

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DEPARTAMENTO ACADÊMICO DE COMPUTAÇÃO
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

ALEXANDRE YUJI KAJIHARA

**INTEGRAÇÃO DO HCF-*WEB* A HERBÁRIOS
VIRTUAIS BRASILEIROS**

MONOGRAFIA

CAMPO MOURÃO

2019

ALEXANDRE YUJI KAJIHARA

**INTEGRAÇÃO DO HCF-*WEB* A HERBÁRIOS
VIRTUAIS BRASILEIROS**

Trabalho de Conclusão de Curso de Graduação, apresentado à disciplina de Trabalho de Conclusão de Curso 2, do Curso de Bacharelado em Ciência da Computação do Departamento Acadêmico de Computação da Universidade Tecnológica Federal do Paraná, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

Orientador: Prof. Dr. André Luis Schwerz

Coorientador: Prof. Dr. Ivanilton Polato

CAMPO MOURÃO

2019



ATA DE DEFESA DO TRABALHO DE CONCLUSÃO DE CURSO

Às **20:00** do dia **2 de julho de 2019** foi realizada na sala **Mini-auditório EAD** da UTFPR-CM a sessão pública da defesa do Trabalho de Conclusão do Curso de Bacharelado em Ciência da Computação do(a) acadêmico(a) **Alexandre Yuji Kajihara** com o título **Integração do HCF-Web a herbários virtuais brasileiros**. Estavam presentes, além do(a) acadêmico(a), os membros da banca examinadora composta por: **Prof. Dr. André Luis Schwerz** (orientador), **Prof. Dr. Ivanilton Polato**, **Prof. Dr. Rafael Liberato Roberto** e **Prof. Dr. Marcelo Galeazzi Caxambu**. Inicialmente, o(a) acadêmico(a) fez a apresentação do seu trabalho, sendo, em seguida, arguido(a) pela banca examinadora. Após as arguições, sem a presença do(a) acadêmico(a), a banca examinadora o(a) considerou _____ na disciplina de Trabalho de Conclusão de Curso **2** e atribuiu, em consenso, a nota _____ (_____). Este resultado foi comunicado ao(à) acadêmico(a) e aos presentes na sessão pública. A banca examinadora também comunicou ao acadêmico(a) que este resultado fica condicionado à entrega da versão final dentro dos padrões e da documentação exigida pela UTFPR ao professor Responsável do TCC no prazo de **onze dias**. Em seguida foi encerrada a sessão e, para constar, foi lavrada a presente Ata que segue assinada pelos membros da banca examinadora, após lida e considerada conforme.

Observações: _____

Campo Mourão, **2 de julho de 2019**

Prof. Dr. Ivanilton Polato
Membro 1

Prof. Dr. Rafael Liberato Roberto
Membro 2

Prof. Dr. Marcelo Galeazzi
Caxambu
Membro 3

Prof. Dr. André Luis Schwerz
Orientador

A ata de defesa assinada encontra-se na coordenação do curso.

Agradecimentos

Ao Prof. Dr. André Luis Schwerz, que com sua experiência, competência, empenho, disponibilidade e paciência, orientou-me no desenvolvimento deste trabalho.

Ao Prof. Dr. Marcelo Galeazzi Caxambu, curador do Herbário da Universidade Tecnológica Federal do Paraná - Campus Campo Mourão, por ter possibilitado a realização deste estudo nessa instituição, auxiliado-me em questões técnicas relacionadas às coleções botânicas e contribuído na banca do TCC1.

Aos Profs. Drs. Ivanilton Polato e Rafael Liberato, pelas importantes contribuições feitas ao meu trabalho, na banca do TCC1.

Ao João Monnerat Lanna, do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, pelo suporte em relação às funcionalidades do Herbário Virtual Re flora.

Aos meus avós, Satomi Kajihara e Segunda Kajihara, pelo amor, carinho e dedicação ao longo da minha vida.

À minha mãe, Olinda Teruko Kajihara, que com seu amor, dedicação, apoio e incentivo, esteve sempre presente em minha trajetória pessoal e escolar.

Aos meus tios, Wilson Koji Kajihara e Cleici G. F. da Silva Kajihara, e ao meu primo, Augusto Koji Kajihara, pelo amor, companheirismo e apoio constante.

Resumo

KAJIHARA, Alexandre Yuji. **Integração do HCF-Web a herbários virtuais brasileiros**. 2019. 97. f. Monografia (Curso de Bacharelado em Ciência da Computação), Universidade Tecnológica Federal do Paraná. Campo Mourão, 2019.

O Herbário da Universidade Tecnológica Federal do Paraná, do Campus Campo Mourão, também conhecido como HCF, compartilha sua coleção, por meio de seu sistema *web* (HCF-Web), com o Herbário Virtual Reflora e o Instituto Nacional de Ciência e Tecnologia - Herbário Virtual da Flora e dos Fungos (INCT-HVFF). Os dois herbários virtuais aceitam sugestões de reidentificação de amostra, elaboradas por especialistas. No entanto, as sugestões não são atualizadas automaticamente, no sistema HCF-Web e, portanto, esses dados são comparados manualmente, com equivalentes, nos herbários virtuais. Como o HCF tem mais de 28 mil exsiccatas, a identificação de inconsistências exige muito tempo e é passível de erros. Este trabalho apresenta o desenvolvimento de uma extensão *web* para o HCF-Web, que compara dados aos seus equivalentes nos herbários virtuais, para que divergências possam ser destacadas. Para implementar essa solução, foi utilizada: (i) a API do Reflora, para atualização imediata ou programada; (ii) e um arquivo CSV, criado pelo sistema speciesLink, do INCT-HVFF. Para testar essa abordagem, nove amostras do HCF, variando de 5 mil a 25 mil espécimes, foram selecionadas, aleatoriamente, para validar as funções. Cada amostra foi testada três vezes para: (i) calcular a média do tempo de execução, em minutos; (ii) e verificar o número de divergências. Também foi coletado o número de falhas, durante a execução da API do Reflora. Como esperado, o tempo para a identificação da divergência variou linearmente, de acordo com o tamanho das amostras. Em relação ao INCT-HVFF, a conexão com a Internet não é necessária para comparar os dados do HCF com os do speciesLink. Portanto, a identificação de divergências é mais rápida que no processo do Reflora. Por fim, as contribuições desta pesquisa são: (i) uma solução para problemas de inconsistência de dados entre o HCF e os herbários virtuais; (ii) aumento da precisão quanto à identificação de divergências; (iii) aprimoramento do fluxo de trabalho do pessoal do HCF, uma vez que a comparação manual de dados, que normalmente leva semanas para ser executada, pode agora ser feita automaticamente, em menor tempo.

Palavras-chaves: HCF. HCF-Web. Herbário Virtual Reflora. INCT-HVFF. speciesLink.

Abstract

KAJIHARA, Alexandre Yuji. **Integration of HCF-Web to virtual herbariums in Brazil**. 2019. 97. f. Monograph (Undergraduate Program in Computer Science), Federal University of Technology - Paraná. Campo Mourão, PR, Brazil, 2019.

The Herbarium of the Federal University of Technology - Paraná, on the Campo Mourão Campus, also known as HCF, shares its collection through its web system (*HCF-Web*) with the Refflora Virtual Herbarium and the National Institute of Science and Technology - Virtual Herbarium of Flora and Fungi (INCT-HVFF). The two virtual herbariums accept suggestions on sample re-identification elaborated by experts. However, suggestions are not automatically updated into the *HCF-Web* system and, therefore, such data are manually compared with equivalents on virtual herbariums. Since HCF has more than twenty-eight thousand exsiccates, the identification of inconsistencies demands a long time and is error-prone. This work presents the development of an *HCF-Web* extension that compares data to their equivalents in virtual herbariums so that divergences can be highlighted. In order to implement this solution, it was used: (i) the Refflora's API for immediate or programmed updating; (ii) and a CSV file created by the speciesLink system from INCT-HVFF. To test this approach, nine HCF samples, ranging from five to twenty-five thousand specimens were randomly selected to validate functions. Each sample was tested three times to: (i) calculate mean execution time, in minutes; (ii) and verify the number of divergences. It was also gathered the number of fails during Refflora's API execution. As expected, time for the divergence identification varied linearly according to the samples' size. Regarding the INCT-HVFF, the Internet connection is not required to compare HCF and speciesLink data. Therefore, identification of divergences is faster than in the Refflora process. Finally, this research contributions are: (i) a solution to issues on data inconsistency between HCF and virtual herbariums; (ii) increase in precision concerning the identification of divergences; (iii) enhancement of the workflow of the HCF personnel, since manual data comparison, which usually took weeks to be performed, can now be automatically made in a shorter time.

Keywords: HCF. *HCF-Web*. Refflora Virtual Herbarium. INCT-HVFF. speciesLink.

Lista de figuras

2.1	Sistemas de <i>software</i> que não são bancos de dados nem planilhas, utilizados pelas coleções participantes da rede speciesLink.	19
2.2	Interface de acesso do BRAHMS, em que é preciso optar entre monousuário e multiusuário.	21
2.3	Interface de configuração do Specify 6. Devem ser preenchidos os campos com o usuário <i>root</i> e o nome do banco de dados que será utilizado.	22
2.4	Interface de configuração do IPT, para conectá-lo a uma coleção.	24
2.5	Interface do SICol, com itens renderizados, em resposta à busca realizada. . .	25
2.6	Interface do speciesBase, com as quatro funcionalidades disponíveis.	26
2.7	Interface de consulta do Jabot, em que é possível selecionar coleção, família, gênero, espécie, localização e coletor.	27
3.1	Estrutura de diretórios do <i>back end</i> - do HCF- <i>Web</i> - formado por: <i>docs</i> (informações do projeto); <i>node_modules</i> (módulos utilizados na aplicação); <i>scripts</i> (código de comunicação com serviço auxiliar); <i>src</i> (código de execução do <i>back end</i>); e <i>test</i> (códigos que verificam se resultados esperados estão sendo gerados).	33
3.2	Estrutura de diretórios do <i>front end</i> - do HCF- <i>Web</i> - formado por: <i>node_modules</i> (módulos utilizados no <i>front end</i>); <i>public</i> (arquivos estáticos presentes no servidores retornados aos clientes); e <i>src</i> (código de execução do <i>front end</i>).	34
3.3	Diagrama ER, do HCF- <i>Web</i> . Cada tombo é o registro de uma amostra da coleção, e a relação que se destaca é a <i>tombo</i> , por causa da quantidade de relacionamentos.	35
3.4	Hierarquia dos usuários no HCF- <i>Web</i> . O curador, que está no topo da pirâmide, tem acesso a todas as funcionalidades. Nos níveis abaixo, de acordo com a prioridade, operadores e visitantes possuem acesso mais restrito.	36
3.5	Interface de listagem de usuários do HCF- <i>Web</i> , de acesso exclusivo ao curador. Além de visualizar, é possível adicionar, buscar, editar e apagar usuários. . .	36
3.6	Cadastro de remessa, realizado para permuta, troca ou doação de exsicatas. .	38

3.7	Exemplo de exsicata do HCF, afixada em cartolina e acompanhada de etiqueta colada no canto inferior direito.	39
3.8	Diagrama de caso de uso do HCF- <i>Web</i> , em que o ator Curador tem acesso a todas as funcionalidades; o Operador utiliza algumas funções, com exceção de criação, remoção, atualização e visualização de usuários e aprovação de pendências; e o Visitante somente visualiza os tombos.	40
4.1	Exemplo de arquivo JSON retornado do Re flora, quando ocorre um erro inesperado no servidor.	44
4.2	Exemplo de arquivo JSON, retornado pela API quando a exsicata solicitada não está disponível no Herbário Virtual Re flora.	45
4.3	Exemplo de arquivo JSON, retornado do Re flora para a exsicata do HCF apresentada na Figura 3.7.	48
5.1	Diagrama de atividades, que descreve a atualização do HCF- <i>Web</i> , a partir dos dados do Herbário Virtual Re flora.	55
5.2	Interface com uma pendência decorrente de uma alteração sugerida no Herbário Virtual Re flora Re flora, e que foi encontrada pelo serviço.	58
5.3	Processo de atualização do HCF- <i>Web</i> , em que dados de seu acervo são enviados via <i>script</i> Darwin Core ao Re flora, que retorna informações equivalentes, por meio de sua API, para que a comparação seja realizada.	58
5.4	Diagrama das atividades para a atualização do HCF- <i>Web</i> , a partir dos dados do speciesLink.	59
5.5	Interface com uma pendência decorrente de uma alteração sugerida no speciesLink, e que foi encontrada pelo serviço.	61
5.6	Processo de atualização do HCF- <i>Web</i> , em que dados de seu acervo são enviados via <i>software</i> spLinker ao speciesLink, que retorna informações equivalentes, por meio de um arquivo, para que a comparação seja realizada.	61
5.7	Interface da funcionalidade relacionada ao Herbário Virtual Re flora, que oferece opções de atualização (imediate ou programada) e de verificação de log. . . .	62
5.8	<i>Feedback</i> dado ao usuário, quando uma atualização imediata é solicitada. . .	63
5.9	Interface da funcionalidade relacionada ao sistema speciesLink, que oferece opções de submissão de arquivos e verificação de logs.	64
5.10	Diagrama de caso de uso, com a implementação dos dois serviços, em que dois novos atores foram adicionados.	65
5.11	Exemplo de arquivo de log, contendo os registros das falhas que ocorreram durante as três requisições ao Re flora.	67
5.12	Médias e desvios-padrões dos tempos de atualização de cada amostra do HCF- <i>Web</i> , em relação aos dados equivalentes no Herbário Virtual Re flora. .	69

5.13 Médias e desvios-padrões dos tempos de atualização de cada amostra do HCF- <i>Web</i> , em relação aos dados equivalentes no speciesLink.	71
--	----

Lista de tabelas

5.1	Atributos da relação <i>reflora</i>	56
5.2	Chaves do JSON retornadas da requisição ao <i>Reflora</i> , e suas respectivas equivalências, nas relações e atributos do <i>HCF-Web</i>	57
5.3	Colunas do arquivo do <i>speciesLink</i> e suas respectivas equivalências, nas relações e atributos do <i>HCF-Web</i>	60
5.4	Atributos da relação <i>configuracao</i>	66
5.5	Atributos da relação <i>configuracao</i> . (continuação)	67
5.6	Médias das divergências relacionadas ao <i>Herbário Virtual Reflora</i>	70
5.7	Médias das falhas relacionadas ao <i>Herbário Virtual Reflora</i>	70
5.8	Médias das divergências relacionadas ao <i>speciesLink</i>	72
A.1	Atributos da relação <i>tombos</i>	82
A.2	Atributos da relação <i>tombos</i> . (continuação)	83
A.3	Atributos da relação <i>tombos_fotos</i>	84
A.4	Atributos da relação <i>colecoes_anexas</i>	84
A.5	Atributos da relação <i>alteracoes</i>	85
A.6	Atributos da relação <i>tombo_alteracoes_antigas</i>	85
A.7	Atributos da relação <i>retirada_exsiccta_tombos</i>	86
A.8	Atributos da relação <i>remessas</i>	86
A.9	Atributos da relação <i>usuarios</i>	87
A.10	Atributos da relação <i>tipos_usuarios</i>	87
A.11	Atributos da relação <i>historico_acessos</i>	87
A.12	Atributos da relação <i>tombos_coletores</i>	88
A.13	Atributos da relação <i>coletores</i>	88
A.14	Atributos da relação <i>locais_coleta</i>	89
A.15	Atributos da relação <i>estados</i>	90
A.16	Atributos da relação <i>paises</i>	90
A.17	Atributos da relação <i>solos</i>	91
A.18	Atributos da relação <i>relevos</i>	91
A.19	Atributos da relação <i>fase_sucessional</i>	91
A.20	Atributos da relação <i>vegetacoes</i>	92

A.21 Atributos da relação autores	92
A.22 Atributos da relação generos	93
A.23 Atributos da relação familias	93
A.24 Atributos da relação sub_familias	93
A.25 Atributos da relação especies	94
A.26 Atributos da relação sub_especies	94
A.27 Atributos da relação variedades	95
A.28 Atributos da relação tipos	95
A.29 Atributos da relação herbarios	96
A.30 Atributos da relação telefonos	96
A.31 Atributos da relação enderecos	97

Siglas

API:	<i>Application Programming Interface</i>
BRAHMS:	<i>The Botanical Research and Herbarium Management System</i>
CBA:	<i>Component-Based Architecture</i>
CERN:	<i>European Center for Nuclear Research</i>
CRUD:	<i>Create, Read, Update and Delete</i>
CSV:	<i>Comma-Separated Values</i>
DOM:	<i>Document Object Model</i>
GBIF:	<i>Global Biodiversity Information Facility</i>
HTML:	<i>HyperText Markup Language</i>
HTTP:	<i>HyperText Transfer Protocol</i>
INCT-HVFF:	Instituto Nacional de Ciência e Tecnologia - Herbário Virtual da Flora e dos Fungos
IPT:	<i>Integrated Publishing Toolkit</i>
Jabot:	Jardim Botânico
JBRJ:	Jardim Botânico do Rio de Janeiro
JS:	<i>JavaScript</i>
JSON:	<i>JavaScript Object Notation</i>
MCTI:	Ministério da Ciência, Tecnologia e Inovação
MCTIC:	Ministério da Ciência, Tecnologia, Inovações e Comunicações
NPM:	<i>Node Package Manager</i>
ODBC:	<i>Open Database Connectivity</i>
ODS:	<i>OpenDocument Spreadsheet</i>
ORM:	<i>Object Relational Mapping</i>
RBH:	Rede Brasileira de Herbários
REST:	<i>Representational State Transfer</i>
RFC:	<i>Request For Comments</i>
SBB:	Sociedade Botânica do Brasil
SiBBR:	Sistema de Informação sobre a Biodiversidade Brasileira
SICol:	Sistema de Informação de Coleções de Interesse Biotecnológico
SQL:	<i>Structured Query Language</i>
UFPE:	Universidade Federal de Pernambuco

URI: *Uniform Resource Identifier*

URL: *Uniform Resource Locator*

UTFPR: Universidade Tecnológica Federal do Paraná

XML: *Extensible Markup Language*

Sumário

1	Introdução	14
2	Trabalhos relacionados	18
2.1	BRAHMS	19
2.2	Specify	21
2.3	IPT	23
2.4	SICol	24
2.5	speciesBase	25
2.6	Jabot	26
2.7	Discussão	28
3	HCF- <i>Web</i>	29
3.1	Tecnologias	29
3.2	Estruturas de diretórios do HCF- <i>Web</i>	32
3.3	Funcionalidades do HCF- <i>Web</i>	34
3.4	Discussão	41
4	Herbários virtuais	42
4.1	Herbário Virtual Reflora	43
4.2	INCT-HVFF	49
4.3	Discussão	52
5	Integração do HCF- <i>Web</i> a herbários virtuais	54
5.1	Herbário Virtual Reflora	54
5.2	Sistema speciesLink	59
5.3	Interfaces	62
5.4	Decisões de projeto	65
5.5	Estimativa de tempo de execução	68
6	Conclusão	73
	Referências	75

Apêndices	81
A Atributos das relações do HCF- <i>Web</i>	82

Introdução

Entre os 13 campi da Universidade Tecnológica Federal do Paraná (UTFPR), Campo Mourão foi o primeiro a formar uma coleção biológica, composta, inicialmente, por espécies de arborização urbana (GLOBAL BIODIVERSITY INFORMATION FACILITY, 2018b). Essas plantas foram reunidas em uma instituição que recebeu o nome de Herbário da Universidade Tecnológica Federal do Paraná - Campus Campo Mourão, conhecido pelo acrônimo HCF (Herbário do Centro Federal)¹. Na época de sua criação, em 2000, o HCF era um laboratório de apoio ao Curso de Tecnologia em Gestão Ambiental (UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ, 2015), e atualmente está vinculado à direção geral do campus de Campo Mourão.

