indivíduos realizadas mensalmente. Como o volume de dados de estudos em fenologia é grande, o banco de dados tem a característica de facilitar o processo de obtenção de tais dados para fazer a validação do monitoramento.

Outra questão importante apresentada é correlacionar as alterações climáticas com a variação de cor das folhas obtida das imagens e os dados de fenologia, a fim de verificar qual impacto tais alterações tem sobre as fenofases dos indivíduos, das espécies e da comunidade como um todo. Se tais questões puderem ser respondidas a partir de consultas no banco de dados, com o uso de processamento de imagens e com a correlação destes dados, então pode-se dizer que da mesma forma como os estudos de fenologia no solo são importantes para analisar os impactos e mudanças ambientais, a fenologia remota pode contribuir ainda mais para entender estes fenômenos e compreender os ecossistemas.

Outro desafio encontrado no desenvolvimento deste trabalho foi projetar uma base de dados que possibilitasse a inclusão de diferentes metodologias de observação de fenologia e dados de áreas de estudos distintas. Também foi preciso projetar a inclusão de câmeras digitais e torres instaladas em outras localidades com vegetações diferentes.

As principais contribuições esperadas e alcançadas neste trabalho foram:

- especificação e implementação de um banco de dados que pudesse integrar os dados de pesquisas de fenologia de plantas de solo, com análises de imagens e dados climáticos;
- inclusão e validação da base de dados a partir de um estudo de caso do projeto e-phenology, que é a área de Cerrado de Itirapina, do Estado de São Paulo, onde são realizadas pesquisas de fenologia de plantas desde setembro de 2004, e com a proposição do projeto e-phenology está sendo monitorada por imagens remotamente desde agosto de 2011;
- implementação de algumas consultas de interesse dos biólogos para validar o funcionamento do banco de dados;

- implementação de um protótipo inicial, a partir do qual os biólogos podem começar a interagir com o banco de dados e realizar algumas consultas.

Com o andamento do trabalho, notou-se demanda para melhorar o banco de dados, tanto do ponto de vista de implementação, quanto de novas pesquisas. As principais são:

- possibilitar que as características ecológicas sejam associadas, também, aos indivíduos;
- melhorar e adaptar o protótipo *Web* apresentado neste trabalho, integrando-o a ferramenta de processamento de imagens e possibilitando a geração de novas consultas e relatórios por macros;
- estudar possibilidades de integrar os dados das espécies com os catálogos taxonômicos existentes, a fim de prover todos os sinônimos e históricos das espécies para as pesquisas;
- expandir o banco de dados para permitir a inclusão de dados espectrais associados a indivíduos e dados ecológicos;
- especificar e implementar o módulo de controle de acesso de usuários de diferentes pesquisas e estudos;
- estudar formas de viabilizar o banco de dados para que seja usado como fonte de informações em uma rede de fenologia;
- fazer novos estudos de casos com dados das novas áreas que serão incorporadas ao projeto e-phenology, onde novas torres e câmeras estão sendo instaladas;
- gerar novas técnicas de visualização de informação para apresentar os dados consultados de maneiras diferentes, facilitando principalmente a interpretação da correlação de múltiplas variáveis, e aplicar técnicas de análises usadas pelos fenólogos;
- especificar e implementar a modelagem do banco de dados de acordo com o uso de metadados em biologia;
- aplicar mineração de dados e algoritmos de análises de séries temporais para extrair dados estatísticos de fenologia;
- implementar e incorporar consultas que envolvem questões espaço-temporais, como por exemplo, “Qual a localização (*Site*) dos indivíduos da espécie *Myrcia guianensis* em 31 de outubro de 2006?” ou “Qual a distância entre dois indivíduos de mesma espécie?”.

Este trabalho teve a seguinte publicação aceita:

- “Desenvolvimento de um Modelo Conceitual do Banco de Dados para o Projeto e-phenology” [28], apresentado no VI Workshop e-Science do XXXII Congresso da Sociedade Brasileira de Computação, realizado em Curitiba, no Estado do Paraná em julho de 2012. O trabalho apresentou o primeiro modelo conceitual desenvolvido para o banco de dados.

# Referências Bibliográficas

- [1] L. F. Alberti and L P. C. Morelatto. Influência da abertura de trilhas antrópicas e clareiras naturais na fenologia reprodutiva de *gymnanthes concolor*. *Revista Brasileira de Botânica*, 31:53–59, 2008.
- [2] B. Alberton, J. Almeida, R. Henneken, R. S. Torres, A. Menzel, and L. P. C. Morelatto. Near remote phenology: Applying digital images to monitor leaf phenology in a brazilian cerrado savanna. *International Conference on Phenology (Phenology'12)*, 2012.
- [3] J. Almeida, J. A. dos Santos, B. Alberton, R. da S. Torres, and L. P. C. Morelatto. Remote phenology: Applying machine learning to detect phenological patterns in a cerrado savanna. In *eScience*, pages 1–8. IEEE Computer Society, 2012.
- [4] J. Almeida, J. A. dos Santos, B. Alberton, R. da S. Torres, and L. P. C. Morelatto. Applying machine learning based on multiscale classifiers to detect remote phenology patterns in cerrado savanna trees. *Ecological Informatics*, (0):–, 2013.
- [5] M. G. G. Camargo, R. M. Souza, P. Reys, and L. P. C. Morelatto. Effects of environmental conditions associated to the cardinal orientation on the reproductive phenology of the cerrado savanna tree *Xylopia aromatica* (annonaceae). *Annals of the Brazilian Academy of Sciences*, 83:1007–1020, 2011.
- [6] R. K. Colwell. Biota: the biodiversity database manager. *Sinauer Associates*, 1996.
- [7] Federal Geographic Data Committee. Fgdc - federal geographic data committee. <http://www.fgdc.gov>. (Acessado em: 16 de agosto de 2013).
- [8] Microsoft Corporation. Sql server 2012. <http://www.microsoft.com/sqlserver/pt/br/default.aspx>, 2005. (Acessado em: 20 de junho de 2013).
- [9] Oracle Corporation. Oracle database. <http://www.oracle.com/>. (Acessado em: 20 de junho de 2013).

