

Relatório Científico - Bolsa de Treinamento Técnico (TT- 4) em "Criação de Chaves Interativas Online para Identificação de Espécies de Bignonieae (Bignoniaceae) "

associada ao auxílio regular do projeto temático Fapesp/Biota "Estruturação e Evolução da Biota Amazônica e seu Ambiente: uma
abordagem integrativa"

(Processo 2012/50260-6)

Bolsista: Raquel Machado de Sousa

Processo: 2016/11626-6

Supervisora: Profa. Dra. Lúcia Garcez Lohmann

Orientador: Profo. Dro. Antonio M. Saraiva

Período do relatório: 01/07/2016 - 31/12/2016

Vigência da Bolsa: 01/07/2016 - 31/08/2017

Introdução

Chaves interativas são fundamentais para a identificação dos espécimes por não-especialistas. Essas chaves são baseadas em um conjunto de critérios de diferenciação que permitem combinar materiais de diferentes espécies de seus nomes científicos precisos. Alguns pacotes de software já estão disponíveis para a identificação dos táxons com chaves interativas. No entanto, a maioria destes softwares não estão livremente disponíveis, exigindo dos desenvolvedores e usuários a compra de licenças cujo custo tornam seu uso proibitivo. Outros pacotes de software são executados em servidores remotos, o que pode causar problemas de desempenho. O objetivo deste projeto é o desenvolvimento de um sistema online de chaves interativas para seleção de táxon como modelo de sistema para identificação de espécies, e o uso livre das chaves interativas pelos usuários para identificação do táxon de interesse. O sistema aqui será aplicado à identificação de plantas da família Bignoniaceae.

As atividades desenvolvidas até o momento incluem: (i) criar uma interface gráfica web interativa e eficiente para os usuários finais para lidar com as chaves interativas; (ii) criar interfaces, protocolos e padrões para especialistas entrarem com os dados de taxon; (iii) desenvolver ferramentas para a identificação de problemas de integridade, consistência, singularidade e precisão em bancos de dados.

Atividades Realizadas

As atividades realizadas até o momento consistem em:

1. Criação de uma interface gráfica web interativa e eficiente para os usuários:

O sistema possui a seguinte interface apresentada nas figuras nº 1, 2, 3, o sistema está disponível em português e em inglês. A página principal apresenta os descritores disponíveis para a identificação das espécies com seus respectivos estados. Ao selecionar os descritores para identificação, as espécies aparecem no campo ao lado de acordo com as características selecionadas.