Até o ano de 2003, a coleção do HCF era composta de apenas 500 exsicatas². A partir dessa época, com o apoio da direção do campus de Campo Mourão, começaram a ser realizadas coletas regulares, no Paraná e em outros estados, o que contribuiu para o crescimento do acervo (GLOBAL BIODIVERSITY INFORMATION FACILITY, 2018b). Além disso, o governo federal modificou sua posição: antes consideradas ônus - em espaço e custo de manutenção - as coleções de herbários passaram a ser valorizadas, pois a biopirataria estava roubando espécies brasileiras, transformando-as em químicos e fármacos patenteados que eram revendidos ao país. No ano de 2013, o acervo do HCF já contava com mais de 12 mil plantas secas (POLISELI, 2013), e três anos depois, em 2016, atingiu o número de 22,5 mil amostras, o que a tornou a sexta maior coleção botânica do Paraná (UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ, 2017). Atualmente, a coleção do HCF é de

¹ Herbário é uma coleção de amostras de plantas ou de fungos, normalmente secos ou preservados em meio líquido, que tem como objetivo documentar a diversidade vegetal ou fúngica de uma região, país ou continente (PEIXOTO; MAIA, 2013).

² Exsicatas são amostras secas de plantas ou fungos, que fazem parte de uma coleção científica de um herbário (PEIXOTO; MAIA, 2013).

cerca de 28 mil espécimes³ (SPECIESLINK, 2018a).

Além da expansão do acervo, outro avanço do HCF, ao longo das duas últimas décadas, foi a inclusão de sua coleção, em 2007, em importantes catálogos, como o da Rede Brasileira de Herbários (RBH), ligada à Sociedade Botânica do Brasil (SBB). No ano de 2015, a coleção do HCF foi registrada e disponibilizada na Plataforma Global de Informação sobre Biodiversidade (GBIF, do inglês *Global Biodiversity Information Facility*) (UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ, 2015), que consiste na maior iniciativa mundial, já realizada, de integração e acesso virtual e gratuito de dados sobre a biodiversidade do planeta (SISTEMA DE INFORMAÇÃO SOBRE A BIODIVERSIDADE BRASILEIRA, 2018a).

O aumento da coleção do HCF, no decorrer do tempo, gerou a necessidade de informatização do acervo. Dessa forma, o *software* Herbarium 1.0 foi desenvolvido por Carneiro (2006) e, ao longo de uma década, tem sido utilizado para o registro e armazenamento das informações referentes às exsicatas. Entretanto, esse sistema não atende a uma necessidade recente, de conversão dos dados do HCF em padrão Darwin Core. Essa padronização é indispensável para que o HCF possa participar do Sistema de Informação sobre a Biodiversidade Brasileira (SiBBBr)⁴, nó brasileiro da GBIF que integra e disponibiliza informações sobre a biodiversidade do país coletadas e publicadas por universidades, institutos de pesquisa e agências governamentais (GADELHA JUNIOR et al., 2014).

Para resolver o problema da conversão dos dados do HCF em padrão Darwin Core, Sangali (2019) propôs algumas reformulações no Herbarium, e a nova versão *web* do sistema recebeu o nome de HCF-*Web*. Atualmente, a coleção do HCF está disponível no Herbário Virtual Refflora⁵ e no Instituto Nacional de Ciência e Tecnologia - Herbário Virtual da Flora e dos Fungos (INCT-HVFF), que utiliza a base informacional do sistema speciesLink⁶.

Uma funcionalidade importante desses dois herbários virtuais é permitir que

³ Espécime é um indivíduo - exemplar ou amostra de planta, animal ou micro-organismo - representativo de uma categoria taxonômica, como classe, gênero ou espécie (CERQUEIRA et al., 2005). O termo espécime difere de espécie, que é uma das categorias taxonômicas - juntamente com reino, filo, classe, ordem, família e gênero - e denomina um conjunto de indivíduos que têm características essenciais semelhantes, e que reproduzem-se entre si (PIRES et al., 2018).

⁴ O SiBBBr é uma iniciativa do Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC), e recebe suporte técnico do Programa das Nações Unidas para o Meio Ambiente e apoio financeiro do Fundo Global para o Meio Ambiente. A plataforma *on-line* SiBBBr pretende dar suporte para pesquisas científicas, cujos resultados subsidiarão a formulação de políticas públicas e tomada de decisões relacionadas à conservação ambiental e ao uso sustentável de recursos naturais (SISTEMA DE INFORMAÇÃO SOBRE A BIODIVERSIDADE BRASILEIRA, 2018b)

⁵ <http://reflora.jbrj.gov.br/reflora/herbarioVirtual/ConsultaPublicoHVUC/ConsultaPublicoHVUC.do>

⁶ O Centro de Referência em Informação Ambiental (CRIA), uma Organização da Sociedade Civil de Interesse Público (CENTRO DE REFERÊNCIA EM INFORMAÇÃO AMBIENTAL, 2018a), com o apoio financeiro da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), desenvolveu o Projeto “Sistema de Informação Distribuído para Coleções Biológicas: a Integração do SpeciesAnalyst e do SinBiota - speciesLink” (SPECIESLINK, 2018b). A rede speciesLink visa integrar e disponibilizar, de forma livre e aberta, na Internet, dados de coleções biológicas do Brasil e do exterior (CENTRO DE REFERÊNCIA EM INFORMAÇÃO AMBIENTAL, 2018b).

especialistas cadastrados contribuam com sugestões de reidentificações de amostras, ao constatarem incorreções, por exemplo, na determinação do gênero ou da espécie de uma planta. A possibilidade de reidentificação é fundamental para a melhoria da qualidade dos dados dos acervos das coleções biológicas do Brasil, ou seja, para o aumento da confiabilidade das identificações e da validade dos dados referentes aos registros de cada espécime.

Considerando que o número de taxonomistas no Brasil ainda é pequeno⁷, e que não há especialistas em todos os grupos taxonômicos⁸, a possibilidade de redução do número de espécimes não identificados nos acervos dos herbários, de correção e de atualização das identificações, a partir de sugestões feitas por pesquisadores brasileiros e estrangeiros, cadastrados no Herbário Virtual Re flora e no INCT-HVFF, é um grande passo para a melhoria da qualidade dos dados das coleções biológicas do país.

A reidentificação das amostras do HCF, até o ano de 2018, gerava registro apenas no banco de dados do herbário virtual onde a sugestão fora realizada, ou seja, no Re flora ou no sistema speciesLink, do INCT-HVFF. Isso produzia inconsistências no HCF-*Web*. Para que as amostras do HCF não ficassem com informações incorretas, era preciso que um funcionário comparasse, manualmente, todos os registros presentes nos herbários virtuais, campo por campo, com os seus equivalentes no HCF-*Web*. Considerando o elevado número de exsicatas da coleção do HCF, essa comparação manual exigia muito tempo, portanto, não podia ser realizada periodicamente. Esse trabalho árduo de busca de inconsistências podia propiciar erros, por causa da grande quantidade de dados e campos que precisavam ser comparados, e da existência de alterações sutis nos registros, que poderiam passar despercebidas.

Assim, neste trabalho foi implementada uma extensão para o HCF-*Web*, que permite atualizar, automática e periodicamente, os dados de suas amostras, comparando-os com os seus equivalentes, no Herbário Virtual Re flora e no sistema speciesLink, do INCT-HVFF, a fim de encontrar inconsistências.

Para a realização dessa extensão, foram utilizadas algumas ferramentas disponibilizadas pelos herbários virtuais, que possibilitam localizar dados referentes ao HCF. O sistema do Herbário Virtual Re flora oferece uma Interface de Programação de Aplicativo (API, do inglês *Application Programming Interface*), e com a extensão implementada, o usuário do HCF-*Web* pode optar entre a atualização imediata ou programada. O sistema speciesLink, do INCT-HVFF, oferece um arquivo com informações das exsicatas disponibilizadas pelos herbários associados, e o usuário deve submetê-lo para encontrar inconsistências.

Pretendeu-se, com a implantação dessas novas funcionalidades: (i) contribuir para otimizar o tempo dos funcionários do HCF em seu trabalho; (ii) tornar mais preciso o

⁷ No ano de 2006, entre os 270 taxonomistas doutores existentes no Brasil, 91 eram bolsistas de produtividade do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq): 55 em angiospermas, 1 em gimnosperma, 5 em pteridófitas, 2 em briófitas, 19 em algas e 10 em fungos (PEIXOTO et al., 2006).

⁸ Grupos taxonômicos ou táxons, são categorias taxonômicas de qualquer nível hierárquico. Os principais táxons, em ordem decrescente de hierarquia, são: reino, divisão ou filo, classe, ordem, família, gênero e espécie (MCNEILL et al., 2013).

resultado da busca de divergências, especialmente daquelas difíceis de serem visualizadas; (iii) e acabar com as inconsistências entre os registros disponíveis no HCF- *Web* e seus equivalentes, nos herbários virtuais.

O restante desta monografia está organizada da seguinte forma: no Capítulo 2 são apresentados trabalhos relacionados, ou seja, as funcionalidades dos seis *software* mais utilizados pelas instituições que participam da rede speciesLink, para o gerenciamento de suas coleções biológicas; no Capítulo 3 são descritas as funcionalidades e tecnologias aplicadas no HCF- *Web* e a sua estrutura de diretórios; no Capítulo 4 são apresentados aspectos históricos do Herbário Virtual Re flora e do INCT-HVFF, e as funcionalidades de seus sistemas, utilizadas para a implementação da extensão para o HCF- *Web*; no Capítulo 5 são descritos os processos de integração do HCF- *Web* aos sistemas dos dois herbários virtuais, as interfaces elaboradas e os testes realizados para estimar os tempos de execução para a identificação das divergências entre os bancos de dados de cada herbário virtual e o do HCF- *Web*; por fim, no Capítulo 6 é apresentada a conclusão deste trabalho de integração do HCF- *Web* aos sistemas do Re flora e do INCT-HVFF.

Trabalhos relacionados

Atualmente, existem vários tipos de *software* destinados exclusivamente ao gerenciamento de banco de dados de biodiversidade: alguns são abrangentes e podem ser utilizados para gestão de diferentes tipos de coleções biológicas; outros são mais específicos, para determinado tipo de coleção, por exemplo, de plantas. Para a escolha de um sistema de gestão de coleções, devem ser analisados: (i) a usabilidade da ferramenta; (ii) se apresenta uma interface intuitiva, pois, frequentemente, a equipe de trabalho é formada por bolsistas temporários que não têm muito tempo para dispendir na aprendizagem de uma ferramenta muito complexa; (iii) a capacidade de suporte do sistema para importar dados pré-existentes - como os de uma planilha de Excel ou de um *software* anterior - e de exportá-los (SISTEMA DE INFORMAÇÃO SOBRE A BIODIVERSIDADE BRASILEIRA, 2018c).

Os participantes da rede speciesLink gerenciam suas coleções biológicas por meio de planilhas eletrônicas, banco de dados e *software* específicos. No final do ano de 2018, quando foi realizado o levantamento para este trabalho, entre as 486 coleções botânicas, zoológicas e microbiológicas disponíveis na rede speciesLink, 484 forneciam os nomes das aplicações utilizadas para o gerenciamento dos acervos, e 55 sistemas de *software* diferentes eram empregados (SPECIESLINK, 2018c).

Nesse subgrupo de 484 coleções, é interessante observar que 87 utilizavam bancos de dados (MySQL¹, PostgreSQL² e Microsoft Access³), o que exige conhecimento da Linguagem de Consulta Estruturada (SQL, do inglês *Structured Query Language*). Outras 182 coleções gerenciavam suas coleções por meio de planilhas eletrônicas, como o LibreOffice Calc⁴, Microsoft Excel⁵ e BrOffice Calc⁶. As demais coleções - 215 - utilizavam os sistemas de

¹ <https://www.mysql.com>

² <https://www.postgresql.org/>

³ <https://products.office.com/pt-br/access>

⁴ <https://pt-br.libreoffice.org/descubra/calc/>

⁵ <https://products.office.com/pt-br/excel>

⁶ A marca BrOffice foi extinta e seu nome foi substituído por LibreOffice (GULGEMIN, 2011).

software apresentados na Figura 2.1, que não são bancos de dados nem planilhas eletrônicas.

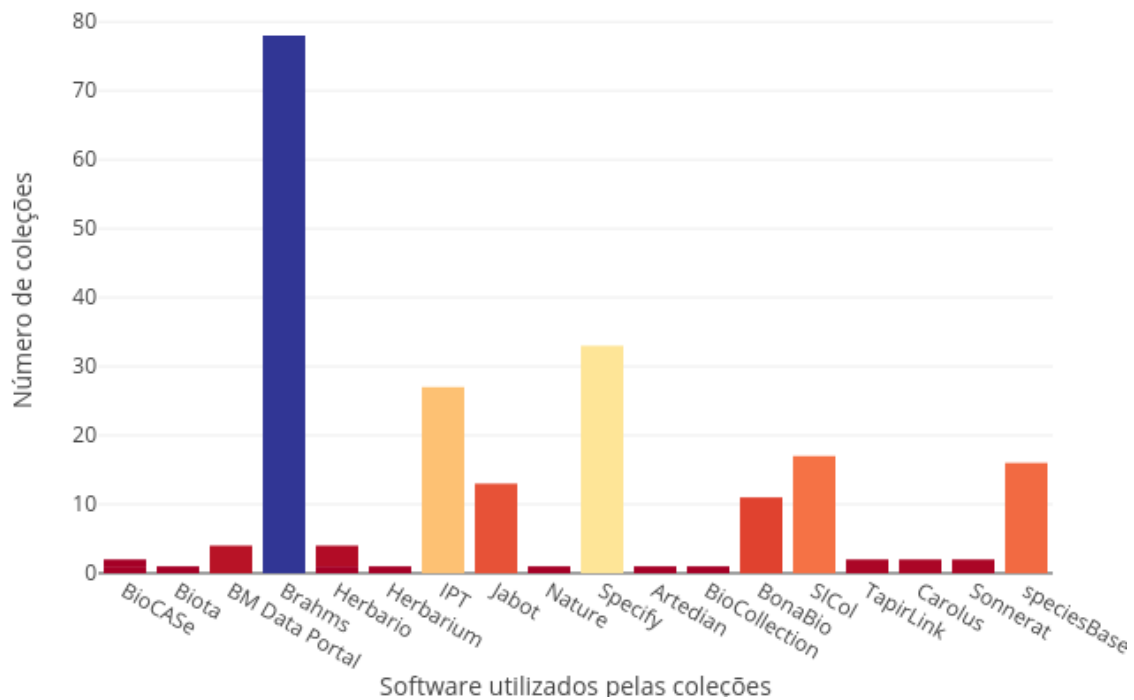


Figura 2.1. Sistemas de *software* que não são bancos de dados nem planilhas, utilizados pelas coleções participantes da rede speciesLink.

Na Figura 2.1 é possível observar que, em 2018, os sete *software* mais utilizados pelos participantes da rede speciesLink foram: *The Botanical Research and Herbarium Management System* (BRAHMS)⁷, Specify⁸, *Integrated Publishing Toolkit* (IPT)⁹, Sistema de Informação de Coleções de Interesse Biotecnológico¹⁰ (SICOI), speciesBase¹¹, Sistema de Gerenciamento de Coleções Botânicas - Jardim Botânico (Jabot)¹² e o BonaBio. Considerando que não foram encontrados sites, documentações ou artigos sobre o BonaBio, nas próximas seções são apresentadas funcionalidades presentes nos outros seis *software* mais utilizados pelos participantes da rede speciesLink. Na Seção 2.7 é realizada uma breve discussão sobre esses sistemas de *software* utilizados para o gerenciamento de coleções biológicas.

2.1. BRAHMS

O *software* mais utilizado pelos participantes da rede speciesLink é o BRAHMS (SPECIESLINK, 2018c). Em 2018, 78 coleções utilizavam esse sistema de gerenciamento de dados

⁷ <https://herbaria.plants.ox.ac.uk/bol/>

⁸ <http://www.sustain.specifysoftware.org/>

⁹ <https://www.gbif.org/ipt>

¹⁰ <http://sicol.cria.org.br/>

¹¹ <http://splink.cria.org.br/speciesbase?criaLANG=pt>

¹² <http://jabot.jbrj.gov.br/v2/consulta.php>

desenvolvido pelo Departamento de Ciências Vegetais, da Universidade de Oxford ¹³. O BRAHMS é utilizado em 80 países, e o Brasil se destaca por ter o maior número de projetos individuais com esse *software* (UNIVERSITY OF OXFORD, 2018a).

O Herbário do Instituto Agrônomo do Norte (IAN), da Embrapa Amazônia Oriental, foi pioneiro no Brasil, na utilização do BRAHMS. Isso ocorreu em 1996, quando a Universidade de Oxford iniciou uma consultoria para a criação do banco de dados desse herbário. A partir do ano de 2000, o IAN passou a colaborar no desenvolvimento do BRAHMS (PEIXOTO, 2005).

Em 2001, o Herbário Bradeanum¹⁴ começou a utilizar o BRAHMS para a informatização de sua coleção. A opção por esse *software* se deu: (i) pelo seu baixo custo; (ii) pela compatibilidade com o Microsoft Excel e com o Microsoft Access; (iii) pela prestação de suporte permanente, por meio de e-mail, de manual atualizado e de consultoria presencial; (iv) e pelo fato de outros herbários brasileiros terem adotado esse *software*, o que facilitava a troca de experiências (PEIXOTO, 2005).