- [10] M. J. Costello and E. V. Berghe. Ocean biodiversity informatics: a new era in marine biology research and management. *Marine Ecology Progress Series*, 316:203–214, 2006.
- [11] R. da S. Torres and A. X. Falcão. Content-Based Image Retrieval: Theory and Applications. *Revista de Informática Teórica e Aplicada*, 13(2):161–185, 2006.
- [12] R. da S. Torres, J. A. M. Zegarra, J. A. dos Santos, C. D. Ferreira, O. A. B. Penatti, F. A. Andaló, and J. G. Almeida Júnior. Recuperação de imagens: Desafios e novos rumos. In *XXXV Seminário Integrado de Software e Hardware (SEMISH)*, pages 223–237, Belém, PA, Brasil, 2008.
- [13] J. Daltio. Um serviço web de ontologias para interoperabilidade em sistemas de biodiversidade. Master’s thesis, Instituto de Computação, Universidade de Campinas, 2007.
- [14] C. J. Date. *Introdução a sistema de banco de dados*. Elsevier, 2004.
- [15] F. F. d’Eça Neves and L. P. C. Morellato. Métodos de amostragem e avaliação utilizados em estudos fenológicos de florestas tropicais. *Acta Botanica Brasileira*, 18(1):99–108, 2004.
- [16] V. Dose and A. Menzel. Bayesian analysis of climate change impacts in phenology. *Global Change Biology*, 10(2):259–272, 2004.
- [17] N. Edelweiss. Bancos de dados temporais: teoria e prática. *XVII Jornada de Atualização em Informática, do XVIII Congresso Nacional da Sociedade Brasileira de Computação*, 2:225–282, 1998.
- [18] R. A. Elmasri and S. B. Navathe. *Fundamentals of Database Systems*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 3rd edition, 1999.
- [19] A. A. Fagundes. Projeto e implementação de um banco de metadados para o sistema de informação de biodiversidade do estado de são paulo. Master’s thesis, Instituto de Computação, Universidade de Campinas, 1999.
- [20] G. S. Fedel. Busca multimodal para apoio à pesquisa em biodiversidade. Master’s thesis, Instituto de Computação, Universidade de Campinas, 2011.
- [21] E. H. Fegraus, S. Andelman, M. B. Jones, and M. Schildhauer. Maximizing the value of ecological data with structured metadata: An introduction to ecological metadata language (eml) and principles for metadata creation. *Bulletin of the Ecological Society of America*, 86(3):158–168, July 2005.

- [22] M. A. Gonçalves. Uso de modelos de hipermídias em bibliotecas digitais para dados geográficos. Master's thesis, Instituto de Computação, Universidade de Campinas, 1997.
- [23] PostgreSQL Global Development Group. Postgre. <http://www.postgresql.org/>, 1996. (Acessado em: 20 de junho de 2013).
- [24] L. C. G. Junior. Uma arquitetura para consultas a repositórios de biodiversidade na web. Master's thesis, Instituto de Computação, Universidade de Campinas, 2007.
- [25] A. A. Kondo. Gerenciamento de rastreabilidade em cadeias produtivas agropecuárias. Master's thesis, Instituto de Computação, Universidade de Campinas, 2007.
- [26] D. R. Maddison and K. S. Schulz. The Tree of Life Web Project, 2007.
- [27] J. E. G. Malaverri. Um serviço de gerenciamento de coletas para sistemas de biodiversidade. Master's thesis, Instituto de Computação, Universidade de Campinas, 2009.
- [28] G. C. Mariano, J. Almeida, B. Alberton, R. da S. Torres, and L. P. C. Morellato. Desenvolvimento de um modelo conceitual de um banco de dados para o projeto e-phenology. In *Anais do XXVIII Congresso da SBC - VI Workshop e-Science*, Curitiba, PR, Brasil, 2012.
- [29] W. K. Michener, J. H. Beach, M. B. Jones, B. Ludäscher, D. D. Pennington, R. S. Pereira, A. Rajasekar, and M. Schildhauer. A knowledge environment for the biodiversity and ecological sciences. *Journal of Intelligent Information Systems*, 29(1):111–126, 2007.
- [30] W. K. Michener, J. W. Brunt, J. J. Helly, T. B. Kirchner, and S. G. Stafford. Non-geospatial metadata for the ecological sciences. *ECOLOGICAL APPLICATIONS*, 7(1):330–342, February 1997.
- [31] C. Mira, P. Feijão, J. Meidanis, T. Duque-Estrada, and C. Joly. Tradução taxonômica: o caso do sinbiota. Relatório técnico, Instituto de Computação, Universidade de Campinas, mar 2013.
- [32] C. V. G. Mira, P. Feijão, T. Duque-Estrada, J. Meidanis, and C. A. Joly. The sinbiota 2.0 biodiversity information system. In *eScience*, pages 142–149. IEEE Computer Society, 2011.

- [33] L. P. C. Morellato. *Phenological data, networks, and research: South America*. Kluwer Academic Publishers, 2003.
- [34] L. P. C. Morellato. *Fenologia de plantas e os efeitos das mudanças climáticas*, pages 181–191. Rima Editora, 2008.
- [35] R. A. Morris, R. D. Stevenson, and W. Haber. An architecture for electronic field guides. *J. Intell. Inf. Syst.*, 29(1):97–110, August 2007.
- [36] P. A. Opler, G. W. Frankie, and H. G. Baker. Rainfall as a factor in the release, timing and synchronization of anthesis by tropical trees and shrubs. *Journal of Biogeography*, 3(3):231–236, 1976.
- [37] P. A. Opler, G. W. Frankie, and H. G. Baker. Comparative phenological studies of treelet and shrubs species in tropical wet and dry forests in the lowlands of costa rica. *Journal of Ecology*, 68:167–188, 1980.
- [38] R. Ortiz. Fenologia de arboles en un bosque semideciduo tropical del estado cojedes. *Acta Botanica Venezuelana*, 16:93–116, 1990.
- [39] C. Parmesan and G. Yohe. A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421:37–42, 01/2003 2003.
- [40] O. A. B. Penatti. Estudo comparativo de descritores para recuperação de imagens por conteúdo na web. Master’s thesis, Instituto de Computação, Universidade de Campinas, 2009.
- [41] C. B. Medeiros R. da S. Torres, M. A. Gonçalves, and E. A. Fox. A Digital Library Framework for Biodiversity Information Systems. *International Journal on Digital Libraries*, 6(1):3–17, February 2006.
- [42] B. Rathcke and E. P. Lacey. Phenological patterns of terrestrial plants. *Annual Review of Ecology and Systematics*, 16:179–214, 1985.
- [43] B. C. Reed, J. F. Brown, D. Vanderzee, T. R. Loveland, J. W. Merchant, and D. O. Ohlen. Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, 5:703–714, 1994.
- [44] G. M. Rego, R. B. Negrelle, and L. P. C. Morellato. *Fenologia: ferramenta para conservação, melhoramento e manejo de recursos vegetais arbóreos*. Embrapa Florestas, 2007.