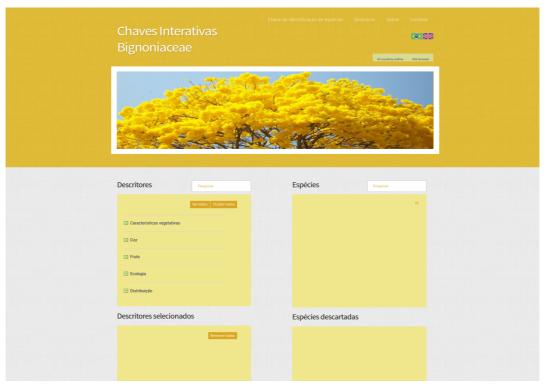


Figura 1: Página inicial do sistema em português

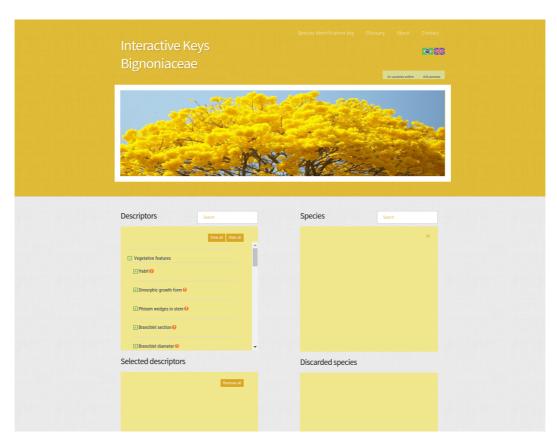


Figura 2: Página inicial do sistema em inglês

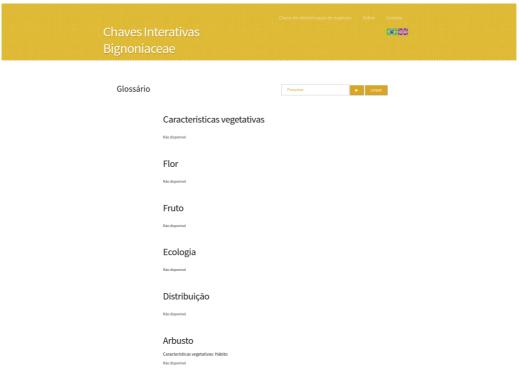


Figura 3: Página do glossário em português

Quando os descritores são selecionados, esses descritores irão aparecer na caixa de descritores selecionados e as espécies que não correspondem a esses descritores irão aparecer na caixa de espécies descartadas.

Já na página de glossário aparece a descrição de todas as características morfológicas dos descritores utilizados para identificação de espécies. Isso é importante principalmente para usuários leigos que não sabem o que significa determinada característica da planta. Atualmente estamos trabalhando na inclusão dessas informações no sistema e finalização das páginas de espécies e espécimes do sistema.

2. <u>Criação interfaces, protocolos e padrões para especialistas entrarem com os dados de taxon:</u>

O primeiro passo para o desenvolvimento das chaves interativas é saber quais informações estariam disponíveis na ferramenta para identificação das espécies de bignoniaceae. Para isso foi preciso definir uma tabela formatada de glossário, de espécies e de espécimes que são preenchidas de forma que o especialista possa inserir as informações.

Ficha de Glossário:

Na ficha de glossário é definido todas as informações de campos do sistema e os descritores que ajudam na identificação das espécies. O glossário é importante pois é onde é definido que informações de descritores estão disponíveis no sistema e o que significa cada descritor e seus estados. Na definição do glossário foi usado o padrão Darwin Core (DwC) [1], que é um corpo de padrões, que incluem um glossário de termos (que em outros contextos podem ser chamados de propriedades, elementos, campos, colunas, atributos ou conceitos) com a intenção de facilitar o compartilhamento de informação sobre diversidade biológica providenciando definições de referência, exemplos e comentários. O DwC é primeiramente baseado em taxon, suas ocorrências na natureza como documentado por observações, espécimes, amostras e informações relacionadas. Além do padrão Dwc, foi utilizado as características morfológicas e caracteres de estados que são baseados em observações originais de especialistas da área.

A tabela de glossário pode ser observada na figura nº 4.

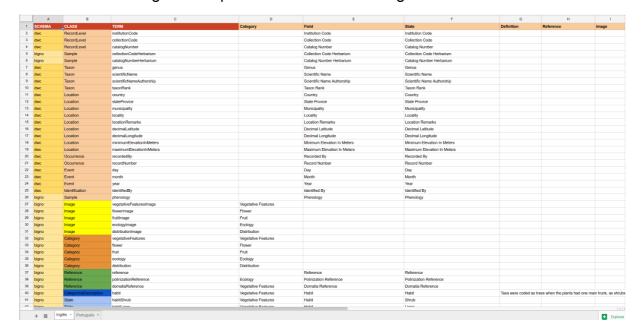


Figura 4: Tabela de glossário

O sistema lê as seguintes informações definidas no glossário:

 SCHEMA: A coluna schema indica os campos que podem ser referenciados no padrão DwC para disponibilização informações em biodiversidade. O que é assinado como *dwc* é a informação mapeada para Darwin Core, e o que é *bigno* são as informações que são relacionados ao padrão criado para as informações que não foram mapeadas para o DwC e estão no sistema.

 CLASS: A coluna class se referência tanto à termos de classes do padrão dwc e termos de classes criadas para o padrão bigno, na tabela nº 1 demonstra a classe de informações para cada padrão.

| Schema | dwc | bigno |
|--------|----------------|------------------------|
| | RecordLevel | Sample |
| | Taxon | Image |
| | Location | Category |
| | Occurrence | Reference |
| Class | Event | CategoricalDescription |
| | Identification | State |
| | | SiteLabel |
| | | ProfilesLabel |
| | | KeyLabel |

Tabela 1: Classes de informação dos padrões do sistema

- TERM: A coluna term se referência a todos os campos possíveis do sistema, existe um termo para cada campo e seus estados, cada termo faz parte de uma classe.
- CATEGORY: Essa coluna determina que tipo de categoria de informação os termos definidos de descritores fazem parte. Aqui foram definidas 5 categorias para os termos que descrevem os descritores morfológicos que identificam uma determinada espécie. Que são: Características Vegetativas (Vegetative Features), Flor (Flower), Fruto (Fruit), Ecologia (Ecology) e Distribuição (Distribution). Esses descritores são representados pela classe Categorical Description e State.