O BRAHMS permite multiusuários, como pode ser visto na Figura 2.2, porém é necessário habilitar essa opção junto à Universidade de Oxford (UNIVERSITY OF OXFORD, 2018b). Esse *software* só pode ser utilizado no sistema operacional Windows, e algumas de suas funcionalidades são (UNIVERSITY OF OXFORD, 2013):

- realiza *Create, Read, Update and Delete* (CRUD): de registros botânicos e coletores; de herbários, para poder utilizar a funcionalidade transações (que são empréstimos entre herbários); de países, com adição de diferentes pronúncias de seus nomes, e criação de grupos com os países cadastrados; de taxonomias (identificação de família, gênero e espécie) e autor (nome da primeira pessoa que descreveu a espécie, ou seja, que a identificou); de tipos nomenclaturais ¹⁵; e de arquivos, fotos e localizações referentes aos registros botânicos;

¹³ <https://herbaria.plants.ox.ac.uk/bol/oxford/brahms>

¹⁴ O Herbário Bradeanum é uma entidade privada, sem fins lucrativos, fundado há mais de 50 anos, e que tem grande reconhecimento mundial na área de pesquisa taxonômica. Há cerca de 20 anos, a Universidade Estadual do Rio de Janeiro, por meio de um convênio de colaboração mútua, acolheu o importante acervo da flora fluminense do Herbário Bradeanum (UNIVERSIDADE ESTADUAL DO RIO DE JANEIRO, 2009).

¹⁵ Tipo nomenclatural (tipo de nome de táxon ou *typus*) é a amostra botânica que foi utilizada na primeira descrição científica de um determinado táxon (MARTINS et al., 2006), ou seja, que serviu de base para o nome dado a uma nova família, gênero ou espécie de planta. Existem seis categorias de tipos nomenclaturais. Por exemplo, um espécime ou mesmo uma ilustração botânica, é considerado: (i) do tipo holótipo, quando foi escolhido pelo autor para representar a nova espécie descrita; (ii) do tipo isótipo, quando for uma duplicata da nova espécie, isto é, uma porção da coleta do holótipo, geralmente depositada em um herbário diferente; (iii) e do tipo parátipo, quando for de uma coleta diferente do holótipo, mas foi examinado pelo autor, e citado na obra em que a nova espécie foi descrita (LÍRIO; SARNAGLIA JÚNIOR, 2014).

- mapeia dados presentes no BRAHMS para obter informações geográficas, por meio de programas, como o Google Earth¹⁶/Maps¹⁷, DIVA GIS¹⁸ e Manifold Mapping¹⁹.



Figura 2.2. Interface de acesso do BRAHMS, em que é preciso optar entre monusuário e multiusuário.

2.2. Specify

O Specify, *software* para a catalogação e gestão de coleções botânicas e zoológicas, criado pela Universidade de Kansas, era utilizado, em 2018, por 33 coleções participantes da rede speciesLink, e a versão 6, a mais empregada (SPECIESLINK, 2018g). Ele é um *software open-source*²⁰, e o seu repositório tem uma comunidade ativa, que conta com 5 contribuidores, mais de 10 mil *commits* e mais de 210 *releases*. A próxima geração do Specify - em versão *web* - está em desenvolvimento, e também será *open-source*²¹, com todas as funcionalidades da anterior. O repositório da versão 7 também terá uma comunidade ativa, que contará com mais: 5 contribuidores, 2 mil *commits* e 10 *releases* (SPECIFY, 2018). Na Figura 2.3 é

¹⁶ <https://www.google.com/earth/>

¹⁷ <https://www.google.com.br/maps>

¹⁸ <http://www.diva-gis.org/>

¹⁹ <http://www.manifold.net/>

²⁰ <https://github.com/specify/specify6>

²¹ <https://github.com/specify/specify7>

apresentada a interface de configuração, necessária para utilizar o Specify 6.

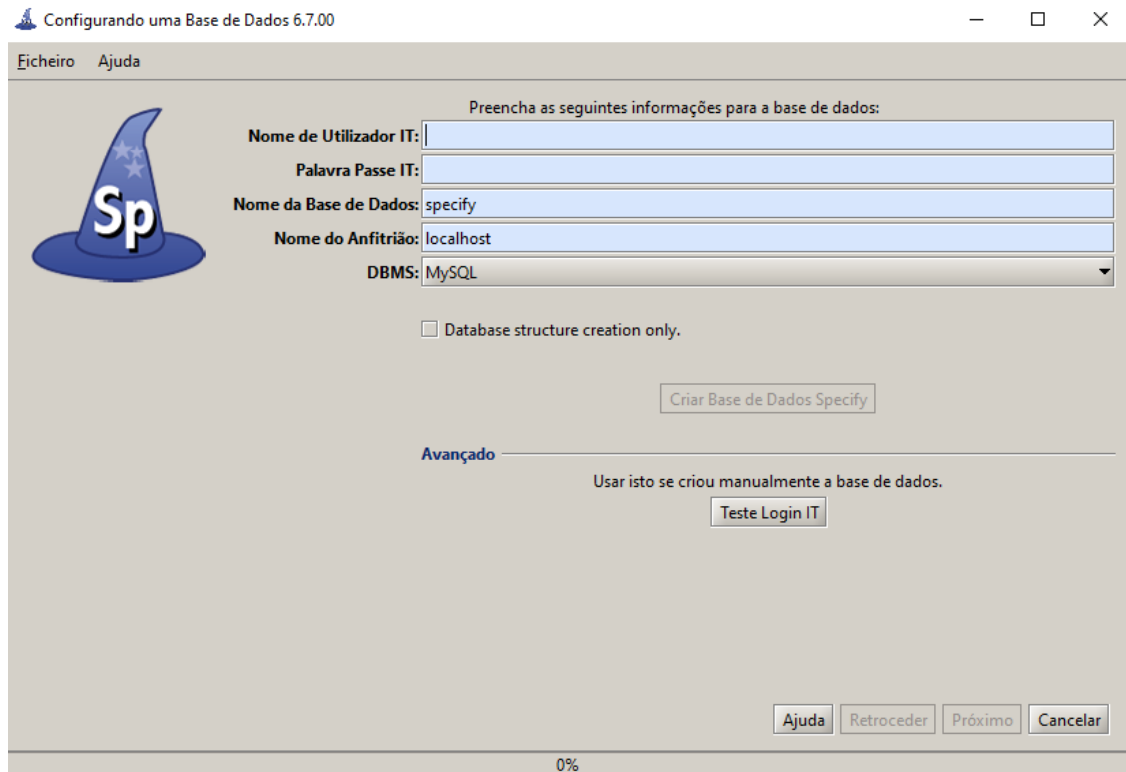


Figura 2.3. Interface de configuração do Specify 6. Devem ser preenchidos os campos com o usuário root e o nome do banco de dados que será utilizado.

Essa aplicação é escrita em Java, pode ser executada nos sistemas operacionais Windows, Linux e Mac OS X, e a versão 6 possui as seguintes funcionalidades (UNIVERSITY OF KANSAS, 2013):

- carrega as coleções de dados escolhidas, configura o relacionamento dessas coleções; por exemplo, taxonomia pode estar presente em mais de uma coleção;
- realiza CRUD de instituições e adiciona hierarquias presentes nelas; usuário define permissões; usuários cadastrados podem ser adicionados em um grupo e dar permissão a outro grupo; realiza e solicita empréstimos para instituições, para realização de pesquisa;
- exibe localizações no Google Earth, com base nos registros;
- renderiza árvore hierárquica de certos tipos de dados, a partir das tabelas táxon e espécimes armazenados;
- realiza pesquisa por meio de: (i) campos pré-definidos; (ii) consulta ao banco de dados.

2.3. IPT

Em 2018, o IPT²² era utilizado por 27 coleções participantes da rede speciesLink. Essa ferramenta é *open-source*²³ e escrita em Java²⁴, e a documentação²⁵ sugere o uso do Apache Tomcat²⁶. Seu repositório tem uma comunidade ativa, com 14 contribuidores, mais de 4 mil *commits* e 23 *releases*. O IPT foi desenvolvido pela GBIF, e visa publicar e compartilhar dados de biodiversidade por meio da GBIF (GLOBAL BIODIVERSITY INFORMATION FACILITY, 2018a).

Como o IPT tem escassa documentação disponível, mas é *open-source*, foi realizado o *download* de seu código-fonte. O *software* foi, então, configurado em um servidor, para verificação de suas funcionalidades: (i) o administrador e os usuários adicionados podem incluir vários herbários, o que permite criar uma rede de instituições associadas; (ii) o administrador tem acesso a todos os herbários cadastrados, e os demais usuários, somente aos que foram criados por eles.

Algumas informações disponíveis quando da inclusão são: sobre o herbário e suas amostras (área geográfica da coleta e taxonomia), mapeamento dos dados para o formato Darwin Core, licença relacionada aos dados e modelo a ser utilizado para referenciar o herbário, em trabalhos científicos. Algumas dessas opções podem ser vistas na Figura 2.4. Esses dados disponibilizados podem ser acessados por meio da URL definida pelo administrador. Os visitantes podem escolher um dos herbários associados e verificar as informações disponibilizados por quem realizou o cadastramento, e também fazer *download* das amostras da coleção.

²² <https://www.gbif.org/ipt>

²³ <https://github.com/gbif/ipt>

²⁴ <https://docs.oracle.com/javase/8/docs/api/>

²⁵ <https://github.com/gbif/ipt/wiki/IPT2ManualNotes.wiki#about-the-gbif-integrated-publishing-toolkit>

²⁶ <https://tomcat.apache.org/>

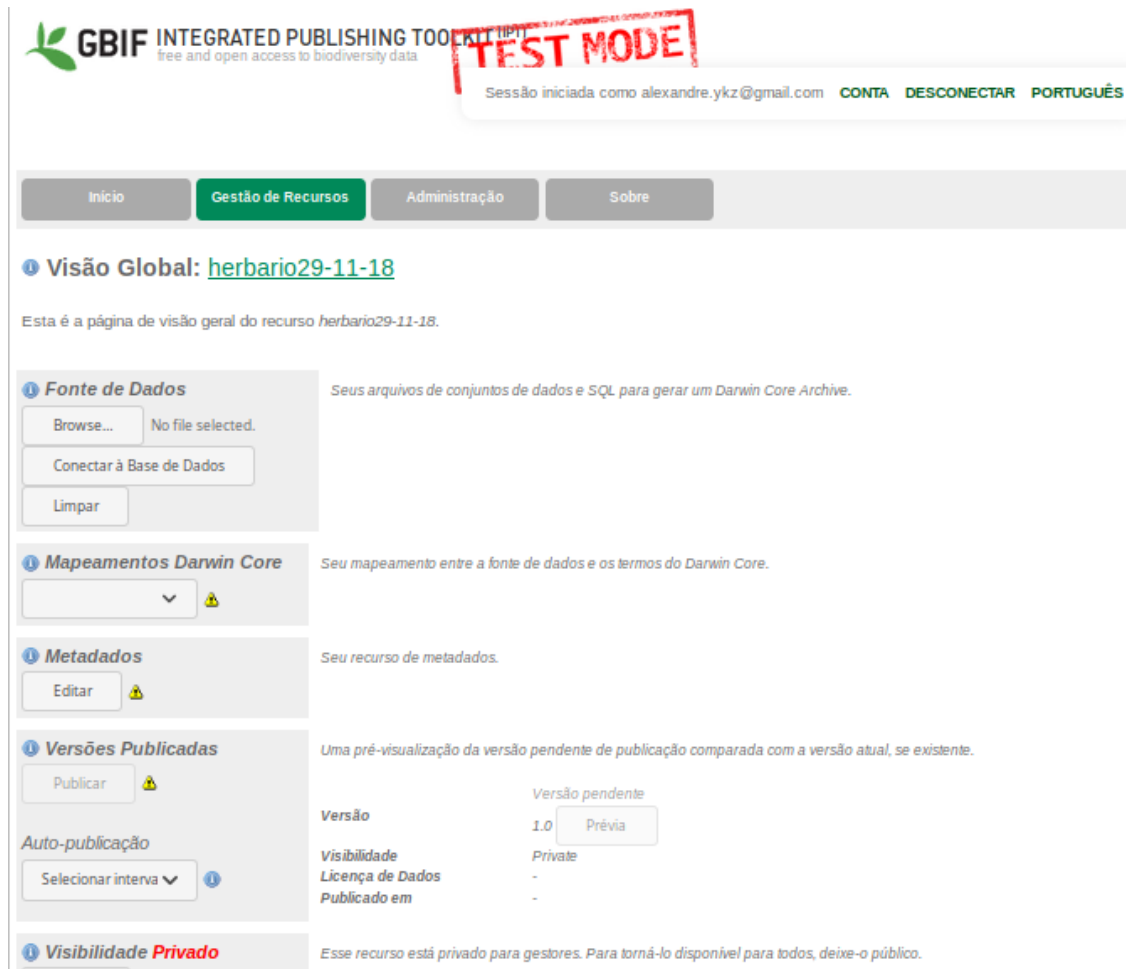


Figura 2.4. Interface de configuração do IPT, para conectá-lo a uma coleção.

2.4. SICol

O SICol, utilizado em 2018 por 17 coleções da rede speciesLink (SPECIESLINK, 2018c), foi desenvolvido pelo CRIA, que também é responsável pela plataforma speciesLink, por isso ambos utilizam o mesmo sistema de informação, para acesso de dados (CENTRO DE REFERÊNCIA EM INFORMAÇÃO AMBIENTAL, 2018b). Esse sistema multiplataforma, ilustrado na Figura 2.5, é destinado ao gerenciamento de acervos de micro-organismos, como fungos filamentosos, culturas de arqueias, bactérias, leveduras e protozoários. O SICol pode ser executado nos sistemas operacionais Windows, Linux e Mac OS X. O SICol é composto por oito módulos (CENTRO DE REFERÊNCIA EM INFORMAÇÃO AMBIENTAL, 2009):

- Linhagem: principal módulo do sistema, que armazena dados sobre linhagem ²⁷, táxon, tipo, coletor, data e local da coleta, coordenadas geográficas do local da coleta, gênero e espécie do micro-organismo;

²⁷ Linhagem: grupo de indivíduos que têm uma ascendência comum (VALOIS et al., 1999).

- Taxa: cadastra grupos taxonômicos referenciados por outros módulos, principalmente pelo Linhagem, em que o usuário precisará selecionar o nome do micro-organismo de uma lista cadastrada no Taxa;
- Instituições: cadastra instituições responsáveis pelas coleções;
- Pessoas: cadastra pessoas ligadas às instituições;
- Documentos: anexa documentos, como sobre meios de cultura e caracterização fisiológica, morfológica e molecular das amostras;
- Referências: cadastra referências bibliográficas;
- Preservação: documenta informações sobre métodos de preservação dos acervos, para garantir a rastreabilidade dentro da coleção e o controle do estoque;
- Distribuição: registra informações sobre linhagens vendidas, doadas e utilizadas em pesquisas, para controle do estoque.

The screenshot shows the SICol web interface. At the top, there is a navigation menu with items like 'microrganismos', 'noticias', 'a rede', 'provedores', 'como participar', 'indicadores', 'uso da rede', 'ferramentas', 'dicas de uso', and 'preferências'. The main header features the 'SICol.net' logo and the text 'Microorganismos' with a search bar and a button 'abrir formulário de busca'. Below the header, there are tabs for 'resumo', 'imagens', 'mapa', 'gráfico', and 'download'. The main content area displays search results, including a table of species and their counts. A detailed view of a specific record is shown, including taxonomic information and collection details.

espécie ^{1,2}	
Em branco	1311
...	1
<i>Abditodendrix rhomboidalis</i>	3
<i>Acremonium sp.</i>	5
<i>Adelosina brongnartii</i>	1
<i>Adelosina sp.</i>	1
<i>Adercotryma glomeratum</i>	1
<i>Alabaminella weddellensis</i>	1
<i>Alternaria alternata</i>	4
<i>Ammostata inepta</i>	1
<i>Ammobaculites exiguus</i>	1
<i>Ammonia beccarii parkinsoniana</i>	1
<i>Ammonia parkinsoniana</i>	3
<i>Ammonia rolshauseni</i>	1
<i>Ammonia tepida</i>	5
<i>Ammoscalaria pseudospiralis</i>	1
<i>Ammoscalaria tenuimargo</i>	1
<i>Ammotium dilatatum</i>	1
<i>Ammotium directum</i>	1
<i>Ammotium distinctum</i>	1
<i>Ammotium salsum</i>	4
<i>Amphicoryna hirsuta</i>	1
<i>Amphicoryna scalaris</i>	3

FUNGI ASCOMYCOTA SACCHAROMYCETES SACCHAROMYCETALES SACCHAROMYCETACEAE
Candida albicans. Det: Dieler Barbieri
 CMRP951 Coleta: Dieler Barbieri
 Loc: Curitiba, Paraná, Brasil
 Coord. munic.: [lat: -25.4277992248535 long: -49.2731018066406 err: ±25180 WGS84]
 Número de catálogo anterior: LMICRO517 / C18 Tipo de relacionamento: Hospedeiro: Humano
 Depósito: Dieler Barbieri.
 Hospedeiro ou Substrato: biofilme dental Isolado por: Dieler Barbieri Método de isolamento: Cultura/Saboraud/Cromoagar Método de cultivo: Sabouraud, 28°C
 Notas: Cód barras: 2014010017
 Tipo de registro: organismo vivo
 © Coleções de Culturas Microbiológicas da Rede Paranaense (CMRP)

FUNGI
CMRP3122
 Número de catálogo anterior: LMICRO2290 / F04
 Notas: Cód barras: 2018020143
 Tipo de registro: organismo vivo
 © Coleções de Culturas Microbiológicas da Rede Paranaense (CMRP)

PROTISTISTA GRANULORETICULOSA FORAMINIFERA ROSALINDAE
Aleoceranthina nitida (Williamson 1858) Det: Erica Riechi

Figura 2.5. Interface do SICol, com itens renderizados, em resposta à busca realizada.

2.5. speciesBase

Em 2018, o speciesBase era utilizado por 16 coleções (SPECIESLINK, 2018c) da rede speciesLink. Esse sistema para registro de informações sobre espécimes, baseado no padrão de dados Darwin Core, foi desenvolvido pelo CRIA para atender às instituições que desejam ingressar na rede speciesLink, mas não têm um sistema de gerenciamento (SPECIESLINK, 2018e). O speciesBase utiliza o Microsoft Access, é gratuito e pode ser livremente alterado e distribuído (CENTRO DE REFERÊNCIA EM INFORMAÇÃO AMBIENTAL, 2012). Apesar de sua documentação afirmar que é um *software open-source*, não foi encontrado

nenhum repositório.

O speciesBase possui apenas quatro funcionalidades, ilustradas na Figura 2.6 e descritas a seguir (CENTRO DE REFERÊNCIA EM INFORMAÇÃO AMBIENTAL, 2012):

- adiciona registros dos espécimes: taxonomia, local da coleta (país, estado, município e continente) e imagens e sons dos espécimes.
- cadastra informações sobre sexo, classe, espécie, ordem, reino, filo e gênero dos espécimes, em tabelas secundárias;
- cadastra, consulta e atualiza empréstimos;
- gera relatórios sobre empréstimos e devoluções, e sobre espécimes e espécies presentes no sistema.