- [45] C. Rosenzweig, D. Karoly, M. Vicarelli, P. Neofotis, Q. Wu, G. Casassa, A. Menzel, T. L. Root, N. Estrella, B. Seguin, P. Tryjanowski, C. Liu, S. Rawlins, and A. Imeson. Attributing physical and biological impacts to anthropogenic climate change. *Nature*, 453:353–357, 2008.
- [46] M. D. Schwartz. Phenoclimatic Measures Phenology: An Integrative Environmental Science. In Mark D. Schwartz, editor, *Phenology: An Integrative Environmental Science*, volume 39 of *Tasks for vegetation science 34*, chapter 21, pages 331–343. Springer Netherlands, Dordrecht, 2003.
- [47] K-T. Shao, C-I. Peng, Eric Yen, K-C. Lai, M-C. Wang, J. Lin, H. Lee, Y. Alan, and S-Y. Chen. Integration of biodiversity databases in taiwan and linkage to global databases. *Data Science Journal*, 6:2–10, 2007.
- [48] D. C. Talora and L. P. C. Morellato. Fenologia de espécies arbóreas em floresta de planície litorânea do sudeste do Brasil 1. *Revta brasil. Bot*, 3(1966):13–26, 2000.
- [49] B. S. C. Vilar. Processamento de consultas baseado em ontologias para sistemas de biodiversidade. Master’s thesis, Instituto de Computação, Universidade de Campinas, 2009.
- [50] G. R. Walther. Plants in a Warmer World. *Perspectives in Plant Ecology, Evolution and Systematics*, 6(3):169–185, 2003.
- [51] G. R. Walther, E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J-M. Fromentin, O. Hoegh-Guldberg, and F. Bairlein. Ecological responses to recent climate change. *Nature*, 416(6879):389–395, March 2002.
- [52] Stuart Weibel. The Dublin Core: A Simple Content Description Model for Electronic Resources. *Bul. Am. Soc. Info. Sci. Tech.*, 24(1):9–11, 1997.
- [53] H. Xu, Z. Gao, D. Xue, and X. Wu. China national biodiversity information query system. *Journal of Environmental Management*, 56(1):45 – 59, 1999.



# Apêndice A

## Ficha de Coleta de Dados no Campo - Cerrado (Borda Leste)

LEGENDA TABELA:

-  Identifica localização. Cada borda e interior possui uma cor diferente.  
Também identifica indivíduos perdidos - quando preenchida a linha
-  Identifica indivíduos mortos.
-  Destaca observações importantes.

Local	Trans	Individu	Coletor:		Data:					Observações		
			Família	Espécie	Bot	Ante	FV	FM	Brot		Quec	Indiv
BL	01 Bc	1	Vochysiaceae	Qualea grandiflora							1	s/ placa
BL	1	2	Myrtaceae	Myrcia guianensis							2	
BL	1	3	Myrtaceae	Myrcia bella							3	
BL	1	4	Rubiaceae	Tocoyena formosa							4	
BL	1	5	Rubiaceae	Tocoyena formosa							5	placa no chão
BL	1	7	Annonaceae	Xylopia aromatica							7	sem placa, atrás da 5
BL	1	8	Verbenaceae	Aegiphila lhotszkyana							8	sem placa, bem baixa, à direita da 5
BL	1	9	Annonaceae	Xylopia aromatica							9	
BL	1	10	Annonaceae	Xylopia aromatica							10	
BL	1	11	Fabaceae	Anadenanthera falcata							11	ao lado da placa 10
BL	1	12	Annonaceae	Xylopia aromatica							12	
BL	1	13	Malpighiaceae	Banisteriopsis campestris							13	
BL	1	14	Araliaceae	Schefflera vinosa							14	ver base
BL	1	15	Fabaceae	Machaerium acutifolium							15	lado da 14
BL	1	16	Ochnaceae	Ouratea spectabilis							16	baixa, pé 14
BL	1	18	Annonaceae	Xylopia aromatica							18	
BL	1	22	Fabaceae	Bauhinia rufa							22	
BL	1	23	Sapotaceae	Pouteria ramiflora							23	
BL	1	24	Sapotaceae	Pouteria ramiflora							24	
BL	1	25	Melastomataceae	Miconia rubiginosa							25	
BL	1	26	Bignoniaceae	Tabebuia ochracea							26	perdida abr/07
BL	1	27	Erythroxylaceae	Erythroxylum suberosum							27	
BL	1	28	Myrtaceae	Campomanesia pubescens							28	atrás da 27
BL	1	29	Sapotaceae	Pouteria ramiflora							29	
BL	1	30	Erythroxylaceae	Erythroxylum suberosum							30	
BL	1	31	Myrtaceae	Myrcia bella							31	morta jan/2010
BL	1	32	Myrtaceae	Myrcia guianensis							32	
BL	1	34	Myrtaceae	Myrcia guianensis							34	
BL	1	35	Myrtaceae	Myrcia bella							35	
BL	1	36	Fabaceae	Anadenanthera falcata							36	
BL	1	38	Myrtaceae	Myrcia bella							38	
BL	1	39	Myrtaceae	Myrcia guianensis							39	
BL	1	40	Fabaceae	Anadenanthera falcata							40	
BL	1	41	Sapotaceae	Pouteria ramiflora							41	atrás da 38
BL	1	42	Sapotaceae	Pouteria ramiflora							42	lado da 41
BL	1	43	Malpighiaceae	Byrsonima coccolobifolia							43	
BL	1	44	Fabaceae	Bauhinia rufa							44	
BL	1	45	Fabaceae	Bauhinia rufa							45	morta jan/2010
BL	1	46	Lauraceae	Ocotea pulchella							46	
BL	1	47	Annonaceae	Xylopia aromatica							47	
BL	1	48	Fabaceae	Stryphnodendron obovatum							48	na borda
BL	1	50	Sapotaceae	Pouteria torta							50	
BL	1	51	Sapotaceae	Pouteria torta							51	
BL	1	52	Sapotaceae	Pouteria torta							52	direita 50, +p/ dentro
BL	1	53	Sapotaceae	Pouteria torta							53	olhar de dentro para a borda
BL	1	54	Sapotaceae	Pouteria torta							54	
BL	1	55	Sapotaceae	Pouteria torta							55	
BL	1	56	Annonaceae	Xylopia aromatica							56	a 2m da borda
BL	1	57	Myrtaceae	Myrcia guianensis							57	
BL	1	58	Sapotaceae	Pouteria torta							58	dir da 57, arv alta
BL	1	59	Myrtaceae	Campomanesia pubescens							59	lado da 65
BL	1	60	Myrtaceae	Myrcia splendens							60	entre 52 e 67
BL	1	61	Ochnaceae	Ouratea spectabilis							61	placa alta, lado da 62
BL	1	62	Annonaceae	Xylopia aromatica							62	
BL	1	63	Melastomataceae	Miconia albicans							63	
BL	1	64	Ochnaceae	Ouratea spectabilis							64	
BL	1	65	Myrtaceae	Campomanesia pubescens							65	lado da 55, sob ela
BL	1	66	Melastomataceae	Miconia rubiginosa							66	
BL	1	67	Fabaceae	Anadenanthera falcata							67	bem alta, atrás 63
BL	1	68	Fabaceae	Anadenanthera falcata							68	placa no galho seco
BL	1	71	Fabaceae	Senna rugosa							71	Perdida?
BL	1	72	Fabaceae	Dimorphandra mollis							72	depois da 74
BL	1	73	Myrtaceae	Myrcia guianensis							73	
BL	1	74	Ochnaceae	Ouratea spectabilis							74	
BL	1	75	Asteraceae	Piptocarpha rotundifolia							75	
BL	1	76	Nyctaginaceae	Guapira noxia							76	borda
BL	1	77	Myrtaceae	Myrcia guianensis							77	
BL	1	78	Loganiaceae	Strychnos pseudoquina							78	
BL	1	80	Myrtaceae	Myrcia guianensis							80	
BL	1	81	Loganiaceae	Strychnos pseudoquina							81	Perdida?
BL	1	82	Myrtaceae	Myrcia guianensis							82	
BL	1	86	Malpighiaceae	Byrsonima basiloba							86	
BL	1	87	Bignoniaceae	Tabebuia ochracea							87	sob Aegiphyla
BL	1	88	Myrtaceae	Myrcia guianensis							88	

Figura A.1: Exemplo de ficha de fenologia em campo.