- FIELD: Essa coluna se referência ao nome do campo de cada termo que aparece no sistema para o usuário.
- STATE: Essa coluna recebe o valores de estados de cada campo field.
- DEFINITION: Essa coluna contém a descrição das definições (significado) de cada valor de field e state.
- REFERENCE: Essa coluna contém a referência das definições.
- IMAGE: Essa coluna contém o caminho das imagens relacionadas a cada valor de field se houver imagem.

- Descritores morfológicos:

Na fase de definição de glossário, de 175 descritores morfológicos utilizados pela especialista, Dr^a. Lúcia Lohmann, para identificação das espécies de Bignoniaceae, foram selecionados 79 e acrescentados mais 6 relacionados a características ecológicas e de distribuição. Todos os descritores podem ser observados no apêndice.

Ficha de Espécies:

A ficha de espécies possui, todas as informações dos descritores morfológicos mais as informações de: *genus, scientific name, scientific name authorship, taxon rank, phenology.*

- Ficha de Espécimes:

A ficha de espécimes possui, todas as informações dos descritores morfológicos mais as informações de: institution code, collection code, catalog number, collection code herbarium, catalog number herbarium, genus, scientific name, scientific name authorship, taxon rank, country, state province, municipality, locality, location remarks, decimal latitude, decimal longitude, minimum elevation in meters, maximum elevation in meters, recorded by, record number, day, month, year, identified by, phenology.

Todas as informações das planilhas são lidas pelo sistema para a formação das funções das chaves de identificação das espécies.

3. <u>Desenvolvimento</u> <u>de</u> <u>ferramentas</u> <u>para</u> <u>a</u> <u>identificação</u> <u>de</u> <u>problemas</u> <u>de</u> <u>integridade</u>, <u>consistência</u>, <u>singularidade</u> <u>e</u> <u>precisão</u> <u>em</u> <u>banco</u> <u>de dados</u>.

Para o desenvolvimento do sistema é utilizado o framework LoopBack [2], que é um conjunto de módulos em node.js que podem usados de forma independente ou em conjunto para construir rapidamente APIs REST. Uma aplicação LoopBack interage com fonte de dados por meio da API de modelo LoopBack, disponível localmente no Node.js e via cliente para HTML5, por exemplo. Usando essas APIs, é possível a aplicação consultar banco de dados, armazenar dados, upload de arquivos, enviar e-mails, criar notificações push, registrar usuários e executar outras ações fornecidas pelos serviços e fonte de dados.

Para armazenamento dos dados está sendo utilizado o banco de dados MongoDB [3]. MongoDB é um banco de dados *open source* orientado a documentos que provê alta desempenho, alta disponibilidade e dimensionamento automático. Ele evita a necessidade de um mapeamento objeto-relacional para facilitar o desenvolvimento de aplicações.

Basicamente o sistema é desenvolvido em cima de modelos que são o coração do LoopBack e estes representam as fonte de dados do back-end, como banco de dados. Esses modelos são objetos Javascript com APIs de Node e REST APIs.

- Implementação dos Modelos:

O sistema de chaves interativas possui os modelos apresentados na figura de nº 5 em seu back-end.

| bdd | |
|---|---|
| Collection | Show/Hide List Operations Expand Operations |
| Dataset | Show/Hide List Operations Expand Operations |
| Identification | Show/Hide List Operations Expand Operations |
| Project | Show/Hide List Operations Expand Operations |
| Schema | Show/Hide List Operations Expand Operations |
| Species | Show/Hide List Operations Expand Operations |
| Specimen | Show/Hide List Operations Expand Operations |
| storage | Show/Hide List Operations Expand Operations |
| User | Show/Hide List Operations Expand Operations |
| [BASE URL: /api , API VERSION: 1.0.0] | |

Figura 5: Modelos implementados no back-end.