Figura 2.6. Interface do speciesBase, com as quatro funcionalidades disponíveis.

2.6. Jabot

O Jabot era utilizado, em 2018, por 13 coleções participantes da rede speciesLink (SPECIESLINK, 2018c). Ele foi desenvolvido pelo Jardim Botânico do Rio de Janeiro (JBRJ), e seu código-fonte não está disponível de forma aberta. O Jabot permite armazenar e publicar dados e imagens *on-line*, e auxilia na curadoria de acervos de amostras botânicas desidratadas e vivas (SILVA et al., 2017).

O Jabot, cuja interface de consulta pode ser vista na Figura 2.7, é composto por oito módulos (SILVA et al., 2017):

1. Espécimes - dados e imagens: cataloga as coletas, para identificação e registro de seu número (tombo) no sistema. O sistema permite selecionar filtros e opções de retorno -

- por exemplo, a localização geográfica da coleta - e imprimir vários formatos de etiquetas, de acordo com o tipo de coleção;
2. Coleção viva: permite acesso aos dados da coleção de plantas vivas e às suas imagens, registradas e armazenadas no servidor;
 3. Curadoria e administração: auxilia digitadores, herborizadores, secretárias, conservadores, supervisores, curadores e subcuradores das coleções, na administração dos herbários. O administrador é responsável por criar e dar permissões. Esse módulo controla, também, remessas e visualização de relatórios;
 4. Publicação de dados: publica dados por meio da exportação de dados, por exemplo, nos formatos XLS, ODS (do inglês, *OpenDocument Spreadsheet*) e Linguagem Extensível de Marcação Genérica (XML, do inglês *Extensible Markup Language*), e também exporta para publicadores de dados nacionais e internacionais, como o Herbário Virtual Refflora e a GBIF;
 5. Qualidade de dados: analisa e acompanha a qualidade de dados referentes à taxonomia e localização geográfica das coletas; renderiza informações, como de gêneros duplicados em diferentes famílias e quantidade de amostras que não apresentam nome do coletor, ano da coleta e coordenadas (latitude, longitude e altitude) do local da coleta;
 6. Lista de espécies: auxilia na construção e gerenciamento de listas taxonômicas regionais;
 7. Espécies (taxonomia): de acesso exclusivo aos supervisores, gerencia os táxons, com a finalidade de manter e aprimorar a qualidade dos nomes adotados no Jabot;
 8. Coleções correlatas: possibilita consulta de diversas coleções, por exemplo, carpoteca, fototeca, exsicatas, sementes, cactário, orquidário e xiloteca.

Figura 2.7. Interface de consulta do Jabot, em que é possível selecionar coleção, família, gênero, espécie, localização e coletor.

2.7. Discussão

Neste capítulo, foram apresentados os seis *software* mais utilizados pelas coleções participantes da rede speciesLink. Alguns pontos em comum entre os sistemas podem ser destacados:

- são aplicações *web*, que permitem acesso remoto por vários usuários;
- gerenciam diversos tipos de coleções, como de plantas, animais e micro-organismos;
- realizam controle de usuários, por meio de concessão de privilégios;
- utilizam registros presentes nas aplicações, para gerar visualizações;
- geram arquivos no formato Darwin Core com dados presentes em seus bancos de dados;
- têm CRUD similares, como de instituições, usuários, espécimes e empréstimos.

Alguns *software* oferecem mais funcionalidades que o HCF-*Web*, que poderão ser implementadas em trabalhos futuros. Entretanto, nenhum desses seis sistemas possibilita a integração dos dados presentes nos herbários virtuais aos dos herbários físicos, como realizado neste trabalho.

HCF-*Web*

Até o ano de 2017, o sistema do HCF, do campus de Campo Mourão da UTFPR, era o Herbarium 1.0 (CARNEIRO, 2006). A ferramenta de programação Delphi (EMBARCADERO, 2018a) e o gerenciador de Banco de Dados Interbase 6 (EMBARCADERO, 2018b) foram utilizados para o seu desenvolvimento.

Sangali (2019) propôs, recentemente, a reformulação de alguns aspectos do Herbarium, para melhorar a eficiência, como do controle dos usuários, da disponibilização do sistema para alunos e servidores, do controle das modificações dos dados e, ainda, para tornar a sua interface gráfica mais intuitiva. Nessa nova versão - de 2017 - denominada HCF-*Web*, necessidades surgidas ao longo do tempo fizeram com que a aplicação se tornasse *web* e algumas funcionalidades fossem adicionadas, por exemplo, dados foram convertidos em formato Darwin Core, para permitir o seu envio para o SiBBr.

O HCF-*Web* é composto por duas partes: o *front end*, que contém os componentes das interfaces gráficas para a interação com o usuário; e o *back end*, responsável pela regra de negócio do HCF e pelo processamento dos dados oriundos do *front end* (SANGALI, 2019). O *JavaScript* (JS) (FREEMAN; ROBSON, 2014) foi a linguagem utilizada para o desenvolvimento do *back end* e do *front end*.

As tecnologias que fundamentam o HCF-*Web* são apresentadas na Seção 3.1; as estruturas dos diretórios da aplicação, na Seção 3.2; as funcionalidades propostas para esse *software*, na Seção 3.3; e uma breve discussão, na Seção 3.4.

3.1. Tecnologias

A *World Wide Web* - ou *web* - é uma estrutura arquitetônica que possibilita o acesso a documentos de diversos assuntos que estão presentes em milhões de máquinas disponíveis na Internet, e que se tornou popular por apresentar uma interface gráfica intuitiva, até para

novatos (TANENBAUM; WETHERALL, 2013).

A *web* foi desenvolvida em 1989, pela Organização Europeia para a Pesquisa Nuclear (CERN, do inglês *European Center for Nuclear Research*), com o objetivo de apoiar a interação de cientistas de diferentes países em experimentos de física de partículas. Essa proposta inicial foi elaborada pelo físico Tim Berners-Lee¹, e consistia em uma rede de documentos interligados. O protótipo inicial, apresentado ao público na Conferência Hypertext, em 1991, motivou o pesquisador Marc Andreessen, da Universidade de Illinois, a desenvolver, em 1993, o Mosaic, que foi o primeiro navegador gráfico². Um ano depois, em 1994, Marc Andreessen fundou a Netscape Communications Corporation, com o objetivo de desenvolver *software* para a *web*. Nos anos seguintes, os navegadores Netscape Navigator e Internet Explorer, da Microsoft, disputaram o domínio do novo mercado (TANENBAUM; WETHERALL, 2013).

Como dito anteriormente, a *web* consiste em uma coleção mundial de documentos, denominadas páginas. O conteúdo de uma página pode ser um documento, um vídeo ou um programa, que resulta em uma interface gráfica, com a qual o usuário pode interagir. Uma página é visualizada com o auxílio de um navegador, ou seja, de um *software* que requisita a página desejada para um ou mais servidores, e que renderiza a resposta na tela do computador (TANENBAUM; WETHERALL, 2013).

O Protocolo de Transferência de Hipertexto (HTTP, do inglês *HyperText Transfer Protocol*), que utiliza um modelo em que o cliente envia uma mensagem de solicitação, e o servidor retorna uma mensagem de resposta, é utilizado para buscar páginas *web* (TANENBAUM; WETHERALL, 2013). O HTTP é especificado em Pedidos de Comentário (RFC, do inglês *Request For Comments*), que determina o formato das mensagens trocadas entre clientes e servidores (FIELDING et al., 1999). Nas mensagens de requisições, por exemplo, é preciso indicar o método solicitado, e os mais empregados são o GET e o POST, responsáveis, respectivamente, pela solicitação ao servidor de uma página e pelo envio de um formulário. Nas mensagens de resposta, o servidor envia um código de status que informa se uma solicitação foi ou não atendida. Em caso de falha, o código, por exemplo, 404, indica que a página não foi encontrada (TANENBAUM; WETHERALL, 2013).

As páginas *web* podem ser estáticas ou dinâmicas. Uma página estática apresenta o mesmo documento, sempre que for buscada e exibida. A página dinâmica é gerada sob demanda, por um programa, ou contém um programa, e pode se apresentar de diferentes maneiras, cada vez que for exibida. As páginas dinâmicas são utilizadas em aplicações e serviços, como comércio eletrônico, leitura e envio de e-mail, pesquisa em catálogos de

¹ Em 2016, Tim Berners-Lee recebeu o prêmio Turing (ASSOCIATION FOR COMPUTING MACHINERY, 2016).

² Em 1993, o Mosaic ganhou o prêmio de Produto do Ano, da revista InfoWorld, e, em 1994, Tecnologia do Ano, da revista IndustryWeek (NACIONAL CENTER FOR SUPERCOMPUTING APPLICATIONS, 2018).

biblioteca e exibição de mapas (TANENBAUM; WETHERALL, 2013).

Nas páginas *web*, a geração de conteúdo pode ocorrer do lado do servidor, como no caso de uso de um formulário, em que os dados nele presentes são empregados como entrada de um programa no servidor. A geração de conteúdo pode ocorrer, também, do lado do cliente, o que permite, por exemplo, respostas a movimentos do mouse ou a interação direta com os usuários, por meio de *scripts* embutidos nas páginas escritas em Linguagem de Marcação de Hipertexto (HTML, do inglês *HyperText Markup Language*). Em relação ao lado do cliente, a linguagem mais popular é o JS. Por ser uma linguagem de alto nível, o JS proporciona facilidades, como manipulação de caixas de diálogo, caixas de entrada de texto e armazenamento de cadeia de caracteres resultantes de uma variável. Dessa forma, o JS é uma excelente opção para projetar páginas *web* interativas (TANENBAUM; WETHERALL, 2013).

Atualmente, com o desenvolvimento da plataforma Node.js³, o JS também tem se tornado popular na implementação de aplicações do lado do servidor (HERMAN, 2013). O Node.js permite executar o JS nos servidores *web*, isto é, fora dos navegadores. Isso ocorre porque o JS é uma linguagem orientada a eventos, e o Node.js aproveita-se disso para permitir que sejam feitos servidores altamente escalonáveis, fáceis e seguros, por meio de uma arquitetura denominada *loop* de eventos⁴, que reduz a complexidade para o desenvolvedor da aplicação (HUGHES-CROUCHER; WILSON, 2012). Além disso, o Node.js pode ser utilizado na espera de resultados de banco de dados, *feedback* de serviços *web* de terceiros e solicitação de conexão de terceiros (WILSON, 2013).

O Node.js é assíncrono (GONZALEZ, 2016), e tem como características principais: biblioteca padrão, sistema de módulos e gerenciador de pacotes do Node (NPM, do inglês *Node Package Manager*). A biblioteca padrão, recurso mais forte e poderoso do Node.js, é composta por um conjunto de bibliotecas binárias e módulos principais (YOUNG; HARTER, 2014). Um módulo é um código que tem apenas um único propósito, e o Node.js utiliza-se disso para melhorar a organização dos códigos (SATHEESH et al., 2015). Os módulos são referenciados pelo caminho do arquivo ou, ainda, pelo seu nome. O NPM é o gerenciador de módulos de terceiros (TEIXEIRA, 2012). Entre os módulos do Node.js, podem ser citados: HTTP, responsável pelo servidor HTTP; Localizador de Recurso Uniforme (URL, do inglês *Uniform Resource Locator*), para análise de informação de Identificador Uniforme de Recurso (URI, do inglês *Uniform Resource Identifier*); e FS, que trabalha com o sistema de arquivos (MARDAN, 2015).

O NPM pode ser utilizado para a instalação do react (AMLER; SONPATKI, 2016), que consiste em um projeto de código aberto⁵, criado pelo Facebook, que oferece uma nova maneira de construir uma interface em JS, apenas se preocupando com a renderização da

³ <https://nodejs.org/en/>

⁴ Essa arquitetura permite que o Node.js execute operações de entrada/saída sem bloqueio (NODE.JS, 2018).

⁵ <https://github.com/facebook/react/>

Interface do Usuário (UI, do inglês *User Interface*). No react é utilizada uma linguagem declarativa, isto é, que descreve, por meio de declarações, como os componentes devem se parecer e se comportar (ANTONIO, 2015).

Os componentes UI, criados com o react, são altamente independentes e têm funcionalidades bem específicas, como (i) selecionar data e endereço, e (ii) comunicar-se com o servidor, com o preenchimento automático, por meio de busca em uma lista no servidor (MARDAN, 2017).

O react contribuiu para a resolução de alguns problemas, como a criação e o gerenciamento de UI complexas da *web*, para aplicativos *front end*. No caso de um grande aplicativo da *web*, como o Facebook, é árduo o trabalho de gerenciamento das mudanças de visualizações em resposta às alterações dos dados. O react contribui, assim, para a construção de grandes aplicações com dados que são alterados com o tempo (MARDAN, 2017).

Vantagens do react podem ser citadas: aplicações mais simples, UI rápidas e menos códigos para escrever, pois a sua comunidade oferece uma variedade de componentes e bibliotecas para desenvolvimento. As aplicações ficam mais simples, porque sua arquitetura é do tipo Arquitetura Baseada em Componentes⁶ (CBA, do inglês *Component-Based Architecture*), com o JS puro e poderosas abstrações do Modelo de Objeto de Documento⁷ (DOM, do inglês *Document Object Model*), sendo mais amigável ao desenvolvedor. As UI ficam mais rápidas por causa do desempenho do DOM virtual e do seu algoritmo *smart-reconciliation*⁸ (MARDAN, 2017).

3.2. Estruturas de diretórios do HCF- Web

O desenvolvimento de um projeto, por meio da tecnologia Node.js e de seu gerenciador de pacotes, requer a execução do comando `npm init -y` dentro de um diretório. A saída desse comando será a criação de um arquivo `package.json`, formado pelos seguintes conteúdos: nome, versão e descrição do projeto; comandos para inicializar, reinicializar ou parar o projeto; e nomes e versões dos módulos utilizados.

Módulos podem ser adicionados ao projeto por meio do comando `npm install x`, em que `x` corresponde ao nome do módulo que o desenvolvedor deseja instalar. A saída desse comando será um novo diretório denominado `node_modules`, composto por todos os módulos instalados.

⁶ É uma arquitetura em que se dividem interesses, baixo acoplamento e reutilização de código (MARDAN, 2017).

⁷ Esse modelo é uma Interface de Programação de Aplicativos (API, do inglês *Application Programming Interface*), que fornece uma coleção de objetos que podem ser utilizados para interagir com o navegador, modificando o documento (BANKS; PORCELLO, 2017).

⁸ Também conhecido como algoritmo *diffing*, que faz com que as atualizações de componentes sejam previsíveis, rápidas para aplicativos de alto desempenho (REACT, 2018).

A proposta feita por Sangali (2019), para a estrutura do *back end* do HCF-Web, pode ser vista na Figura 3.1, e é composta pelos seguintes diretórios:

- **docs**: dedica-se às informações do projeto;
- **scripts**: dispõe de código que se comunica com outro serviço auxiliar;
- **src**: contém o código de execução do *back end*;
- **test**: possui os códigos para verificar se as funcionalidades estão sendo desempenhadas da maneira esperada.

Alguns dos `node_modules` do *back end* do HCF-Web são:

- **http**⁹: para criação do servidor;
- **mysql2**¹⁰: para conexão com o banco de dados MySQL (MYSQL, 2018), definição de entidades com seus atributos e respectivos tipos, e estabelecimento das chaves primárias e estrangeiras do banco de dados;
- **jsonwebtoken**¹¹: para gerar *tokens* para autenticação do sistema;
- **bcrypt-nodejs**¹²: para verificar senha;
- **sequelize**¹³: consiste em uma técnica de Mapeamento Objeto Relacional (ORM, do inglês *Object Relational Mapping*). Essa técnica possibilita a conexão de um modelo de um objeto a um banco de dados relacional, utilizando metadados para descrever o objeto e os dados (MEHTA, 2008). No caso do HCF-Web, o **sequelize** permite ao *back end* comunicar-se com o banco de dados MySQL, e realizar certas ações: adicionar, consultar e apagar registros.

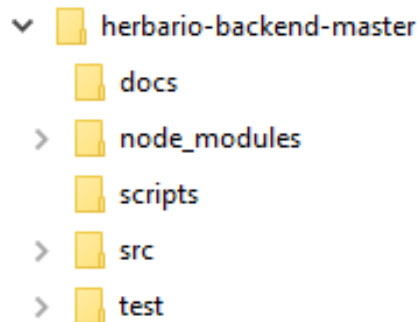


Figura 3.1. Estrutura de diretórios do *back end* - do HCF-Web - formado por: **docs** (informações do projeto); **node_modules** (módulos utilizados na aplicação); **scripts** (código de comunicação com serviço auxiliar); **src** (código de execução do *back end*); e **test** (códigos que verificam se resultados esperados estão sendo gerados).

⁹ <https://nodejs.org/api/http.html>

¹⁰ <https://github.com/sidorares/node-mysql2/tree/master/documentation>

¹¹ <http://self-issued.info/docs/draft-ietf-oauth-json-web-token.html>

¹² <https://www.npmjs.com/package/bcrypt-nodejs>

¹³ <http://docs.sequelizejs.com/>

Sangali (2019) propôs uma estrutura de *front end* para o HCF-*Web*, que pode vista na Figura 3.2, e que é composta pelos seguintes diretórios:

- **src**: contém o código de execução do *front end*;
- **public**: contém arquivos estáticos que são requisitados ao servidor pelo cliente;
- **node_modules**: formado, por exemplo, pelo **react**¹⁴, que permite criar componentes que podem ser utilizados em diferentes interfaces; pelo **moment**¹⁵, feito para a formatação dos dias, horas e telefones; pelo **axios**¹⁶, que permite realizar requisições HTTP, como GET, POST, PUT, DELETE; e pelo **react-router-dom**¹⁷, que habilita a navegação entre as interfaces.

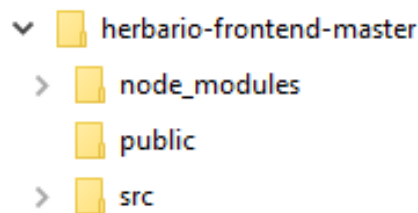


Figura 3.2. Estrutura de diretórios do *front end* - do HCF-*Web* - formado por: **node_modules** (módulos utilizados no *front end*); **public** (arquivos estáticos presentes no servidores retornados aos clientes); e **src** (código de execução do *front end*).

3.3. Funcionalidades do HCF-*Web*

Para que as funcionalidades implementadas no HCF-*Web* operem perfeitamente, é necessário um banco de dados. No caso dessa aplicação, o banco tem 31 relações, como ilustrado na Figura 3.3. Considerando o porte desse banco de dados, não é possível exibir todos os atributos das relações, porém elas estão disponíveis no Apêndice A. É possível observar que a relação **tombo** se destaca das demais, por causa da quantidade de relacionamentos que apresenta, os quais lhe conferem grande importância no banco de dados.