# Apêndice B

## Descrição de Dados e-phenology

### B.1 Módulo de Dados Localização

Descrição das tabelas do módulo de localização:

**TABELA: State**

**Descrição:** Esta tabela lista os Estados brasileiros.

*idState*: código sequencial numérico que identifica unicamente os dados de Estado no banco de dados

*state*: nome completo do Estado

*acronym*: sigla do Estado

**TABELA: City**

**Descrição:** Esta tabela lista os Municípios brasileiros.

*idCity*: código sequencial numérico que identifica unicamente os dados de Municípios no banco de dados

*city*: nome completo do município

*idState*: identifica Estado ao qual o município pertence

**TABELA: Area**

**Descrição:** Esta tabela lista os campos onde os estudos de fenologia são realizados.

*idArea*: código sequencial numérico que identifica unicamente os dados de Área no banco de dados

*areaName*: nome da área

*latitude*: define coordenada geográfica de localização da área

*longitude*: define coordenada geográfica de localização da área

*idCity*: identifica o município ao qual a área pertence

*idVegetation*: identifica o tipo da vegetação presente na maior parte da área

**Exemplo:** Para a cidade de Itirapina, a área é Botelho; para a cidade de Jundiá, a área chama-se Serra do Japi.

#### TABELA: **SubArea**

**Descrição:** Esta tabela lista as sub-áreas dentro de uma área de estudos de fenologia.

*idSubArea*: código sequencial numérico que identifica unicamente os dados de Sub-Áreas no banco de dados

*subAreaName*: nome da sub-área

*latitude*: define coordenada geográfica de localização da sub-área

*longitude*: define coordenada geográfica de localização da sub-área

*idArea*: identifica a área com a qual a sub-área está relacionada.

*idSamplingMethod*: identifica tipo de coleta de dados realizada (forma como a área é dividida para a coleta)

*idVegetation*: vegetação predominante na sub-área.

**Exemplo:** Para o Município de Itaripina, a área Botelho tem uma sub-área que é a própria Botelho; no Município de Jundiá, a área Serra do Japi tem duas sub-áreas: Mirant e Ermita.

#### TABELA: **Vegetation**

**Descrição:** Esta tabela lista os tipos de vegetação.

*idVegetation*: código sequencial numérico que identifica unicamente os dados de Vegetação no banco de dados

*vegetationName*: nome da vegetação

*vegetationType*: tipo da vegetação

#### TABELA: **ParamGPS**

**Descrição:** Esta tabela lista os parâmetros informados de GPS utilizado para obter a latitude e longitude de localização.

*idParamGPS*: código sequencial numérico que identifica unicamente os dados de Parâmetros do GPS no banco de dados

*idArea*: área ao qual os parâmetros de GPS faz referência

*idSpatialReference*: faz referência à tabela do SQL Server que lista os objetos de referência para o GPS

*precisionMeter*: precisão do GPS em metros

*precisionKm*: precisão do GPS em quilômetros

#### TABELA: **SamplingMethod**

**Descrição:** Esta tabela lista os métodos de amostragem utilizados para realizar as divisões de localização de um indivíduo.

*idSamplingMethod*: código sequencial numérico que identifica unicamente os dados de Método de Amostragem no banco de dados

*samplingMethodName*: nome da forma como é feita a coleta de dados

**Exemplo:** Aleatória ou Transecto.

#### TABELA: **SamplingEnvironment**

**Descrição:** Esta tabela lista regiões dentro de uma sub-área de acordo com um método de amostragem.

*idSamplingEnvironment*: código sequencial numérico que identifica unicamente os dados de ambientes no banco de dados

*samplingEnvironmentName*: nome para as divisões das regiões onde os dados são coletados

*idSamplingMethod*: identifica o método de amostragem utilizado no ambiente

**Exemplo:** No Município de Itirapina, na área Botelho, a coleta tem como método de amostragem Transecto, que se divide em dois ambientes: Borda e Interior.

#### TABELA: **SamplingAreaEnvironment**

**Descrição:** Esta tabela lista sub-divisões nos ambientes divididos de acordo com o método de amostragem.

*idSamplingAreaEnvironment*: código sequencial numérico que identifica unicamente os dados de ambientes de amostragem no banco de dados

*samplingAreaEnvironmentName*: nome da divisão do ambiente de amostragem

*idSamplingEnvironment*: identifica qual região geral de amostragem se refere

**Exemplo:** Para o Município de Itirapina, a área Botelho, cuja coleta é por Transecto em Borda e Interior, tem subdivisões. Para Borda, as subdivisões são: Borda Leste (BL), Borda Sul (BS); e para Interior: Interior Leste (IL) e Interior Sul (IS).

#### TABELA: **Site**

**Descrição:** Esta tabela lista todos os locais, de acordo com as áreas dos ambientes de amostragem, onde os indivíduos podem ser encontrados.

*idSite*: código sequencial numérico que identifica unicamente os dados de Local no banco de dados

*siteName*: nome para a subdivisão dentro das áreas de ambiente de amostragem, permite a localização de um indivíduo

*latitude*: define coordenada geográfica de localização do local

*longitude*: define coordenada geográfica de localização do local

*idSamplingAreaEnvironment*: identifica em qual região o local se encontra, de acordo com o tipo de amostragem

**Exemplo:** Para o Município de Itirapina, área Botelho, com amostragem por Transecto, dividida em Borda e Interior, e sub-dividida em Borda Leste e Borda Sul, e Interior Leste e Interior Sul, respetivamente, os transectos são: (i) para borda: BL1, BL2, BL3, BL4, BL5, BL6, BL7, BL8, BL9, BL10 e BS1, BS2, BS3, BS4, BS5, BS6, BS7, BS8, BS9, BS10; (ii) para Interior: IL1, IL2, IL3, IL4, IL5, IL6, IL7, IL8 e IS1, IS2, IS3, IS4, IS5, IS6, IS7, IS8.