Nesses modelos é que estão implementados as principais funções utilizadas nas chaves interativas. Por exemplo, no modelo schema é implementado a função de leitura da planilha do glossário, no modelo species é implementado a função de leitura da planilha espécie e no modelo specimen é implementado a função de leitura da planilha de espécimes. Em resumo, as atividades de desenvolvimento do sistema já executadas foram:

- Definição das fichas de glossário, espécies e espécimes.
- Definição e implementação do layout do sistema.
- Implementação da página principal e página de glossário no front-end.
- Implementação da função de leitura das fichas de espécies, espécime e schema para o padrão definido nas planilhas.
- Integração entre as funções implementadas do back-end e front-end.
- Configuração dos parâmetros de redirecionamento das páginas no server.
- Configuração do google analytics para acompanhamento de acesso ao sistema.

As atividades em desenvolvimento são:

- Implementação da página de espécies e espécimes no front-end do sistema.
- Função de agregação dos dados de espécies e espécimes no back-end.

• Função para leitura de metadados de imagens no back-end.

Atividades Complementares

Além das atividades do projeto, durante este período foram realizadas outras atividades de âmbito acadêmico que contribuíram para formação da bolsista.

Disciplinas:

"Introdução à Ciência de Dados". Escola Politécnica - USP Carga horária:
 120 horas. (Setembro - Dezembro/2016)

Palestras e Oficinas:

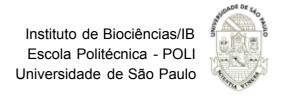
- "Plataforma Analítica para Gestão de Grandes Volumes de Dados" ministrada na Escola Politécnica da Universidade de São Paulo.
- "Nuvem e Transformação Digital Na Busca por Competitividade e Diferenciação" ministrada na Escola Politécnica da Universidade de São Paulo.
- "Oficina Aspectos de Desenvolvimento do Portal de Dados da Biodiversidade Brasileira" realizada na Escola Politécnica da Universidade de São Paulo.

Workshops:

- "V Workshop de Pós-Graduação em Engenharia de Computação", com duração de 08 (oito) horas, realizado na Escola Politécnica da Universidade de São Paulo.
- Monitora na Semana USP de Ciência e Tecnologia 2016 realizada na Universidade de São Paulo.

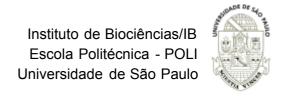
Defesas de Mestrado e Doutorado:

1. [Doutorado] Allan Koch Veiga. "A Conceptual Framework on Biodiversity Data Quality".



Referências

- [1] Darwin Core Terms: A quick reference guide. Disponível em: http://rs.tdwg.org/dwc/terms/.
- [2] LoopBack 2.x. Disponível em: http://loopback.io/doc/en/lb2/index.html.
- [3] Introduction to MongoDB. Disponível em: https://docs.mongodb.com/getting-started/shell/introduction/



APÊNDICE

1 - Lista de 175 descritores morfológicos iniciais:

| Habit | Adaxial side of the leaflets - Pellucid punctations | Staminode - height of insertion of the staminode |
|--|--|--|
| Size | Adaxial side of the leaflets - Indument | Staminode - Anther position in relation to corolla lobes |
| Dimorphic growth form | Adaxial side of the leaflets - Hair type | Anther - indumentum |
| Phloem wedges in stem | Adaxial side of the leaflets - Indumentum position | Anther - thecae |
| Width of phloem wedges | Inflorescence - position | Anther - length |
| Pith | Inflorescence - type | Anther - color |
| Sclerotic pith | Inflorescence - shape | Anther - Anther connective texture |
| Septate fibers | Inflorescence - Flower density | Anther - Shape of connective |
| Marginal parenchyma | Peduncle - length | Pollen - type |
| Vessel dimorphism | Rachis - length | Pollen - apertures |
| Crystals in vessels | Rachis - indument | Pollen - exine ornamentation |
| Crystals in wood ray cells | Rachis - Hair type | Pollen - Nectar disk |
| Width of wood ray | Bracts - persistence | Pollen - Nectar disk color |
| Perforated ray cells | Bracts - shape | Pollen - Nectar disk height |
| Notable odors | Bracts - length | Pollen - Shipe at the base of ovary |
| Branchlet - section | Bracts - Glands | Pollen - Pistil lenght |
| Branchlet - diameter | Flower bracts - size | Ovary - shape |
| Branchlet - striation | Flower bracts - shape | Ovary - length |
| Branchlet - indumentum | Flower bracts - length | Ovary - width |
| Branchlet - Hair type | Flower bracts - Glands | Ovary - surface |
| Branchlet - lenticels | Pedicel - length | Ovary - indumentum |
| Branchlet - Interpetiolar gland fields | Calyx - shape | Ovary - Number of ovule series in each placenta |
| Branchlet - Interpetiolar ridge | Calyx - apex | Style - length |
| | 1 | • |