Um **tombo** de um herbário é o número de registro de uma amostra seca de uma planta ou de um fungo, que pertence a uma família, subfamília, gênero, espécie, subespécie etc., e que faz parte de uma determinada fase sucessional de uma formação florestal. Essa amostra, que pode ser permutada ou emprestada para outro herbário, foi coletada em uma localidade com determinada coordenada geográfica, tipo de solo, relevo e vegetação (PEIXOTO; MAIA, 2013).

¹⁴ <https://reactjs.org/docs/hello-world.html>

¹⁵ <https://momentjs.com/docs/>

¹⁶ <https://github.com/axios/axios>

¹⁷ <https://reacttraining.com/react-router/web/guides/quick-start>

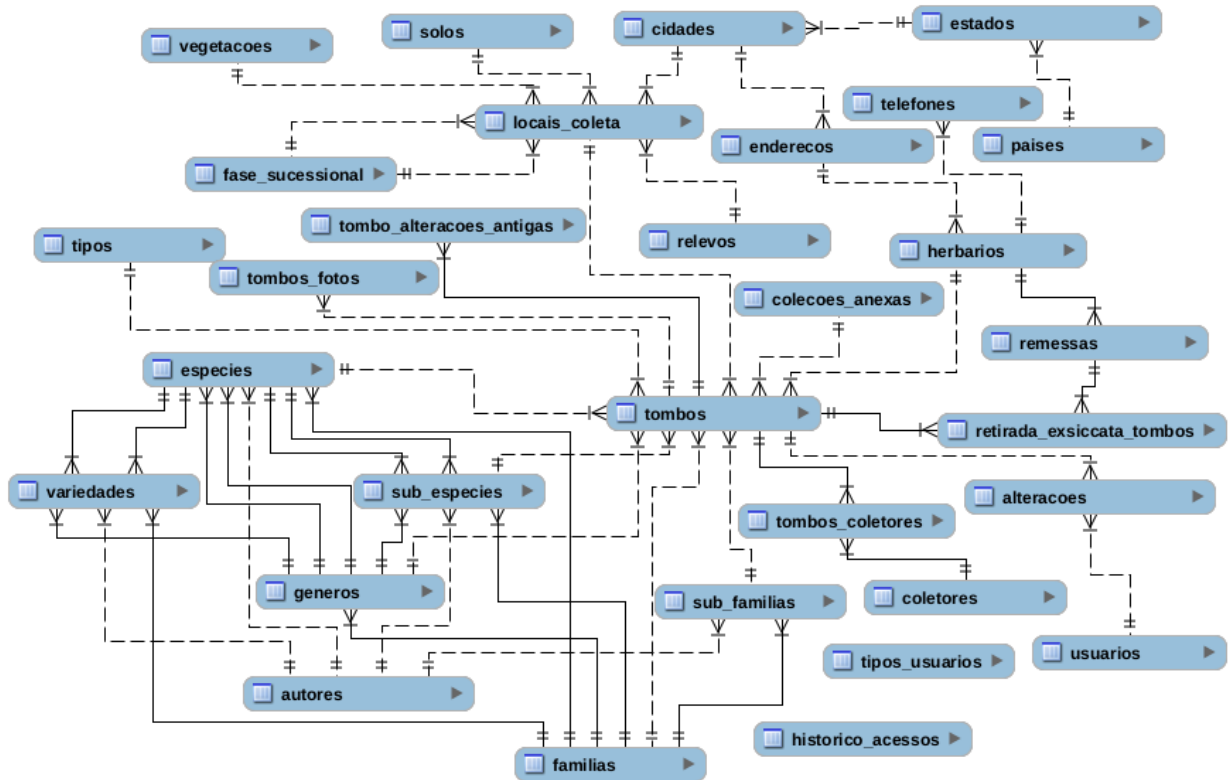


Figura 3.3. Diagrama ER, do HCF-Web. Cada tombo é o registro de uma amostra da coleção, e a relação que se destaca é a `tombos`, por causa da quantidade de relacionamentos.

O HCF-Web apresenta dez funcionalidades, mas algumas não estão disponíveis para todos os usuários, porque existe uma hierarquia, que pode ser visualizada na Figura 3.4. Somente o curador, que ocupa o nível mais alto dessa estrutura, tem acesso a todas as funcionalidades do herbário e, conseqüentemente, a todas as relações do banco de dados. O curador é o único que tem acesso às seguintes funcionalidades: adicionar, buscar, visualizar, editar e apagar usuários que utilizam a aplicação; apagar um tombo; aprovar uma alteração; e aprovar a adição de uma nova exsiccata, feita por um operador. Um exemplo de usuário cadastrado pode ser visto na Figura 3.5.

No nível hierárquico intermediário, abaixo do curador, encontram-se os operadores, que têm acesso à maioria das funcionalidades, com exceção das funções relacionadas aos usuários e aprovações. O nível hierárquico inferior é composto pelos visitantes, que somente têm acesso à listagem dos tombos presentes no herbário, por isso, só podem realizar consultas à relação `tombos`.



Figura 3.4. Hierarquia dos usuários no HCF-Web. O curador, que está no topo da pirâmide, tem acesso a todas as funcionalidades. Nos níveis abaixo, de acordo com a prioridade, operadores e visitantes possuem acesso mais restrito.

Buscar usuário

Nome: Email: Tipo:

Telefone:

Nome	Tipo	Email	Telefone	Data Criação	Ação
Marcelo Caxambu	curador	mgcaxambu@yahoo.com.br		02/09/2018 17:58	

Figura 3.5. Interface de listagem de usuários do HCF-Web, de acesso exclusivo ao curador. Além de visualizar, é possível adicionar, buscar, editar e apagar usuários.

Todas as funcionalidades do HCF-Web podem ser acessadas por meio de um menu lateral, à esquerda da tela. Para a adição de uma nova amostra ao sistema, é preciso utilizar a função cadastro de tombo. O cadastramento pode ser feito pelo curador e pelo operador, mas quando este o realiza, o registro somente irá para o banco de dados após a sua aprovação, por aquele. Enquanto esse novo dado não for aprovado pelo curador, o operador que o adicionou poderá apagá-lo. Nesse período em que estiver aguardando aprovação, o dado ficará em um atributo da relação `alteracoes`, no formato Notação de Objetos JavaScript (JSON, do inglês

JavaScript Object Notation); e depois de aprovado, tornar-se-á um registro na relação **tombos**. Quando o curador adiciona uma nova amostra, o registro é incluído, diretamente, ao banco de dados.

Todos os usuários podem visualizar **tombos** nas páginas. Eles podem, também, visualizar detalhes do registro e filtrar **tombos** específicos, selecionando número de tomo, tipo, situação da amostra (regular, doada, emprestada ou permutada), nome científico ou nome popular. O curador e os operadores podem exportar dados dos **tombos**, selecionando campos por meio do *check box*. A única pessoa que pode excluir **tombos** é o curador.

Em relação à funcionalidade taxonomia, as amostras do acervo do HCF- *Web* podem ser cadastradas nos táxons família, subfamília, gênero, espécie, subespécie e variedade, nas relações, respectivamente, **familia**, **subfamilia**, **genero**, **especie**, **subespecie** e **variedade**. As funcionalidades de cada uma dessas relações são semelhantes às de **tombos**: além de cadastrar, é possível visualizar táxons e filtrar buscas, indicando família, subfamília, gênero, espécie, subespécie ou variedade. Entretanto, não é possível exportar dados. Em relação ao cadastramento dos táxons, todos os campos são obrigatórios, e são preenchidos quando o curador ou operadores adicionam novos **tombos**.

Essas utilidades, disponíveis em **tombos** e taxonomia, também são oferecidas na opção **remessas** para outros herbários, da relação **remessas**. No cadastramento de uma remessa, que pode ser visto na Figura 3.6, os campos doador e receptor são obrigatórios. Além de cadastrar, também é possível: (i) visualizar remessas; (ii) filtrar remessas, selecionando seu número, número do tomo ou nome do herbário; (iii) visualizar se uma remessa foi um empréstimo, doação ou permuta; (iv) realizar edições; (v) apagar remessas; (vi) e exportar dados.

Para cadastrar uma remessa é preciso indicar o herbário de origem e o de destino, por isso, é imprescindível ter uma lista dos herbários com os quais o HCF mantém contato. A aplicação proporciona uma listagem de herbários, a partir da relação **herbarios**. É possível visualizar, filtrar, ordenar e cadastrar novos herbários. Edições, remoções e exportações de registros dos herbários presentes no banco de dados também podem ser realizados pelo curador. Para cadastrar um herbário, grande parte dos campos são obrigatórios, com exceção de sigla, logotipo e observações.

The image shows a web interface for recording a shipment. On the left is a dark sidebar with the HCF logo and navigation icons. The main content area is titled 'Cadastrar Remessa'. It contains several input fields: 'Doador' (a dropdown menu with 'HCF - Herbário da Universidade Tecnológica Federal do Paraná Campus Campo Mourão' selected), 'Receptor' (a dropdown menu with 'HUEM - Herbário da Universidade Estadual de Maringá' selected), 'Data de envio' (a date picker showing '27-06-2019'), and 'Observação' (a text area with the text 'Exsicatas doadas para o Herbário da Universidade Estadual de Maringá'). Below these fields is a button labeled '+ Adicionar tombo'. At the bottom, there is a table header with four columns: 'Tombo', 'Tipo', 'Data Vencimento', and 'Ação'.

Figura 3.6. Cadastro de remessa, realizado para permuta, troca ou doação de exsicatas.

Na relação **alteracoes**, quando o curador ou o operador modificam alguma informação, é gerada uma pendência no formato JSON. A efetivação de uma alteração realizada pelo operador requer a aprovação do curador. Caso o operador desista da alteração, poderá removê-la, se o status for pendente.

A lista renderizada na funcionalidade pendências é elaborada a partir dos registros presentes na relação **alteracoes**. Nessa lista são renderizadas informações, como números dos tombos, usuários, datas de criação, status (aprovado, reprovado ou pendente) e registros de observações. Se o curador desejar exibir detalhes de um item da lista pendentes, poderá clicar nele e visualizar o valor presente na relação **tombos** e o valor sugerido, e decidir pela aprovação ou pela reprovação justificada. O operador pode visualizar a lista e o comentário realizado pelo curador, mas não pode comentar, aprovar ou reprovar a pendência. É possível buscar itens pelos nomes de usuários e status.

O HCF-Web pode gerar relatórios sobre: (i) datas das coletas das amostras; (ii) nomes das espécies, famílias e gêneros das exsicatas e respectivas quantidades, presentes no banco de dados. Os relatórios são produzidos a partir da exportação dos conteúdos das relações **especies**, **familia**, **genero** e **tombos**.

Cada exsicata do acervo do HCF possui uma etiqueta impressa, colada no canto inferior direito da cartolina onde a amostra foi afixada, como pode ser visto na Figura 3.7. As informações contidas nas etiquetas são obtidas a partir da relação **tombos**. Na funcionalidade relatórios, não é possível realizar cadastro, edição ou remoção, mas somente buscar tombos, selecionando intervalo da data de coleta, gênero, espécie ou família. O *checkbox* permite

selecionar campos cujos dados podem ser exportados: código da família, data da coleta, família, gênero, espécie e quantidade.



Figura 3.7. Exemplo de exsicata do HCF, afixada em cartolina e acompanhada de etiqueta colada no canto inferior direito.

No HCF-Web, arquivos no padrão Darwin Core são obtidos ao serem realizadas operações de select em certas relações, como tombos, coletores, herbarios, cidade, estado, pais, familia, genero e especie. O arquivo gerado é semelhante a um arquivo do tipo Valores Separados por Vírgula (CSV, do inglês *Comma-Separated Values*). Todas as funcionalidades descritas nesta seção são representadas no diagrama de caso de uso, da Figura 3.8.

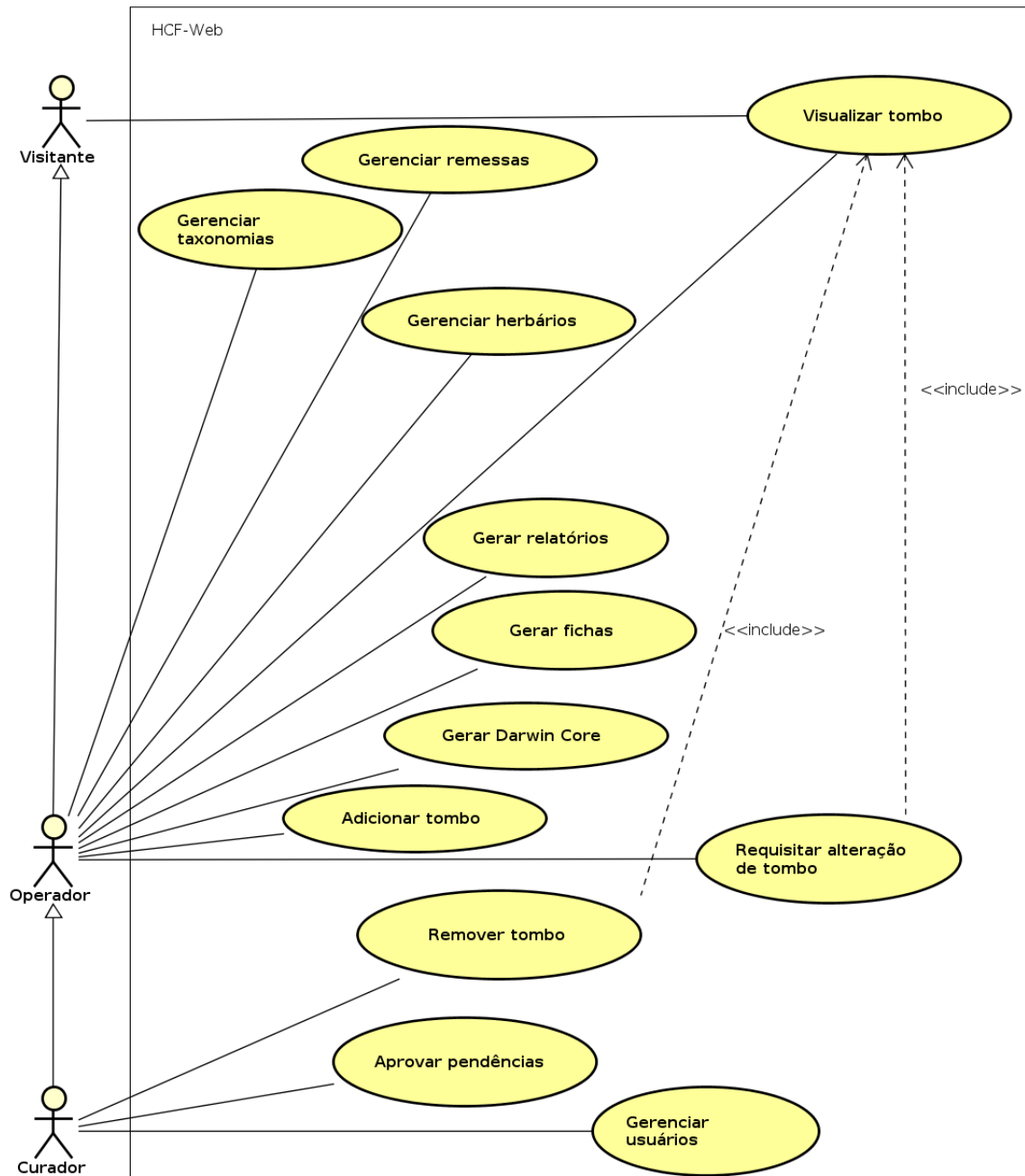


Figura 3.8. Diagrama de caso de uso do HCF-Web, em que o ator Curador tem acesso a todas as funcionalidades; o Operador utiliza algumas funções, com exceção de criação, remoção, atualização e visualização de usuários e aprovação de pendências; e o Visitante somente visualiza os tombos.

3.4. Discussão

Neste capítulo, foram descritas as mudanças ocorridas, ao longo do tempo, com o sistema do HCF, de *desktop* para uma versão *web*, que permite multiusuários. Além disso, foram apresentadas as tecnologias que fundamentam o HCF-*Web*, assim como a sua estrutura de diretórios (*back end* e *front end*) e as suas funcionalidades, entre elas, a que permite a conversão de seus dados em formato Darwin Core, para que o seu acervo possa ser disponibilizado em herbários virtuais.

No próximo capítulo, são apresentados o Herbário Virtual Re flora e o INCT-HVFF, onde, atualmente, é possível acessar a coleção de plantas do HCF. Ainda nesse capítulo, são relatados aspectos históricos da criação desses dois herbários virtuais, e descritos os processos de submissão dos dados do HCF para eles.

Herbários virtuais

Um herbário virtual é uma fonte qualificada de pesquisa e de informações, composta por dados e imagens de amostras depositadas em um herbário físico, disponibilizados na Internet e acessíveis a todos. Atualmente, um herbário virtual, além de individualmente disponibilizar seus dados *on-line*, também pode fazê-lo por meio de redes colaborativas (MAIA et al., 2017), como a do Herbário Virtual Refflora e a do INCT-HVFF.

Esses dois herbários virtuais, juntamente com outros projetos da área de biodiversidade que começaram a ser implantados no início do século XXI, foram fruto da inclusão, pelo governo federal, dessa temática nas prioridades de pesquisas a serem realizadas no país. No final do século XX, a biodiversidade foi inserida na agenda política internacional (MAIA et al., 2017). Em 1992, o Brasil sediou a Conferência das Nações Unidas sobre o Meio Ambiente e o Desenvolvimento, em que foi produzida a Convenção sobre Diversidade Biológica, um importante tratado internacional que tem como objetivos: (i) conservar a diversidade biológica; (ii) usar, de forma sustentável, a biodiversidade; (iii) e repartir, de forma justa e equitativa, os benefícios provenientes da utilização dos recursos genéticos (PEIXOTO et al., 2006).

A partir desse tratado, o governo federal, em parceria com as sociedades científicas de botânica, zoologia e microbiologia, começaram a trabalhar em um plano estratégico para conhecer, conservar e usar, de forma sustentável, a biodiversidade. Esse foi um grande desafio, pois o Brasil contava com um grupo muito pequeno de taxonomistas, e a maioria das coleções biológicas ainda não estava estruturada para atender à demanda por sistemas integrados de informação sobre biodiversidade (PEIXOTO et al., 2006). Esse reconhecimento da importância da biodiversidade resultou em (MAIA et al., 2017):

- maiores recursos para pesquisas;
- estímulo à formação de recursos humanos e redes de pesquisa;
- estímulo ao desenvolvimento e utilização de novas ferramentas computacionais;
- estímulo ao desenvolvimento de projetos, por exemplo, de herbários virtuais.