## B.2 Módulo de Dados Taxonômicos

Descrição das tabelas do módulo de taxonomia:

### TABELA: **Individual**

**Descrição:** Esta tabela lista todos os indivíduos.

*idIndividual*: código sequencial numérico que identifica unicamente os dados de Indivíduos no banco de dados

*individualNumber*: identifica individuo em um transecto, pode se repetir para transectos diferentes.

*comments*: campo para observações sobre indivíduo, pode estar vazio ou não.

*diameter*: grossura do tronco

*height*: altura do indivíduo

*latitude*: define coordenada geográfica de localização do individuo

*longitude*: define coordenada geográfica de localização do individuo

*idSite*: identifica a localização de um individuo

*idCanopyStatus*: identificação do estrato - ambiente onde a planta está

*idHabit*: identificação do habito - tipo de planta

*idSpecie*: identifica a espécie ao qual o indivíduo pertence

*basalArea*: área basal do indivíduo

**TABELA: Method**

**Descrição:** Esta tabela lista todas as maneiras de se fazer a coleta de dados (observar o campo).

*idMethod*: código sequencial numérico que identifica unicamente os dados de Método de observação no banco de dados

*methodName*: nome do método de observação de campo

*beginDate*: data de início que o método começou a ser usado para observação

*endDate*: data de fim que o método parou de ser usado para observação

**Exemplo:** Presença/Ausência; quantificação de fenofases.

**TABELA: MethodIndividual**

**Descrição:** Esta tabela relaciona a tabela Individuo e a tabela Metodo.

*idIndividual*: identificador do indivíduo

*idMethod*: identificador do método

**TABELA: CanopyStatus**

**Descrição:** Esta tabela identifica a divisão do ambiente de acordo com a altura do indivíduo.

*idCanopyStatus*: código sequencial numérico que identifica unicamente os dados de Estrato de indivíduos no banco de dados

*canopyStatusName*: nome do ambiente da planta

*abbreviation*: abreviatura do nome do estrato

**Exemplo:** Dossie/emergentes; sub-bosque e herbáceas.

**TABELA: Habit**

**Descrição:** Esta tabela define o que um indivíduo é.

*idHabit*: código sequencial numérico que identifica unicamente os dados de Habito de indivíduos no banco de dados

*habitName*: nome do habito

*abbreviation*: abreviatura do nome do hábito

**Exemplo:** arbustivo; arbóreo.

**TABELA: Specie**

**Descrição:** Esta tabela lista todas as espécies.

*idSpecie*: código sequencial numérico que identifica unicamente os dados de Espécies no banco de dados

*includeHistoryDate*: parte integrante da identificação de uma espécie, para obtenção de histórico de quando a nomenclatura passou a ter validade

*lastDate*: parte integrante da identificação de uma espécie, para obtenção de histórico de quando a nomenclatura deixou de ter validade

*genus*: identifica o gênero da espécie

*family*: identifica a família da espécie

*specieName*: identifica o epíteto específico da espécie - compõem nome binomial com gênero

*author*: identifica quem descreveu a espécie pode ter mais de um autor. Para casos que a espécie contém mais de um autor. Os nomes devem ser colocados juntos, em sequência - concatenados

*commomName*: nome geral (popular) da espécie. Pode ser multivalorado (haver mais de um nome para a mesma espécie), caso em que os nomes devem ser separados por ponto e vírgula

*sexualityFloral*: identifica se a espécie tem flor masculina, feminina ou hermafrodita. Obs: Monóica: possui flores apenas masculinas ou femininas; Dióica: masculina e femina ou hermafrodita

*imagePath*: identifica localização da imagem para a espécie. Apenas 1 imagem por espécie

*editUser*: responsável pela inclusão do registro

*idDeciduous*: identifica a deciduidade de folhas

*idPhytosociologic*: identifica características específicas da espécie

*idLeaf*: identifica época que há folhas

*Flower*: identifica tipo de flor da espécie

*Fruit*: identifica tipo de fruto da espécie

### B.3 Módulo de Dados Ecológicos

Descrição das tabelas do módulo de dados ecológicos:

**TABELA: Deciduous**

**Descrição:** Esta tabela lista a deciduidade de folhas: se perde, se perde parcialmente ou totalmente folhas ou ainda se não perde.

*idDeciduous*: código sequencial numérico que identifica unicamente os dados de Deciduidade de espécies e indivíduos no banco de dados

*deciduousName*: nome da deciduidade

*abbreviation*: abreviatura do nome

#### TABELA: **Flower**

**Descrição:** Esta tabela lista todas as flores das espécies.

*idFlower*: código sequencial numérico que identifica unicamente os dados de Flor no banco de dados

*flowerName*: nome da flor

*abbreviation*: abreviatura do nome da flor

*color*: cor da flor

*typeFlower*: tipo da flor

*symmetry*: simetria da flor

*diameter*: diâmetro da flor

*deepCorolla*: profundidade da corola

*widthCorolla*: largura da corola

*extentPistil*: comprimento do pistilo

*extentStame*: comprimento do estame

*numberEggCell*: número de óvulos

*hue*: matiz, estado puro da cor

*chrome*: croma, saturação da cor

*brightness*: brilho ou intensidade acromática

*odor*: odor da flor

*volumeNectar*: volume de néctar

*concentrationNectar*: concentração de néctar

#### TABELA: **Fruit**

**Descrição:** Esta tabela lista todos os frutos das espécies e suas características.

*idFruit*: código sequencial numérico que identifica unicamente os dados de Fruta no banco de dados

*fruitName*: nome do fruto

*abbreviation*: abreviatura do nome do fruto

*diameter*: diâmetro do fruto

*extent*: comprimento do fruto

*weight*: peso do fruto

*diameterPip*: diametro da semente (média)

*weightPip*: peso da semente (média)

*extentPip*: comprimento da semente (média)

*numberPip*: número de sementes (média) por fruto  
*colorFruit*: nome da cor do fruto  
*hue*: matiz, estado puro da cor  
*chrome*: croma, saturação da cor  
*brightness*: brilho ou intensidade acromática  
*protein*: define a porcentagem média de proteína no fruto  
*lipid*: define a porcentagem média de lipídio no fruto  
*fructose*: define a quantidade de frutose  
*sucrose*: define a quantidade de sacarose  
*glucose*: define a quantidade de glicose  
*water*: define a quantidade de água  
*calories*: define a quantidade de calorias  
*idAccessoryFruit*: identifica informações de estrutura acessória para o fruto. Fruto pode não ter estrutura acessória.