Instituto de Biociências/IB Escola Politécnica - POLI Universidade de São Paulo



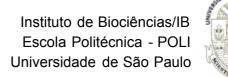
| Branchlet - Prophylls of axillary buds | Calyx - texture | Style - indumentum |
|--|----------------------------------|----------------------------------|
| Branchlet - Glands on prophylls | Calyx - color | Stigma lobes - shape |
| Branchlet - Prophylls indumentum | Calyx - outer indumentum | Stigma lobes - indumentum |
| Branchlet - Hair type | Calyx - Hair type | Fruit - splitting |
| Branchlet - Petiole and petiolule bases | Calyx - glands | Fruit - lenght |
| Branchlet - Petiole lenght | Calyx - length | Fruit - width |
| Branchlet - Petiolule lenght | Calyx - width | Fruit - shape |
| Branchlet - Similarity of the petiolules | Calyx - Double-calyx | Fruit - Shape of the extremities |
| Branchlet - Petiole and petiolule indumentum | Calyx - Calyx arrangement in bud | Fruit - thickness |
| Branchlet - Indumentum position | Corolla - tube color | Fruit - texture |
| Branchlet - Hair type | Corolla - lobes color | Fruit - curvature |
| Branchlet - Leaf type | Corolla - mouth color | Fruit - valve indumentum |
| Branchlet - Tendril type | Corolla - nectar guides | Fruit - valve lenticels |
| Branchlet - Hook at tendril tip | Corolla - nectar guides color | Fruit - valve glands |
| Branchlet - Adhesive disks at tendril tip | Corolla - shape | Fruit - ridges |
| Lateral leaflets - shape | Corolla - whole length | Fruit - valve margin |
| Lateral leaflets - lenght | Corolla - width | Fruit - valve wings |
| Lateral leaflets - width | Corolla - curvature | Fruit - Fruit ornamentation |
| Lateral leaflets - texture | Corolla - texture | Fruit - Calyx in fruit |
| Lateral leaflets - margin | Corolla - tube | Seed - lenght |
| Lateral leaflets - apex shape | Corolla - length | Seed - width |
| Lateral leaflets - base shape | Corolla - lobes | Seed - wing shape |
| Apical leaflets - shape | Corolla - length | Seed - texture |
| Apical leaflets - length | Corolla - width | Seed - body texture |
| Apical leaflets - width | Corolla - surface | Seed - body indumentum |
| | | |



| Apical leaflets - texture | Corolla - apex shape | Seed - Seedling hypocotyl elongation |
|--|---|--------------------------------------|
| Apical leaflets - margin | Corolla mouth - at anthesis | |
| Apical leaflets - apex shape | Corolla mouth - indumentum | |
| Apical leaflets - base shape | Corolla mouth - Corolla indumentum (external) | |
| Leaflets - primary venation | Corolla mouth - Corolla indumentum (internal) at insertion of staminode | |
| Leaflets - Number of first veins | Corolla mouth - Corolla glands | |
| Leaflets - secondary venation | Corolla mouth - Corolla lobe aestivation | |
| Leaflets - tertiary venation | Stamens - shorter filaments lenght | |
| Abaxial side of the leaflets - Glands | Stamens - longer filaments lenght | |
| Abaxial side of the leaflets - Pellucid punctations | Stamens - position of the insertion of the stamens cycles | |
| Abaxial side of the leaflets - Domatia | Stamens - height of insertion of the longer stamens | |
| Abaxial side of the leaflets - Indument | Stamens - height of insertion of the shorter stamens | |
| Abaxial side of the leaflets - Hair type | Stamens - filaments indumentum | |
| Abaxial side of the leaflets - Indumentum position | Staminode - lenght | |
| Adaxial side of the leaflets - Glands | Staminode - indumentum | |