O Herbário Virtual Reflora faz parte do Projeto “Plantas do Brasil: resgate histórico e herbário virtual para a conservação da flora brasileira – Reflora”, do CNPq/Ministério da Ciência, Tecnologia e Inovação (MCTI). Esse Herbário Virtual foi criado com o objetivo de repatriar imagens e informações textuais de plantas brasileiras, e disponibilizá-las, de forma acessível e aberta, além de promover a integração e disseminação desses dados da flora, que até então somente estavam disponíveis nas instituições estrangeiras onde as amostras estavam depositadas (MAIA, 2010).

O INCT-HVFF faz parte do Programa Institutos Nacionais de Ciência e Tecnologia, do CNPq/MCTI, que reúne os mais importantes grupos de pesquisa que atuam em áreas fundamentais para o desenvolvimento sustentável do país, com o objetivo de impulsionar a pesquisa científica e tecnológica, tornando-a competitiva internacionalmente. Na primeira fase desse Programa, foram criados 122 institutos, entre eles o INCT-HVFF, que iniciou suas atividades em 2009, é coordenado pela Universidade Federal de Pernambuco (UFPE), e tem como meta principal reunir dados de plantas e fungos de herbários brasileiros na plataforma speciesLink.

Em 2010, o INCT-HVFF foi convidado pelo CNPq a participar do Projeto Reflora (PEIXOTO; MAIA, 2013), e essa junção INCT-HVFF/Reflora tem como objetivo principal aumentar a base informacional sobre a flora e fungos do Brasil (INCT - HERBÁRIO VIRTUAL DA FLORA E DOS FUNGOS, 2018). A partir dessa inserção, o INCT-HVFF passou a ser responsável (MAIA, 2010):

- pela integração dos dados textuais repatriados sobre a flora brasileira aos acervos dos herbários parceiros;
- pela disseminação digital dos dados repatriados, por meio de sua disponibilização para acesso livre e aberto na Internet;
- pelo desenvolvimento de ferramentas e plataformas informatizadas, indispensáveis para a instalação de salas de visitas públicas em vários pontos do país.

Nas Seções 4.1 e 4.2 são descritas as funcionalidades, respectivamente, do Herbário Virtual Reflora e do INCT-HVFF, que foram utilizadas para a obtenção de informações referentes às exsicatas do HCF; e na Seção 4.3, é realizada uma breve discussão.

4.1. Herbário Virtual Reflora

Durante o período colonial, entre os séculos XVIII e XIX, missões estrangeiras coletaram uma imensa quantidade de amostras de plantas brasileiras, e as enviaram aos herbários europeus. A necessidade de resgatar e disponibilizar ao público brasileiro esses dados botânicos históricos, repatriados de outros países, levou o CNPq a iniciar, em 2010, o Projeto Reflora, que desenvolveu um Herbário Virtual, administrado pelo Instituto de Pesquisas Jardim Botânico do Rio de Janeiro (REFLORA, 2018a).

Os dois primeiros parceiros do Herbário Virtual Reflora foram o Kew Gardens (Inglaterra) e o *Muséum National d'Histoire Naturelle* (França), que permitiram o acesso às informações textuais e a digitalização dos materiais referentes à flora brasileira. A partir do ano de 2014, o Herbário Virtual Reflora, com o apoio do SiBBR, começou a disponibilizar dados de coleções da flora brasileira depositados em outros oito herbários europeus e americanos e, também, em herbários brasileiros, entre eles, no HCF. O acervo atual do Reflora é composto por 3.519.680 imagens de espécimes (REFLORA, 2018a).

O Herbário Virtual Reflora foi construído de forma a permitir que os usuários possam acessar imagens em alta resolução, em uma plataforma *on-line* que oferece várias funcionalidades, entre elas, **a consulta, reidentificação e tipificação das amostras**. O sistema proporciona, também, aos curadores das instituições parceiras, acesso a relatórios para realização de atualização em seus herbários (REFLORA, 2018a).

O Reflora publica, atualmente, imagens e dados de 59 herbários parceiros do Brasil (REFLORA, 2018a). Para que um especialista cadastrado nesse herbário virtual possa sugerir reidentificação de exsicatas, é necessário que o herbário associado disponibilize, no Reflora, os dados de seu acervo em formato Darwin Core, conforme explicado na Seção 3.3.

O Herbário Virtual Reflora oferece uma API¹, do tipo Transferência de Estado Representacional (REST, do inglês *Representational State Transfer*)². Essa API permite realizar requisições por meio do método HTTP GET, para obter diversas informações - por exemplo, sobre famílias, gêneros e espécies de plantas e fungos - disponíveis nos herbários parceiros. Para a implementação da extensão para o HCF-*Web*, essa API foi utilizada para retornar informações das amostras, por meio da URI:

`http://servicos.jbrj.gov.br/v2/herbarium/x,`

em que **x** é o número do catálogo da exsicata, e equivale, no HCF-*Web*, ao registro no atributo `codigo_barra`, da relação `tombos_fotos`.

Quando o usuário do Herbário Virtual Reflora realiza uma busca, a API retorna arquivos JSON com três tipos de conteúdos. Na situação em que ocorre uma falha no servidor, o conteúdo retornado é o ilustrado na Figura 4.1.

```
{"erro": "500", "message": "Oops, something's gone wrong in server!"}
```

Figura 4.1. Exemplo de arquivo JSON retornado do Reflora, quando ocorre um erro inesperado no servidor.

¹ <http://servicos.jbrj.gov.br/v2/flora/>

² O estilo arquitetural REST auxilia na criação e organização de sistemas distribuídos (DOGLIO, 2015).

Quando a exsicata solicitada está disponível no banco de dados do herbário parceiro, mas não no do Herbário Virtual Reflora, a API retorna um arquivo JSON com o seguinte conteúdo, ilustrado na Figura 4.2.

```
{"success": true, "result": []}
```

Figura 4.2. Exemplo de arquivo JSON, retornado pela API quando a exsicata solicitada não está disponível no Herbário Virtual Reflora.

Se a exsicata solicitada está disponível nos bancos de dados do herbário associado e do Reflora, a API retorna um JSON do tipo exemplificado na Figura 4.3. Nesse caso, são retornadas as chaves (**success**) e (**result**), e dentro destas, são retornadas várias chaves semelhantes às de um arquivo Darwin Core. Como ilustrado na Figura 4.3, durante o processo de atualização dos dados do HCF-*Web* são retornadas as seguintes chaves do arquivo JSON do Reflora (SISTEMA DE INFORMAÇÃO SOBRE A BIODIVERSIDADE BRASILEIRA, 2018a):

- **occurrenceid:** identificador obrigatório e único (sem valores em duplicata), para o registro de ocorrências de espécies. Quando não há um identificador único, global e persistente, pode ser utilizada a forma: "urn:catalog:[institutionCode]:[collectionCode]:[CatalogNumber]";
- **rightsholder:** pessoa ou organização que tem direitos sobre o recurso;
- **bibliographiccitation:** indicação de como deve ser realizada uma referência bibliográfica de um registro utilizado;
- **modified:** data mais recente da modificação do recurso;
- **institutioncode:** nome da instituição que possui a custódia de um objeto ou de uma informação referente ao registro;
- **collectioncode:** nome da coleção ou conjunto de dados ao qual pertence o registro;
- **catalognumber:** identificador do registro dentro de uma coleção (número de tombo);
- **occurrenceremarks:** comentário ou nota sobre a ocorrência;
- **recordedby:** lista de pessoas ou organizações responsáveis pelo registro de uma ocorrência original (coletor da amostra);
- **recordnumber:** identificador dado a uma ocorrência, no momento em que foi registrada;
- **othercatalognumbers:** lista de números de catálogos anteriores, alternativos ou de outros identificadores utilizados para a mesma ocorrência;
- **eventdate:** data (ano/mês/dia) ou intervalo de tempo em que um evento ocorreu;
- **verbatimeventdate:** citação literal e original de como a data de um evento foi descrita;
- **fieldnotes:** anotações que indicam a existência de um evento ou de uma referência (bibliográfica ou URI), ou texto sobre anotações feitas em campo sobre um evento;

- **country**: país da localização da ocorrência;
- **countrycode**: código padrão (sigla) do país da localização da ocorrência;
- **stateprovince**: estado da localização da ocorrência;
- **municipality**: distrito da localização da ocorrência;
- **locality**: descrição específica do local da ocorrência;
- **minimumelevationinmeters**: limite inferior da amplitude da elevação, em metros, da localização da ocorrência;
- **maximumelevationinmeters**: limite superior da amplitude da elevação, em metros, da localização da ocorrência;
- **verbatimlatitude**: latitude, literal e original, da localização da ocorrência;
- **verbatimlongitude**: longitude, literal e original, da localização da ocorrência;
- **decimallatitude**: latitude geográfica, em graus decimais, da localização da ocorrência;
- **decimallongitude**: longitude geográfica, em graus decimais, da localização da ocorrência;
- **identifiedby**: lista de nomes de determinadores, grupos ou organizações responsáveis pela identificação do táxon a que pertence o espécime;
- **identificationverificationstatus**: indicador categórico que assinala a precisão com que a classificação taxonômica foi realizada;
- **dateidentified**: data (ano/mês/dia) em que o espécime foi identificado como sendo representante de um táxon;
- **identificationremarks**: anotações sobre a identificação do espécime;
- **identificationqualifier**: frase que expressa dúvida do determinador (especialista que identificou o material) sobre uma identificação;
- **typestatus**: lista de tipos nomenclaturais aplicados ao espécime;
- **scientificname**: nome científico completo do espécime e, quando conhecidas, autoria e informação da data em que a identificação foi realizada;
- **family**: nome da família do espécime;
- **genus**: nome do gênero do espécime;
- **specificepithet**: epíteto específico, ou seja, nome da espécie do espécime;
- **infraspecificepithet**: epíteto infraespecífico, ou seja, nome da subespécie, variedade, forma etc. do espécime, que correspondem aos níveis taxonômicos inferiores. Para saber se o epíteto infraespecífico do espécime é de uma subespécie, variedade, forma etc., basta consultar o valor do campo **taxonrank**;
- **taxonrank**: nível taxonômico mais específico até onde foi identificado o espécime coletado, por exemplo, gênero, espécie, subespécie, variedade, forma etc. ³;
- **scientificnameauthorship**: autor do nome científico;

³ Por exemplo, se somente tiver sido identificado a família e o gênero de um espécime, seu **taxonrank** será gênero; se tiver sido identificado a família, gênero, espécie e subespécie, seu **taxonrank** será subespécie.

- **type:** tipo de registro realizado (imagem, video ou áudio);
- **rights:** informação sobre direito de propriedade de um recurso;
- **associatedmedia:** lista de identificadores da mídia (publicação, identificador único global, URI) associada a uma ocorrência.
- **eventremarks:** comentários ou notas a respeito de um evento.

```

{
  "success": true,
  "result": [
    {
      "occurrenceid": 4701136,
      "rightsholder": "HCF",
      "bibliographiccitation": "Herbário Virtual REFLORA Disponível em:
        http://herbariovirtualreflora.jbrj.gov.br/reflora/herbarioVirtual
        /ConsultaPublicoHVUC/ConsultaPublicoHVUC.do?
        idTestemunho=4701136",
      "modified": "2018-05-02 08:17:09.214",
      "institutioncode": "HCF",
      "collectioncode": "HCF",
      "catalognumber": "HCF000007987",
      "occurrenceremarks": "",
      "recordedby": "H.B. Nogueira L.Rebellato;
        J.G.A. Lima Jr.",
      "recordnumber": "293",
      "othercatalognumbers": "",
      "eventdate": "28/4/2004",
      "year": 2004,
      "month": 4,
      "day": 28,
      "verbatimeventdate": "28/4/2004",
      "fieldnotes": "",
      "country": "Brasil",
      "countrycode": "BR",
      "stateprovince": "Mato Grosso",
      "municipality": "Barão de Melgaço",
      "locality": "",
      "minimumelevationinmeters": null,
      "maximumelevationinmeters": null,
      "verbatimlatitude": null,
      "verbatimlongitude": null,
      "decimallatitude": null,
      "decimallongitude": null,
      "identifiedby": "Wallnöfer, B.",
      "identificationverificationstatus": "Determinação online Reflora",
      "dateidentified": "2/5/2018",
      "identificationremarks": null,
      "identificationqualifier": null,
      "typestatus": null,
      "scientificname": "Ebenaceae Diospyros lasiocalyx (Mart.) B.Walln.",
      "family": "Ebenaceae",
      "genus": "Diospyros",
      "specificepithet": "lasiocalyx",
      "infraspecificepithet": null,
      "taxonrank": "ESPECIE",
      "scientificnameauthorship": "(Mart.) B.Walln.",
      "type": null,
      "rights": "Creative Commons",
      "associatedmedia": "http://imagens1.jbrj.gov.br/fsi/server?
        type=image&source=h/hcf/05001_10000/HCF000007987.JPG",
      "eventremarks": null
    }
  ]
}

```

Figura 4.3. Exemplo de arquivo JSON, retornado do Reflora para a exsicata do HCF apresentada na Figura 3.7.

A seguir, são apresentadas as funcionalidades do sistema speciesLink, do INCT-HVFF, necessárias para a implementação da extensão para o HCF- *Web*.

4.2. INCT-HVFF

O INCT-HVFF foi criado pelo governo federal em 2009, com o objetivo de reunir herbários brasileiros em rede, para disponibilizar informações textuais contidas nas etiquetas dos espécimes em uma plataforma *on-line*, acessível e aberta (MAIA et al., 2017). Em 2011, com o ingresso no Projeto Reflora, o INCT-HVFF passou a trabalhar no desenvolvimento de um servidor de imagens e serviço *web* associados, para armazenar, gerenciar e disponibilizar imagens *on-line* de plantas, em alta resolução, integrados com os seus dados textuais. Essa nova ferramenta permite ao usuário observar detalhes e medir partes de um espécime, e comparar imagens de amostras diferentes (INCT - HERBÁRIO VIRTUAL DA FLORA E DOS FUNGOS, 2018).

A iniciativa de disponibilizar imagens *on-line* até de amostras não identificadas, está contribuindo para o desenvolvimento da e-taxonomia ou *cybertaxonomy* no Brasil (INCT - HERBÁRIO VIRTUAL DA FLORA E DOS FUNGOS, 2018), visto que **têm permitido aos especialistas avaliarem, diretamente, imagens dos espécimes e sugerirem, via Internet, identificações e reidentificações aos curadores dos herbários** (BRAGANETO, 2013).

No período de 2012 a maio de 2019, 2,8 mil sugestões de reidentificação de nome de espécie já foram realizadas na rede speciesLink (CENTRO DE REFERÊNCIA EM INFORMAÇÃO AMBIENTAL, 2019). Apesar disso, a necessidade de atualização de informações e correção de erros referentes às exsicatas é um desafio a ser enfrentado pelo INCT-HVFF, para aprimorar e qualificar suas amostras: em setembro de 2018, entre os 6.173.278 registros disponíveis, 299.101 apresentavam identificação em branco, 227.475 somente o nome da família, e 744.713, o nome do gênero (CENTRO DE REFERÊNCIA EM INFORMAÇÃO AMBIENTAL, 2018c).

A plataforma speciesLink, que utiliza o padrão Darwin Core para a estruturação de dados, foi adotada pelo INCT-HVFF como seu sistema de informação. Isso ocorreu porque a speciesLink era a única plataforma no Brasil que apresentava abrangência nacional e reconhecimento internacional. Algumas características da arquitetura da rede implantada foram fundamentais para o sucesso do INCT-HVFF (MAIA et al., 2017):

- utilização de padrões e protocolos internacionais;
- política de dados das coleções biológicas definida por cada provedor, que determina quais informações são sigilosas e aquelas de acesso público, e as condições de uso dos dados;
- aceitação de dados sem necessidade de filtro de qualidade;

- realização de alterações e correções de dados de competência exclusiva do curador de cada herbário associado;
- reconhecimento de créditos ao trabalho realizado por cada herbário.

A plataforma speciesLink oferece um aplicativo denominado spLinker, desenvolvido em Java⁴. A instituição parceira deve poder utilizar o *software* que melhor atenda às suas necessidades de informatização do acervo, por isso, o spLinker possibilita: (i) que a leitura dos dados seja feita por meio de um mapeamento dos campos do *software* da coleção do herbário parceiro com o modelo de dados Darwin Core; (ii) a comunicação, por meio de uma ponte de dados em padrão Conectividade de Banco de Dados Aberta (ODBC, do inglês *Open Database Connectivity*)⁵ ou de um extrato dos dados em um arquivo texto, em que a primeira linha deve conter os nomes dos campos separados por tabulação (*tab delimited*). Uma vez configurado e os campos mapeados, o responsável pela coleção poderá, por meio do spLinker, enviar os dados não sensíveis da coleção para o servidor regional (SPECIESLINK, 2018f).

Dessa forma, o spLinker necessita ser configurado de acordo com o banco de dados da aplicação utilizada pelo herbário associado à rede speciesLink. Isso deve ser feito para que o spLinker possa acessar, no herbário parceiro, registros de certas relações que serão enviadas ao sistema speciesLink.

Outra diferença entre o sistema do Herbário Virtual Reflora e o speciesLink, é que este não fornece uma API. Entretanto, o site permite realizar consultas de amostras por meio de preenchimento de alguns campos, por exemplo, amostras disponibilizadas pelos herbários parceiros, determinador, data da coleta, município e país da coleta e taxonomia (classificação do espécime). Como resultado, é renderizada uma lista de amostras, com base nas opções escolhidas. O site permite, também, baixar o resultado dessa busca em arquivos de formatos XLS, XLSX e TXT. Para isso, o usuário deve informar o seu e-mail para receber um *link* para o *download* do arquivo escolhido, que é composto por: (i) 51 campos baseados no padrão de dados Darwin Core, com informações sobre as amostras; (ii) e por um número de linhas que corresponde à quantidade de amostras encontradas. A seguir, são descritos os campos presentes nos arquivos retornados (SPECIESLINK, 2018d):

- **datelastmodified**: data e hora em que foi realizada a modificação mais recente do registro;
- **institutioncode**: código padrão que identifica a instituição onde está depositado o item catalogado;
- **collectioncode**: valor alfanumérico único que identifica a coleção dentro da instituição;

⁴ O spLinker foi desenvolvido pelo CRIA, para que as instituições que não têm os seguintes pré-requisitos para servir dados diretamente à rede speciesLink, possam espelhar seus dados em servidores regionais: (i) conexão Internet rápida e estável; (ii) equipamento e *software* adequados para servir os dados à rede 24 horas por dia; (iii) e equipe qualificada para a manutenção do sistema (SPECIESLINK, 2018f).

⁵ ODBC é uma API que permite que o *software* invoque um sistema gerenciador de bancos de dados (ELMASRI; NAVATHE, 2011).