#### TABELA: **AccessoryFruit**

**Descrição:** Esta tabela lista dados sobre a estruturas acessórias dos frutos.

*idAccessoryFruit*: código sequencial numérico que identifica unicamente os dados de Estrutura Acessória de frutos no banco de dados

*weight*: define peso da estrutura acessória

*numberFruits*: número de frutos na estrutura acessória

*hue*: matiz, estado puro da cor

*chrome*: croma, saturação da cor

*brightness*: brilho ou intensidade acromática

*contrastLeaf*: número advindo de cálculo para verificar o quanto o fruto se destaca, refere-se ao contraste cromático entre a cor dos frutos e seu segundo plano

*contrastSecondAccessory*: referem-se ao contraste cromático entre a cor dos frutos e seu segundo plano (geralmente folhas) ou estruturas secundárias. Sendo as estruturas secundárias estruturas dos frutos associadas ao display dos mesmos, como frutos imaturos e pecíolos, cápsulas ou arilos coloridos, etc

#### TABELA: **Leaf**

**Descrição:** Esta tabela lista características da folha de espécies.

*idLeaf*: código sequencial numérico que identifica unicamente os dados de Folha no banco de dados

*typeLeaf*: tipo de folha

*hue*: matiz, estado puro da cor

*chrome*: croma, saturação da cor

*brightness*: brilho ou intensidade acromática

#### TABELA: **Phytosociologic**

**Descrição:** Esta tabela lista dados estatísticos de uma espécie.

*idPhytosociologic*: código sequencial numérico que identifica unicamente os dados de Fitossociologia no banco de dados

*totalIndividuals*: quantidade de indivíduos

*absoluteDensity*: Densidade Absoluta (DeAb) relação do número total de indivíduos de uma espécie por área; obtida pela divisão do número total de indivíduos desta espécie encontrados na área amostral, por unidade de área (1 ha) - Valor: número por 1ha

*relativeDensity*: Densidade Relativa (DeRel) representa a porcentagem com que uma espécie aparece na amostragem em relação ao total de indivíduos do componente amostrado. A probabilidade de, se amostrado um indivíduo aleatoriamente, ele pertença à espécie em questão - Valor: número em %

*absoluteFrequency*: Frequência Absoluta (FrAb) a porcentagem de amostras em que foi registrado uma determinada espécie, ou a probabilidade de uma parcela aleatoriamente sorteada conter a dada espécie. Expressa pela porcentagem do número de unidades amostrais em que a espécie ocorre dividido pelo número total de unidades amostrais - Valor: número em %

*relativeFrequency*: Frequência Relativa (FrRel) relação em porcentagem da ocorrência de uma espécie pela somatória de ocorrências para todas as espécies do componente analisado - Valor: número em %

*basalArea*: é o valor da área de secção do tronco a 1,3 m de altura, serve de indicativo para a dominância das espécies - Valor número em m<sup>2</sup>

*absoluteDominance*: Dominância Absoluta (DoAb) é a área basal total em m<sup>2</sup> que uma determinada espécie ocupa na amostra, por unidade de área (1 ha), calculada pela somatória da área de todos os indivíduos desta espécie - Valor: número em m<sup>2</sup> por ha

*relativeDominance*: Dominância Relativa (DoRel) é a área total da secção do caule que todos os indivíduos de uma determinada espécie ocupam, dividido pelo total de todos os indivíduos amostrados e expressa em porcentagem. Representa a contribuição da biomassa do táxon em relação ao total da biomassa do componente analisado. - Valor: número em %

*ivi*: Índice de Valor de Importância (IVI) - Estima a importância ecológica de um táxon (espécie, família, etc) dentro de uma comunidade florestal; é calculado pela soma da densidade relativa, da dominancia relativa e ainda da frequência relativa, seu valor máximo, portanto, corresponde a 300%

*ivc*: Índice de Valor de Cobertura (IVC) - Estima a importância ecológica de um táxon (espécies, família, etc) dentro de uma comunidade florestal; é calculado pela soma da densidade relativa com a dominância relativa de determinada espécie, sendo seu valor máximo possível igual a 200%

#### TABELA: **PollinationSyndrome**

**Descrição:** Esta tabela lista todos os tipos de síndromes de polinização.

*idPollinationSyndrome*: código sequencial numérico que identifica unicamente os dados de Síndrome de Polinização no banco de dados

*pollinationSyndromeName*: nome para síndrome de polinização

*abbreviation*: abreviatura do nome da síndrome de polinização

#### TABELA: **DispersalSyndrome**

**Descrição:** Esta tabela lista todos os tipos de síndrome de dispersão de sementes.

*idDispersalSyndrome*: código sequencial numérico que identifica unicamente os dados de Síndrome de Dispersão no banco de dados

*dispersalSyndromeName*: nome para síndrome de dispersão

*abbreviation*: abreviatura do nome da síndrome de dispersão

#### TABELA: **PollinationSyndromeSpecie**

**Descrição:** Esta tabela relaciona tabela Especie e a tabela de Síndrome de Polinização.

*idPollinationSyndromeSpecie*: código sequencial numérico que identifica unicamente os dados de síndrome de polinização de espécies no banco de dados

*idSpecie*: identificador de espécie

*idPollinationSyndrome*: identificador da síndrome de polinização

*majorPollinationSyndrome*: define se é a polinização principal

**Exemplo:** 'S' ou 'N'.

#### TABELA: **DispersalSyndromeSpecie**

**Descrição:** Esta tabela relaciona tabela Especie e a tabela Dispersao.

*idDispersalSyndromeSpecie*: código sequencial numérico que identifica unicamente os dados de síndrome de dispersão de espécies no banco de dados

*idSpecie*: identificador de espécie

*idDispersalSyndrome*: identificador da síndrome de dispersão

*majorDispersalSyndrome*: define se é a dispersão é principal

**Exemplo:** 'S' ou 'N'.

#### TABELA: **Pollinator**

**Descrição:** Esta tabela relaciona os visitantes polinizadores.

*idPollinator*: código sequencial numérico que identifica unicamente os dados de Polinizadores no banco de dados

*idIndividual*: identificador de indivíduo

*idSpecie*: identificador de Espécie

*namePollinator*: nome do visitante polinizador

*watchDate*: data que a visita foi observada

#### TABELA: **Disperser**

**Descrição:** Esta tabela relaciona os visitantes dispersores.

*idDisperser*: código sequencial numérico que identifica unicamente os dados de Dispersores no banco de dados

*idIndividual*: identificador de indivíduo

*idSpecie*: identificador de Espécie

*nameDisperser*: nome do visitante dispersor

*watchDate*: data que a visita foi observada

## B.4 Módulo de Dados Climáticos

Descrição das tabelas do módulo de dados de clima:

#### TABELA: **Tower**

**Descrição:** Esta tabela lista as torres instaladas.

*idTower*: código sequencial numérico que identifica unicamente os dados de Torre no banco de dados

*idCity*: identificador da cidade onde a torre está localizada

*towerName*: nome da torre

*latitude*: define coordenada geográfica de localização da torre

*longitude*: define coordenada geográfica de localização da torre

*height*: altura da torre

*installationDate*: data de instalação da torre

**TABELA: SensorType**

**Descrição:** Esta tabela lista os possíveis tipos de sensores.

*idSensorType*: código sequencial numérico que identifica unicamente os dados de Tipo de Sensor no banco de dados