2 - Lista dos 85 descritores morfológicos utilizados:

| Category | Categorical Description |
|---------------------|------------------------------------|
| Vegetative Features | Dimorphic Growth Form |
| Vegetative Features | Phloem Wedges In Stem |
| Vegetative Features | Branchlet Section |
| Vegetative Features | Branchlet Diameter |
| Vegetative Features | Branchlet Indument |
| Vegetative Features | Branchlet Hair Type |
| Vegetative Features | Branchlet Lenticels |
| Vegetative Features | Branchlet Interpetiolar Gland |
| Vegetative Features | Branchlet Interpetiolar Ridge |
| Vegetative Features | Branchlet Odors |
| Vegetative Features | Prophylls Of Axillary Buds |
| Vegetative Features | Prophylls Of Axillary Buds Glands |
| Vegetative Features | Leaf Type |
| Vegetative Features | Leaf Tendril Type |
| Vegetative Features | Leaf Tendril Hook |
| Vegetative Features | Leaf Tendril Adesive Disk |
| Vegetative Features | Leaf Leaflets Texture |
| Vegetative Features | Leaf Leaflets Margin |
| Vegetative Features | Leaf Leaflets Glands |
| Vegetative Features | Leaf Leaflets Pellucid Punctations |
| Vegetative Features | Leaf Leaflets Domatia |
| Vegetative Features | Leaf Leaflets Primary Venation |
| Vegetative Features | Leaf Leaflets Secondary Venation |
| Vegetative Features | Leaf Leaflets Tertiary Venation |
| Flower | Inflorescence Position |



| | 45 |
|--------|---|
| Flower | Inflorescence Type |
| Flower | Inflorescence Bract Shape |
| Flower | Inflorescence Bract Glands |
| Flower | Calyx Shape |
| Flower | Calyx Apex |
| Flower | Calyx Double |
| Flower | Calyx Texture |
| Flower | Calyx Outer Indument |
| Flower | Calyx Glands |
| Flower | Calyx Arrangement In Bud |
| Flower | Corolla Color |
| Flower | Corolla Nectar Guides |
| Flower | Corolla Shape |
| Flower | Corolla Curvature |
| Flower | Corolla Texture |
| Flower | Corolla Lobes |
| Flower | Corolla Circunference |
| Flower | Corolla Outer Indument |
| Flower | Corolla Inner Indument At The Insertion Of Staminode |
| Flower | Corolla Glands |
| Flower | Corolla LobeAestivation |
| Flower | Corolla Mouth At Anthesis |
| Flower | Stamens Anther Indument |
| Flower | Stamens Anther Thecae |
| Flower | Stamens Anther Connective Texture |
| Flower | Stamens Anther Shape Of Connective |
| Flower | Stamens Corolla Indument At Insertion Of Staminode |
| Flower | Stamens Pollen Type |
| Flower | Stamens Pollen Apertures |
| | |

| Flower | Stamens Pollen Exine Ornamentation |
|--------------|---|
| Flower | Stamens Pollen Nectar Disk |
| Flower | Stamens Pollen Stipe At The Base Of Ovary |
| Flower | Ovary Surface |
| Flower | Ovary Indument |
| Flower | Ovary Number Of Ovule Series |
| Flower | Ovary Stigma Shape |
| Flower | Ovary Stigma Indument |
| Fruit | Splitting |
| Fruit | Shape |
| Fruit | Thickness |
| Fruit | Texture |
| Fruit | Curvature |
| Fruit | Indument |
| Fruit | Lenticels |
| Fruit | Glands |
| Fruit | Ridges |
| Fruit | Margin |
| Fruit | Wings |
| Fruit | Ornamentation |
| Fruit | Calyx In Fruit |
| Fruit | Seed Shape |
| Fruit | Seed Texture |
| Fruit | Seed Indument |
| Ecology | Floral Type |
| Ecology | Pollination Syndrome |
| Ecology | Nectar Robber Protection |
| Ecology | Dispersal Syndrome |
| Distribution | Habitat |
| Distribution | Distribution |
| l . | ı |

Instituto de Biociências/IB Escola Politécnica - POLI Universidade de São Paulo