- **catalognumber**: valor alfanumérico único que identifica um registro individual de uma coleção;
- **scientificname**: nome científico completo do item catalogado;
- **basisofrecord**: abreviatura que indica se o registro é uma observação (O), um organismo vivo coletado (L), um espécime de uma coleção/museu (S), uma foto (P), uma semente coletada (G) etc.;
- **kingdom**: reino ao qual o espécime pertence;
- **phylum**: filo ao qual o espécime pertence;
- **class**: nome da classe do espécime;
- **ordem**: nome da ordem do espécime;
- **family**: nome da família do espécime;
- **genus**: nome do gênero do espécime;
- **species**: nome da espécie do espécime;
- **subspecies**: nome da subespécie do espécime;
- **scientificnameauthor**: nome do autor do nome científico do espécime;
- **identifiedby**: nome do determinador que identificou o espécime;
- **yearidentified**: ano em que o item da coleção foi identificado;
- **monthidentified**: mês em que o item da coleção foi identificado;
- **dayidentified**: dia em que o item da coleção foi identificado;
- **typestatus**: categoria de tipo nomenclatural (holótipo, parátipo, isótopo etc.) que o espécime representa;
- **collectornumber**: número de identificação aplicado a um espécime no momento da coleta;
- **fieldnumber**: número criado no momento da coleta para identificar todo o material resultante de um evento de coleta;
- **collector**: nome da pessoa ou organização responsável pela coleta;
- **yearcollected**: ano em que o espécime foi coletado;
- **monthcollected**: mês em que o espécime foi coletado;
- **daycollected**: dia em que o espécime foi coletado;
- **julianday**: dia ordinal do ano em que ocorreu a coleta, ou seja, o número de dias desde o 1º de janeiro daquele ano;
- **timeofday**: hora do dia em que o espécime foi coletado, expressa em hora decimal, a partir da meia-noite. Por exemplo: 13,5 = 13h:30min;
- **continentocean**: continente ou oceano onde o espécime foi coletado;
- **country**: país onde o espécime foi coletado;
- **stateprovince**: estado onde o espécime foi coletado;
- **county**: município onde o espécime foi coletado;
- **locality**: descrição do local onde o espécime foi coletado;

- **longitude**: longitude do local onde o espécime foi coletado;
- **latitude**: latitude do local onde o espécime foi coletado;
- **coordinateprecision**: estimativa da precisão com que uma localidade foi expressa, nos campos latitude e longitude;
- **boundingbox**: ponto de acesso que oferece um mecanismo para realização de busca, usando uma caixa geográfica delimitadora;
- **minimumelevation**: distância mínima, em metros, acima ou abaixo do mar, da localidade onde a coleta foi realizada;
- **maxmimelevation**: distância máxima, em metros, acima ou abaixo do mar, da localidade onde a coleta foi realizada;
- **minimumdepth**: distância mínima, em metros, abaixo da superfície da água, onde a coleta foi realizada;
- **maximumdepth**: distância máxima, em metros, abaixo da superfície da água, onde a coleta foi realizada;
- **sex**: sexo do espécime;
- **preparationtype**: tipo de preparação empregado para a preservação da amostra;
- **individualcount**: número de indivíduos presentes no lote coletado;
- **previouscatalognumber**: número anterior de catálogo, se o item já foi identificado por outro número, no catálogo atual ou de outra instituição;
- **relationshipstype**: valor nomeado ou codificado que identifica o tipo de relacionamento entre um item da coleção e o item referência da coleção;
- **relatedcatalogitem**: identificador completo de um item do catálogo, composto por código da instituição, código da coleção e número de catálogo;
- **notes**: anotações sobre o espécime.

Não foi encontrada documentação com informações sobre as colunas `latitude_mun`, `longitude_mun` e `barcode` da planilha.

4.3. Discussão

Nas seções anteriores, foram apresentados: (i) o Herbário Virtual Refflora, que em maio de 2019 oferecia imagens e informações textuais de 25.584 exemplares do HCF (REFLORA, 2018b); (ii) e o INCT-HVFF, que disponibiliza informações textuais de toda a coleção do HCF, que em maio de 2019 era composta de 28.476 exsicatas (SPECIESLINK, 2018a).

Esses dois herbários virtuais permitem reidentificação de espécimes: (i) quando um determinador descobre uma incorreção nas informações disponíveis sobre as exsicatas; (ii) quando o curador de um herbário associado encontra um erro e solicita a um especialista, por e-mail, uma nova identificação.

Dessa forma, a sugestão de reidentificação pode ser feita quando for observada

incorreção no nome científico de um espécime, ou em sua classificação indicativa de família, gênero ou espécie. Segundo o curador do HCF, essa possibilidade de reidentificar é de grande importância, porque contribui para a validação das informações referentes às exsicatas da coleção.

Até o ano de 2018, as reidentificações eram registradas apenas nos bancos de dados dos herbários virtuais, mas não nos sistemas dos herbários físicos onde as exsicatas estão depositadas. A comparação dos dados do HCF-*Web* com os do Re flora e do INCT-HVFF tinha, portanto, de ser feita manualmente, por funcionários que verificavam se todas as informações presentes nesses dois herbários virtuais eram as mesmas do banco de dados local. Essa comparação demandava tempo, por isso a atualização não era realizada periodicamente.

Considerando que a atualização manual era um processo demorado e exaustivo, por causa da enorme quantidade de exsicatas que fazem parte do acervo do HCF e, ainda, a importância de sua realização, porque permite corrigir informações sobre as amostras, **foi implementado neste trabalho, uma extensão para o HCF-*Web*, que permite atualizar seus dados, comparando-os com os seus equivalentes, no Herbário Virtual Re flora e no sistema speciesLink, do INCT-HVFF, a fim de identificar inconsistências.** No próximo capítulo, será apresentado esse trabalho de integração.

Integração do HCF-*Web* a herbários virtuais

Informações textuais de todas as 28.476 exsicatas do HCF estão disponíveis no acervo *online* do INCT-HVFF, e cerca de 90% das informações textuais e imagens da coleção, no Herbário Virtual Refflora. Sugestões de reidentificações das amostras do HCF, realizadas por especialistas nos sites desses dois herbários virtuais, contribuem para a validação das informações referentes, por exemplo, à classificação das famílias e gêneros dos espécimes da coleção. Como essas reidentificações geravam registros apenas nos herbários virtuais, era preciso atualizar manualmente o banco de dados do HCF. A verificação das informações de mais de 28 mil espécimes era um processo lento - que demorava semanas - e estava sujeito a erros. A implementação de uma extensão para o HCF-*Web*, portanto, era uma questão importante a ser resolvida.

Para a implementação dessa extensão, que permite encontrar inconsistências entre os dados da coleção do HCF-*Web* e os presentes no Herbário Virtual Refflora e no sistema speciesLink, do INCT-HVFF, foram utilizadas funcionalidades disponíveis nos dois herbários virtuais. Dessa forma, na Seção 5.1 é descrito o processo de comparação dos dados do HCF-*Web* com os seus equivalentes, no Herbário Virtual Refflora; na Seção 5.2 é apresentado o processo de comparação dos dados do HCF-*Web* com os seus equivalentes, no sistema speciesLink; na Seção 5.3 são descritas as interfaces elaboradas para o usuário atualizar os dados do HCF-*Web*; na Seção 5.4 são apresentadas as decisões de projeto para a implementação da extensão; e na Seção 5.5 são apresentadas estimativas dos tempos de execução para a identificação de divergências entre os dados do HCF-*Web* e os dos herbários virtuais.

5.1. Herbário Virtual Refflora

Para identificar divergências existentes entre os dados dos espécimes disponíveis no HCF-*Web* e os seus equivalentes, no Refflora, foi desenvolvido um algoritmo, cuja representação está

ilustrada no diagrama de atividades, apresentado na Figura 5.1.

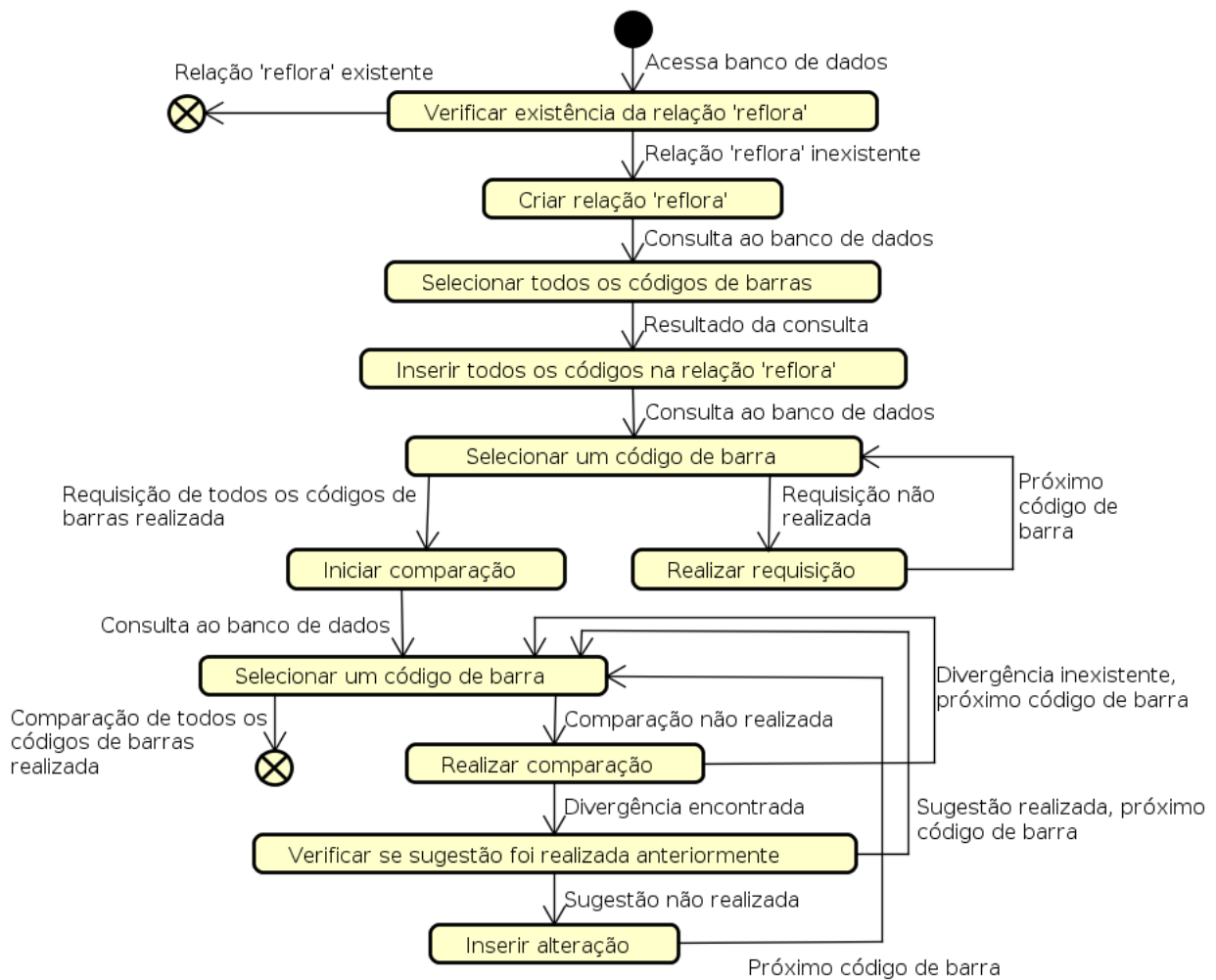


Figura 5.1. Diagrama de atividades, que descreve a atualização do HCF-Web, a partir dos dados do Herbário Virtual Reflora.

Primeiramente, o sistema verifica se a relação **reflora** está presente no banco de dados do HCF-Web. Essa relação é criada temporariamente, ou seja, somente ocorre quando o processo de identificação de divergências está em curso. Os atributos presentes nessa relação podem ser visualizados na Tabela 5.1. O objetivo dessa relação é garantir que: (i) todos os códigos de barras sejam requisitados; (ii) todas as comparações sejam realizadas; (iii) duas ou mais execuções não ocorram simultaneamente; (iv) e o código de barra seja requisitado por até três vezes, quando não ocorrer a resposta esperada para uma requisição.

Tabela 5.1. Atributos da relação *reflora*.

Atributos	Descrição
<i>id</i>	Identificador da relação.
<i>cod_barra</i>	Código de barra do tombo do HCF.
<i>tombo_json</i>	Resposta da requisição feita ao Refflora.
<i>ja_comparou</i>	Valor, do tipo booleano: true indica que a comparação de informações já foi realizada; e false , que não foi feita.
<i>ja_requisitou</i>	Valor, do tipo booleano: true indica que a requisição das informações já ocorreu; e false , que não ocorreu.
<i>nro_requisicoes</i>	Registro de tentativas de requisições realizadas, que pode variar de 1 a 3.

Se a relação *reflora* não existir no banco de dados, ela será criada. A seguir, ocorrerá consulta ao banco de dados do HCF-*Web*, para obtenção dos valores referentes aos códigos de barras de todas as exsicatas do acervo, que serão, então, inseridos na relação *reflora*.

Os códigos de barra serão selecionados um a um. A requisição de um código de barra será feita à API do Herbário Virtual Refflora se o sistema verificar que isso ainda não ocorreu. O JSON retornado pela API do Refflora, contendo informações sobre a exsicata selecionada, será salvo no atributo da relação *reflora*, como explicado anteriormente, na Seção 4.1. Caso a API não retorne informações da exsicata selecionada, o valor do atributo *nro_requisicoes* será decrementado, até atingir o valor zero. Quando isto ocorrer, a requisição para esse código de barra cessará. A decrementação evitará que muitas requisições de um código de barra sejam realizadas, até a informação da exsicata ser obtida e, conseqüentemente, que o tempo do processo de comparação seja elevado.

Uma vez terminado o processo de seleção, requisição e registro dos dados das exsicatas na relação *reflora*, será inicializada a comparação das informações. Uma nova consulta à relação *reflora*, do banco de dados do HCF-*Web*, será realizada. O sistema retornará, um a um, códigos de barra cujas informações não tiverem sido comparadas (*ja_comparou=false*). A seguir, os valores do JSON retornados do Herbário Refflora, referentes a um código de barra, serão comparados com os valores equivalentes, nas relações e atributos do HCF-*Web*, ilustrados na Tabela 5.2.

Tabela 5.2. Chaves do JSON retornadas da requisição ao Refflora, e suas respectivas equivalências, nas relações e atributos do HCF-*Web*.

Chave do JSON (Refflora)	Relações (HCF- <i>Web</i>)	Atributos (HCF- <i>Web</i>)
<code>family</code>	<code>familias</code>	<code>nome</code>
<code>genus</code>	<code>generos</code>	<code>nome</code>
<code>specificepithet</code>	<code>especies</code>	<code>nome</code>
<code>infraespecificepithet</code>	<code>subespecies</code> ou <code>variedades</code>	<code>nome</code>

Apesar de o HCF-*Web* permitir a reidentificação da subfamília de um espécime, não é possível comparar informações desse nível hierárquico, porque a API do Refflora não retorna dados sobre subfamílias das amostras de nenhum herbário parceiro¹.

A API do Refflora retorna uma chave que indica o epíteto infraespecífico, ou seja, uma palavra que corresponde à subespécie, à variedade ou à forma de um determinado espécime da coleção do HCF. Para saber se essa palavra refere-se à subespécie ou à variedade, é preciso analisar a chave `taxonrank` do Refflora, que indica o nível taxonômico mais específico até onde o espécime foi identificado, por exemplo, gênero, espécie, subespécie, variedade, forma etc. Se o `taxonrank` da planta for subespécie ou variedade, será comparado o valor retornado pela chave `infraespecificepithet` do Refflora com o valor, respectivamente, da relação `subespecie` ou `variedade`, do HCF-*Web*. Se o `taxonrank` do espécime for outro, a comparação não ocorrerá.

Para cada tombo do HCF, cujos valores forem comparados com os do Refflora, será gerado um JSON, que apontará as divergências existentes. Por exemplo, no Refflora foi identificada a seguinte sugestão de reidentificação de gênero, no tombo 5735 (código de barra HCF000015559):

```
{"especie_nome": "falcata"}
```

Antes de incluir uma sugestão de alteração no HCF-*Web*, é preciso averiguar se ela não foi realizada anteriormente. Dessa forma, serão analisadas, na relação `alteracoes`, todas as modificações existentes no tombo, para que não ocorram sugestões duplicadas. A proposta de reidentificação será inserida nessa relação, somente quando o encaminhamento não tiver sido feito, e ficará disponível no HCF-*Web*, para análise pelo curador, como pode ser visto na Figura 5.2.

¹ Foi realizada consulta em sete herbários que disponibilizam seus acervos no Herbário Virtual Refflora: para nenhum deles é retornada informação sobre subfamília.

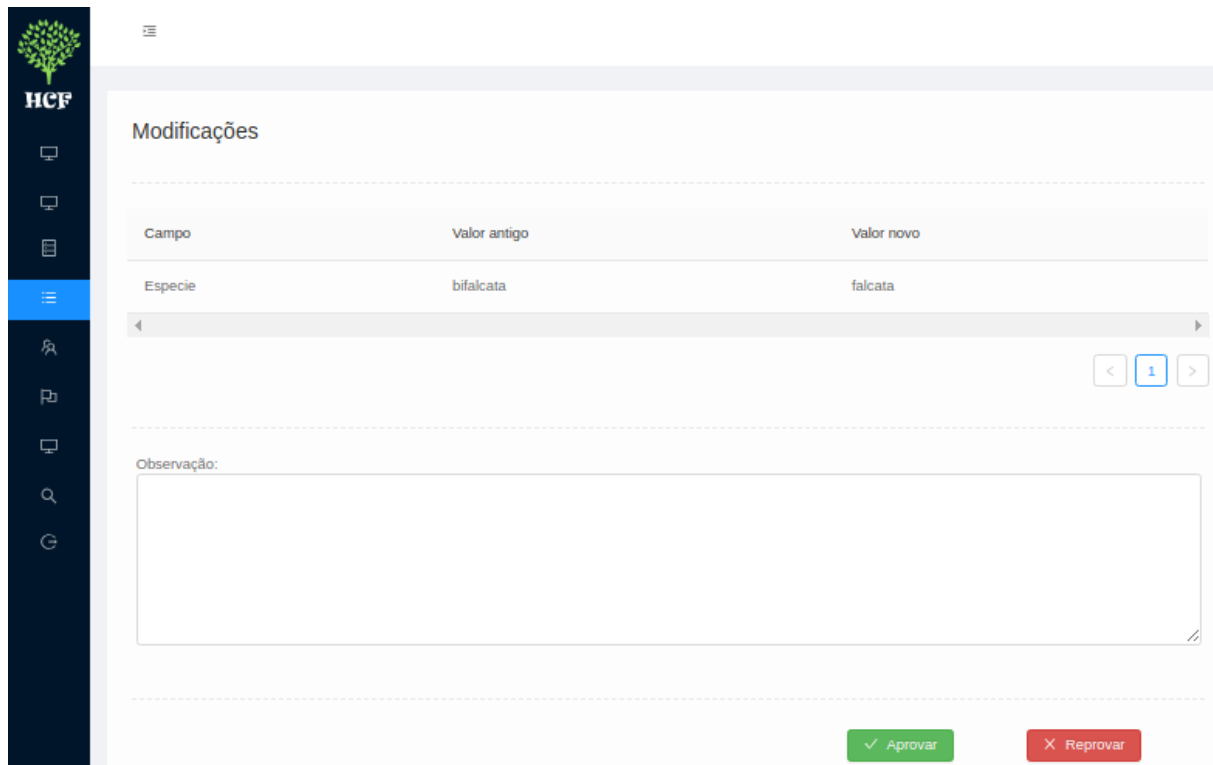


Figura 5.2. Interface com uma pendência decorrente de uma alteração sugerida no Herbário Virtual Reflora Reflora, e que foi encontrada pelo serviço.