*sensorType*: define tipo de sensor

**TABELA: Sensor**

**Descrição:** Esta tabela lista todos os sensores.

*idSensor*: código sequencial numérico que identifica unicamente os dados de Sensor no banco de dados

*idTower*: identificador da torre

*idSensorType*: identificador do tipo de sensor

*sensorName*: nome do sensor

*model*: modelo do sensor

**TABELA: ClimateData**

**Descrição:** Esta tabela armazena valores coletados a partir dos sensores localizados na Torre - dados HOBOLINK. Será atualizada com arquivo CSV enviado via FTP

*idClimateData*: código sequencial numérico que identifica unicamente os dados de Clima dos sensores no banco de dados

*idSensor*: identificador do sensor

*value*: valor lido pelo sensor

*realDateHour*: data e hora de leitura do valor pelo sensor

## B.5 Módulo de Dados Fenológicos

Descrição das tabelas do módulo de fenologia:

**TABELA: Phenophase**

**Descrição:** Esta tabela lista todas as fenofases.

*idPhenophase*: código sequencial numérico que identifica unicamente os dados de Fenofase no banco de dados

*phenophaseName*: nome da fenofase

*comments*: campo para observações

**TABELA: DataPhenology**

**Descrição:** Esta tabela armazena os dados das observações de fenologia realizadas em campo.

*idDataPhenology*: código sequencial numérico que identifica unicamente os dados de Dados de Fenologia no banco de dados

*watchDate*: data de observação dos dados da fenofase

*observer*: responsável pela observação em campo

*idPhenophase*: identificador da fenofase

*idIndividual*: identificador do individuo

*value*: valorem de 0 a 2 que identificam a intensidade com que a fenofase ocorreu.

**Exemplo:** 0 representa que não há indícios da fenofase para aquele indivíduo; 1 representa que a fenofase está presente em menor intensidade (até 50%) e 2 representa que a fenofase está presente em maior intensidade (de 51% a 100%).

**TABELA: DataCycle**

**Descrição:** Esta tabela armazena valores de análises dos ciclos anuais da fenofase para um indivíduo.

*idDataCycle*: código sequencial numérico que identifica unicamente os dados estatísticos de Ciclo no banco de dados

*begin*: identifica o mês de início de um ciclo da fenofase. Uma fenofase pode ter mais de um início, no período de um ano, para o mesmo individuo

*durationYear*: identifica quantos meses no ano o ciclo foi observado

*durationCycle*: identifica quantos meses o ciclo durou

*peakYear*: identifica o mês onde a fenofase teve maior intensidade dentro do ano

*peakCycle*: identifica o mês onde a fenofase teve maior intensidade dentro de um ciclo

## B.6 Módulo de Imagens e Metadados

Descrição das tabelas do módulo de imagens:

**TABELA: Camera**

**Descrição:** Esta tabela armazena dados sobre câmeras digitais instaladas em torres.

*idCamera*: código sequencial numérico que identifica unicamente os dados de Câmera no banco de dados

*idTower*: identificador da torre

*maker*: fabricante da câmera

#### TABELA: **Image**

**Descrição:** Esta tabela armazena dados sobre as imagens tiradas pelas câmeras digitais.

*idImage*: código sequencial numérico que identifica unicamente os dados de Imagem no banco de dados

*idCamera*: identificador da câmera digital

*path*: caminho da imagem no sistema de arquivos

*name*: nome para a imagem

*dateHour*: data e hora que a imagem foi tirada

#### TABELA: **Descriptor**

**Descrição:** Esta tabela armazena quais descritores serão usados para realizar a recuperação de imagens por conteúdo.

*idDescriptor*: código sequencial numérico que identifica unicamente os dados de Descritor de imagem no banco de dados

*descriptorName*: nome do descritor

*authors*: autor ou autores do descritor

*type\_2*: tipo do descritor (cor, forma, etc.)

#### TABELA: **DescriptorImage**

**Descrição:** Esta tabela armazena os dados que relacionam imagens a seus descritores.

*idDescriptorImage*: código sequencial numérico que identifica unicamente os dados de Descritor associado às imagens no banco de dados

*idDescriptor*: identificador do descritor

*idImagem*: identificador da imagem

#### TABELA: **FeatureVector**

**Descrição:** Esta tabela armazena os dados referentes aos vetores de características extraídos das imagens com os descritores.

*idFeatureVector*: código sequencial numérico que identifica unicamente os dados de Vector de Característica no banco de dados

*idDescriptor*: identificador do descritor

*idImage*: identificador da imagem

*vector*: vetor de características da imagem de acordo com o descritor

*idImageMaskSpecie*: máscara utilizada para obter o vetor de características

#### TABELA: **ImageMaskSpecie**

**Descrição:** Esta tabela armazena as imagens de máscaras das espécies.

*idImageMaskSpecie*: código sequencial numérico que identifica unicamente os dados das Imagens de Máscara das espécies no banco de dados

*idSpecie*: identificador da espécie a qual a máscara pertence

*path*: caminho da imagem da máscara no sistema de arquivos

*name*: nome para a imagem da máscara

#### TABELA: **Matrix**

**Descrição:** Esta tabela armazena os valores de distância referentes à matriz de distância de uma imagem obtida por descritores.

*idMatrix*: identificador sequencial numérico

*idDescriptor*: identificador do descritor

*idImage*: identificador da imagem

*distanceValue*: valor obtido pelo cálculo de similaridade (função de distância) da imagem com relação a todas as demais

#### **SENSORES TORRE ITIRAPINA**

*WATER\_CONT*: umidade do solo a vinte centímetros

*WETNESS*: umidade foliar - fica em uma árvore e capta a umidade da área foliar

*PAR - RA PAR*: para cima - recebe a irradiação

*PAR - RA PAR REFLETIDA*: reflete a irradiação para baixo (o que volta das plantas)

*RAIN*: pluviômetro

*SOLAR RAD - RA SOLAR REFLETIDA*: para baixo - radiação solar

*SOLAR RAD - RA SOLAR*: para cima

*WIND SPEED*: Velocidade do vento

*GUST SPEED*: Rajada de vento

*WIND DIR*: Direção de vento

*TEMP*: Temperatura do ar

*RH*: Umidade Relativa do Ar

*DEWPT*: Ponto de Orvalho

*PRESSURE*: Pressão Atmosférica

*TEMP*: Temperatura do solo a vinte centímetros

*BATT*: Bateria



## Apêndice C

### Planilha com Dados Climáticos

Plot Title: Cerrado

#	Date Time, GMT-03:00	Water Content, m <sup>3</sup> /m <sup>3</sup> (LBL: Umd. Solo 20 cm)	Wetness, % (LBL: Umd. Foliar)	PAR, uE (LBL: Ra. Par)	PAR, uE (LBL: Ra. Par Refletida)	Rain, mm (LBL: Chuva / mm)	Solar Radiation, W/m <sup>2</sup> (LBL: Ra. Solar Refletida)	Solar Radiation, W/m <sup>2</sup> (LBL: Ra. Solar)
1	09/12/2011 10:50	0,02876	29,41	481,2	26,2		39,4	186,9
2	09/12/2011 10:55	0,02819	28,24	663,7	33,7	0	50,6	251,9
3	09/12/2011 11:00	0,02819	27,06	666,2	33,7	0	51,9	251,9
4	09/12/2011 11:05	0,02762	25,29	691,2	33,7	0	51,9	260,6
5	09/12/2011 11:10	0,02819	24,12	556,2	28,7	0	41,9	210,6
6	09/12/2011 11:15	0,02762	22,35	503,7	26,2	0	38,1	189,4
7	09/12/2011 11:20	0,02762	21,76	351,2	18,7	0	28,1	133,1
8	09/12/2011 11:25	0,02704	21,76	303,7	16,2	0	23,1	114,4
9	09/12/2011 11:30	0,02704	21,18	318,7	16,2	0	25,6	121,9
10	09/12/2011 11:35	0,02647	21,18	388,7	21,2	0	29,4	145,6
11	09/12/2011 11:40	0,02647	20,59	396,2	21,2	0	29,4	148,1
12	09/12/2011 11:45	0,02647	20	411,2	21,2	0	30,6	154,4
13	09/12/2011 11:50	0,02647	20	516,2	26,2	0	38,1	193,1
14	09/12/2011 11:55	0,0259	19,41	491,2	23,7	0	36,9	184,4
15	09/12/2011 12:00	0,0259	19,41	386,2	21,2	0	30,6	146,9
16	09/12/2011 12:05	0,0259	20,59	533,7	26,2	0,2	38,1	194,4
17	09/12/2011 12:10	0,02533	20,59	688,7	33,7	0	49,4	246,9
18	09/12/2011 12:15	0,02533	21,76	691,2	33,7	0,4	49,4	251,9
19	09/12/2011 12:20	0,02533	22,35	668,7	33,7	0	49,4	243,1
20	09/12/2011 12:25	0,02476	22,94	678,7	33,7	0,2	49,4	248,1
21	09/12/2011 12:30	0,02476	22,94	521,2	28,7	0	40,6	199,4
22	09/12/2011 12:35	0,02476	23,53	471,2	23,7	0	36,9	181,9
23	09/12/2011 12:40	0,02419	22,35	603,7	28,7	0	43,1	218,1
24	09/12/2011 12:45	0,02419	22,35	333,7	16,2	0	24,4	121,9
25	09/12/2011 12:50	0,02419	22,35	416,2	21,2	0,2	28,1	143,1
26	09/12/2011 12:55	0,02419	22,94	466,2	23,7	0,2	30,6	158,1
27	09/12/2011 13:00	0,02362	24,71	456,2	23,7	0,6	30,6	156,9
28	09/12/2011 13:05	0,02362	26,47	403,7	21,2	0,6	25,6	135,6
29	09/12/2011 13:10	0,02362	30,59	306,2	16,2	1,2	18,1	99,4

Figura C.1: Exemplo de planilha com dados climáticos coletados pelo sensores da Torre de Itrapina - parte 1. Campos: Data e hora e Sensores com valores coletados.

Figura C.2: Exemplo de planilha com dados climáticos coletados pelo sensores da Torre de Itirapina - parte 2. Campos: Data e hora e Sensores com valores coletados.

Continuação Leitura por Sensores

#	Date Time, GMT-03:00	Wind Speed, m/s (LBL: Velo. do Vento)	Gust Speed, m/s (LBL: Rajada de Vento)	Wind Direction, ø (LBL: Direcao do Vento)	Temp, °C (LBL: Tem. Ar)	RH, % (LBL: Umd. Ar)	Pressure, mbar (LBL: Pressao Atmosferica)	Temp, °C (LBL: Tem. Solo 20 cm)
1	09/12/2011 10:50	2,04	3,34	25,3	20,96	96,2	921,05	20,698
2	09/12/2011 10:55	2,04	3,71	22,5	20,96	94,8	920,95	20,722
3	09/12/2011 11:00	2,6	4,27	14	21,127	95	920,95	20,746
4	09/12/2011 11:05	2,78	4,45	16,8	21,246	94,2	920,95	20,746
5	09/12/2011 11:10	2,97	4,64	2,8	21,27	93,6	921,05	20,77
6	09/12/2011 11:15	3,71	5,57	355,2	21,103	93,5	921,05	20,77
7	09/12/2011 11:20	2,6	4,08	344	21,127	94,2	921,05	20,793
8	09/12/2011 11:25	2,41	3,9	345,4	21,056	94,6	921,05	20,793
9	09/12/2011 11:30	3,15	4,82	344	20,793	95,8	920,95	20,817
10	09/12/2011 11:35	2,97	4,64	352,4	20,77	94,6	920,95	20,817
11	09/12/2011 11:40	3,34	5,01	352,4	20,77	95,2	920,85	20,841
12	09/12/2011 11:45	3,34	5,2	2,8	20,627	95	920,75	20,841
13	09/12/2011 11:50	3,15	4,45	5,6	20,555	95,5	920,75	20,865
14	09/12/2011 11:55	2,78	4,64	9,8	20,579	96,3	920,75	20,865
15	09/12/2011 12:00	2,97	4,27	5,6	20,412	96,7	920,65	20,889
16	09/12/2011 12:05	2,41	3,15	4,2	20,436	96,6	920,65	20,889
17	09/12/2011 12:10	2,23	3,53	4,2	20,579	97	920,55	20,913
18	09/12/2011 12:15	2,04	3,15	353,8	20,698	97,4	920,55	20,913
19	09/12/2011 12:20	1,48	2,41	355,2	20,913	97,8	920,55	20,936
20	09/12/2011 12:25	1,67	2,97	5,6	20,96	97,2	920,55	20,936
21	09/12/2011 12:30	1,48	2,23	5,6	21,127	97,9	920,45	20,96
22	09/12/2011 12:35	2,04	3,15	11,2	21,056	97,9	920,35	20,96
23	09/12/2011 12:40	2,41	4,45	8,4	21,032	96,3	920,25	20,984
24	09/12/2011 12:45	3,53	6,12	7	20,793	96,1	920,35	20,984
25	09/12/2011 12:50	3,53	4,82	14	20,65	97	920,25	21,008
26	09/12/2011 12:55	2,6	4,27	2,8	20,627	96,1	920,25	21,008
27	09/12/2011 13:00	3,34	4,82	355,2	20,46	96,4	920,15	21,032
28	09/12/2011 13:05	2,78	3,9	1,4	20,388	97,5	920,15	21,032
29	09/12/2011 13:10	1,86	2,78	345,4	20,531	98,3	920,05	21,032