Portanto, com a integração, o processo de atualização dos dados do HCF-Web pode ser resumido da seguinte forma: registros do acervo do HCF-Web são disponibilizados via *script* Darwin Core ao Reflora, e este retorna informações equivalentes via API, para que a comparação dos dados possa ser feita, como ilustrado na Figura 5.3.

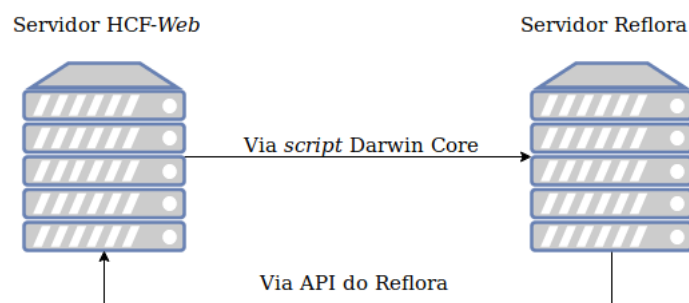


Figura 5.3. Processo de atualização do HCF-Web, em que dados de seu acervo são enviados via *script* Darwin Core ao Reflora, que retorna informações equivalentes, por meio de sua API, para que a comparação seja realizada.

A seguir, é descrito o processo de integração do HCF-Web ao sistema speciesLink, do INCT-HVFF.

5.2. Sistema speciesLink

Considerando que o sistema speciesLink, do INCT-HVFF, não fornece uma API, o seu algoritmo, representado na Figura 5.4, é menos complexo que o do Herbário Virtual Refflora.

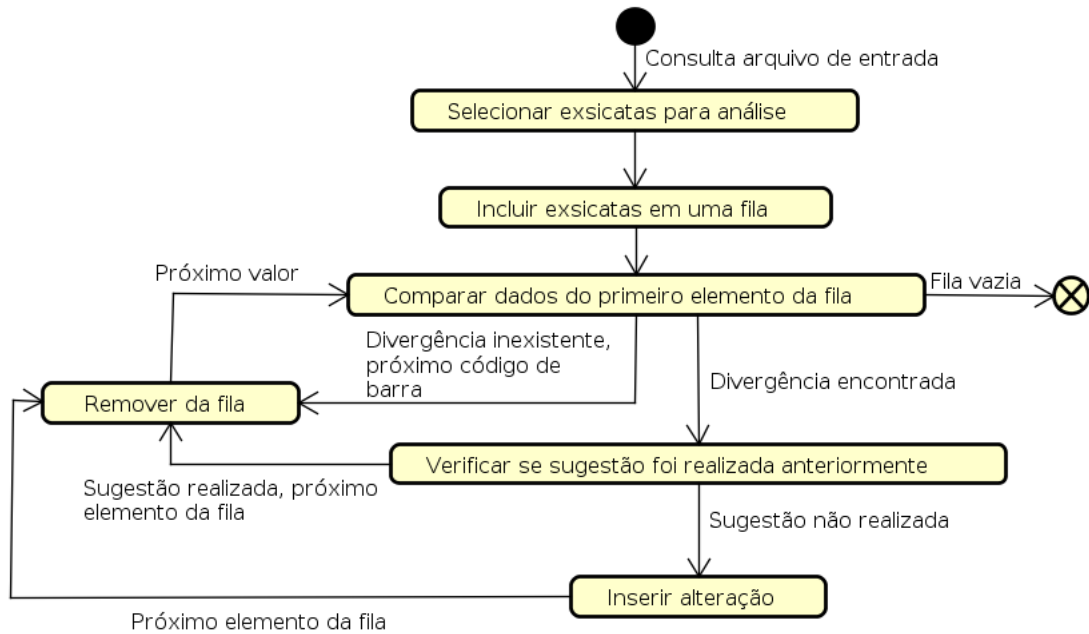


Figura 5.4. Diagrama das atividades para a atualização do HCF-Web, a partir dos dados do speciesLink.

Para identificar divergências entre os dados do HCF-Web e os seus equivalentes, no INCT-HVFF, é preciso utilizar o arquivo retornado pelo speciesLink, em que cada linha corresponde a uma exsicata e apresenta 51 campos, com informações sobre, por exemplo, o nome científico da planta, o município e o país da coleta, a data da coleta e o nome do determinador.

Primeiramente, como ilustrado no diagrama da Figura 5.4, o sistema seleciona, para análise, as exsicatas do arquivo retornado. As linhas desse arquivo são, então, carregadas em uma fila, e os valores retornados dos campos **family**, **genus**, **species** e **subspecies** são comparados com os equivalentes, nas relações e atributos do HCF-Web apresentados na Tabela 5.3.

Tabela 5.3. Colunas do arquivo do speciesLink e suas respectivas equivalências, nas relações e atributos do HCF-*Web*.

Colunas do arquivo (speciesLink)	Relações (HCF- <i>Web</i>)	Atributos (HCF- <i>Web</i>)
family	familias	nome
genus	generos	nome
species	especies	nome
subspecies	subspecies	nome

O sistema speciesLink, assim como o do Herbário Virtual Reflora, não retorna informações sobre subfamília. Além disso, como é possível constatar no manual do *software spLinker*², não há opção para enviar informações sobre variedade.

O restante do processo de comparação de dados é semelhante ao do Herbário Virtual Reflora: para cada tombo do HCF, cujos valores forem comparados com os do speciesLink, será gerado um JSON, que apontará as divergências existentes. Por exemplo, no tombo 25493 (código de barra HCF000026009) foi identificada a seguinte sugestão de reidentificação de espécie:

```
{"especie_nome": "aphylla"}
```

Antes de incluir uma sugestão de alteração no HCF-Web, é preciso averiguar se ela não foi realizada anteriormente. Dessa forma, serão analisadas, na relação `alteracoes`, todas as modificações existentes no tombo, para que não ocorram sugestões duplicadas. A proposta de reidentificação será inserida nessa relação, somente quando o encaminhamento não tiver sido feito, e ficará disponível no HCF-*Web*, para análise pelo curador, como pode ser visto na Figura 5.5.

² <http://splink.cria.org.br/files/splinker/Manual.pdf>

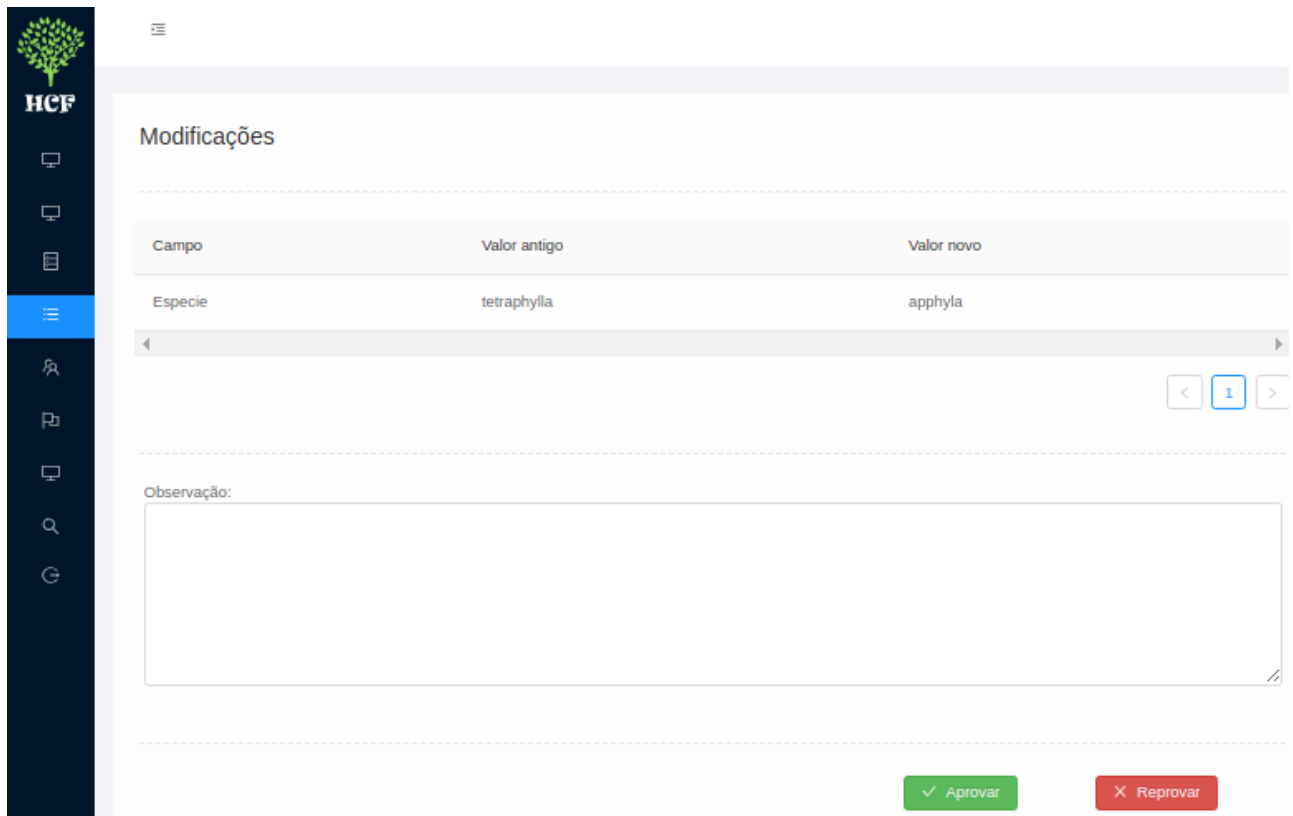


Figura 5.5. Interface com uma pendência decorrente de uma alteração sugerida no speciesLink, e que foi encontrada pelo serviço.

Portanto, com a integração, o processo de atualização dos dados do HCF-Web pode ser resumido da seguinte forma: registros do acervo do HCF-Web são disponibilizados via *software* spLinker ao speciesLink, do INCT-HVFF, e este retorna informações equivalentes por meio de um arquivo, para que a comparação dos dados possa ser feita, como ilustrado na Figura 5.6.

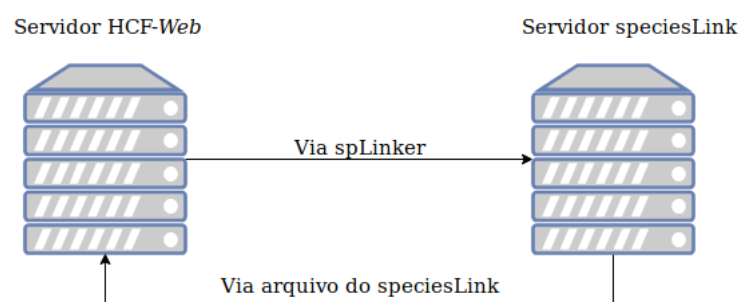


Figura 5.6. Processo de atualização do HCF-Web, em que dados de seu acervo são enviados via *software* spLinker ao speciesLink, que retorna informações equivalentes, por meio de um arquivo, para que a comparação seja realizada.

Na próxima seção são apresentadas as novas interfaces, disponíveis aos usuários do HCF-Web.

5.3. Interfaces

Na interface relacionada ao Herbário Virtual Refflora, o usuário pode optar entre a atualização imediata ou programada. Neste caso, a periodicidade poderá ser semanal, mensal ou bimestral. Além disso, é fornecida a informação da data da atualização mais recente e de seu tempo de duração. É possível, também, selecionar e visualizar os logs gerados na execução. A interface relacionada ao Herbário Virtual Refflora é apresentada na Figura 5.7.

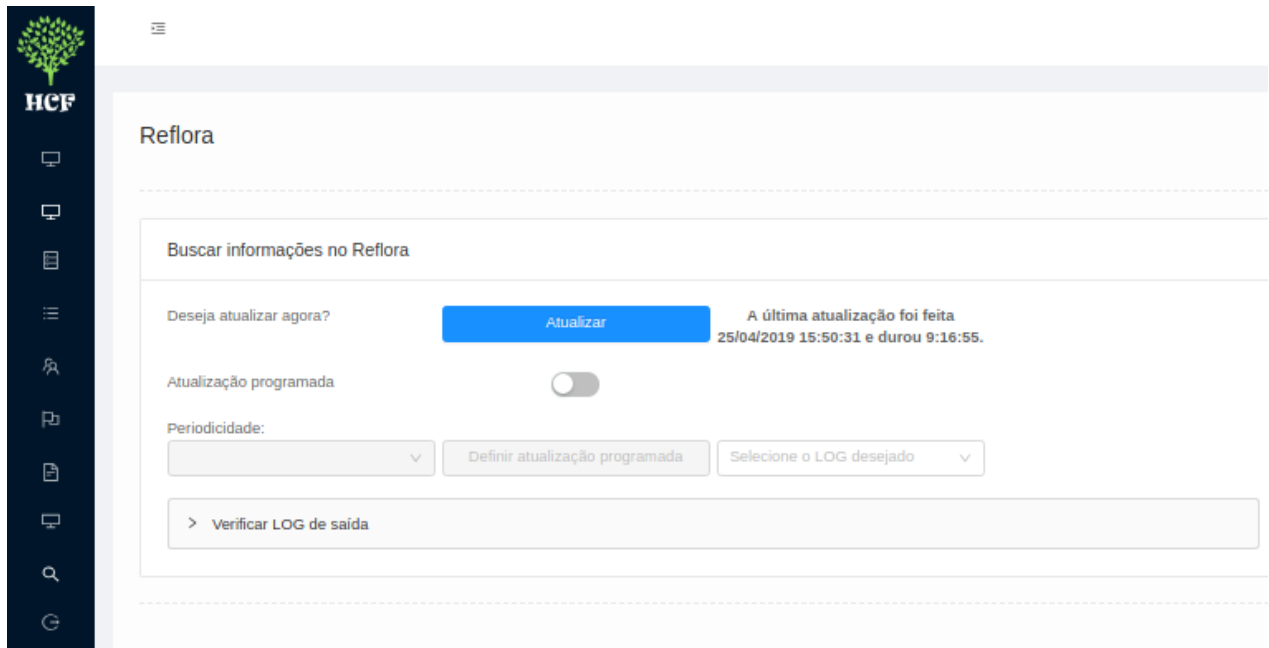


Figura 5.7. Interface da funcionalidade relacionada ao Herbário Virtual Refflora, que oferece opções de atualização (imediata ou programada) e de verificação de log.

Ao solicitar uma atualização imediata ou programada, o usuário recebe, no lado direito e superior da tela, um *feedback*, ou seja, uma notificação de que a atualização será inicializada em breve, como é possível visualizar na Figura 5.8. Cabe ressaltar que quando uma atualização imediata está sendo executada, todos os botões permanecem desabilitados, e somente é possível selecionar e visualizar logs. Uma atualização programada mas ainda não realizada, pode ser cancelada pelo usuário, e uma atualização imediata pode, então, ser executada.

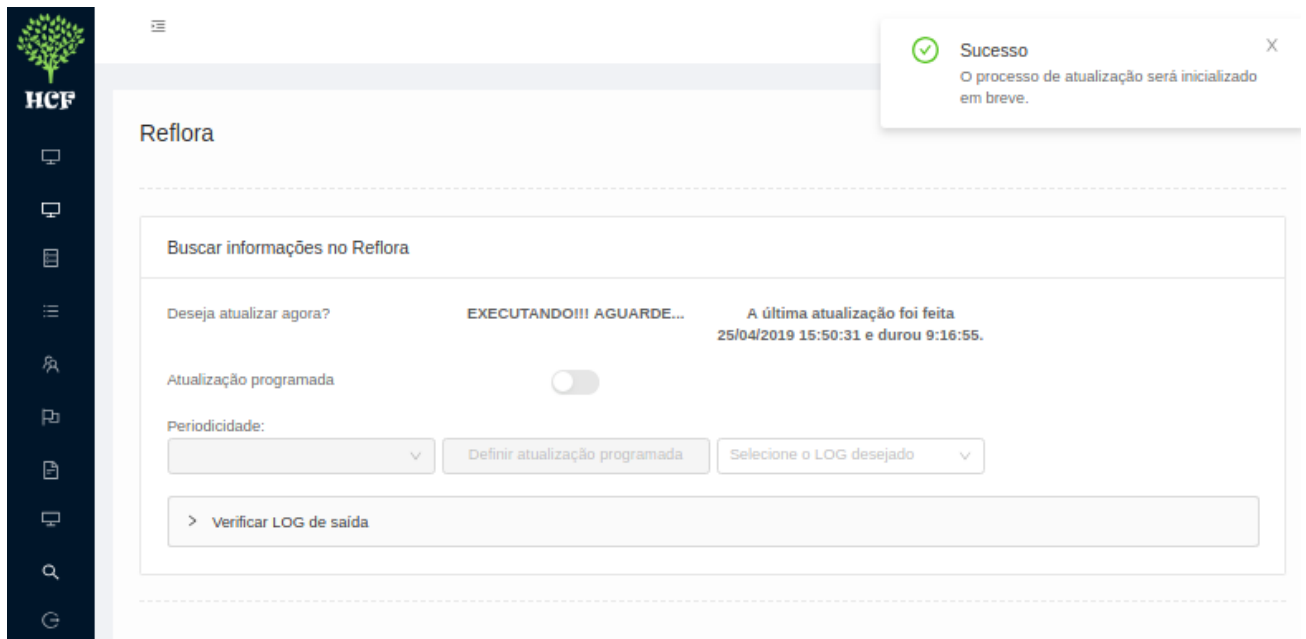


Figura 5.8. *Feedback* dado ao usuário, quando uma atualização imediata é solicitada.

A interface relacionada ao sistema speciesLink é muito semelhante à implementada para a integração ao Herbário Virtual Reflora. As duas interfaces: (i) informam a data de realização da atualização mais recente e o seu tempo de duração; (ii) permitem a seleção e a visualização dos logs das execuções do herbário virtual; (iii) notificam o usuário de que a atualização será iniciada brevemente; (iv) e mantêm os botões desabilitados enquanto estiver ocorrendo a execução. A única diferença entre as duas interfaces é que somente naquela relacionada ao speciesLink está disponível o botão upload, como pode ser visto na Figura 5.10.

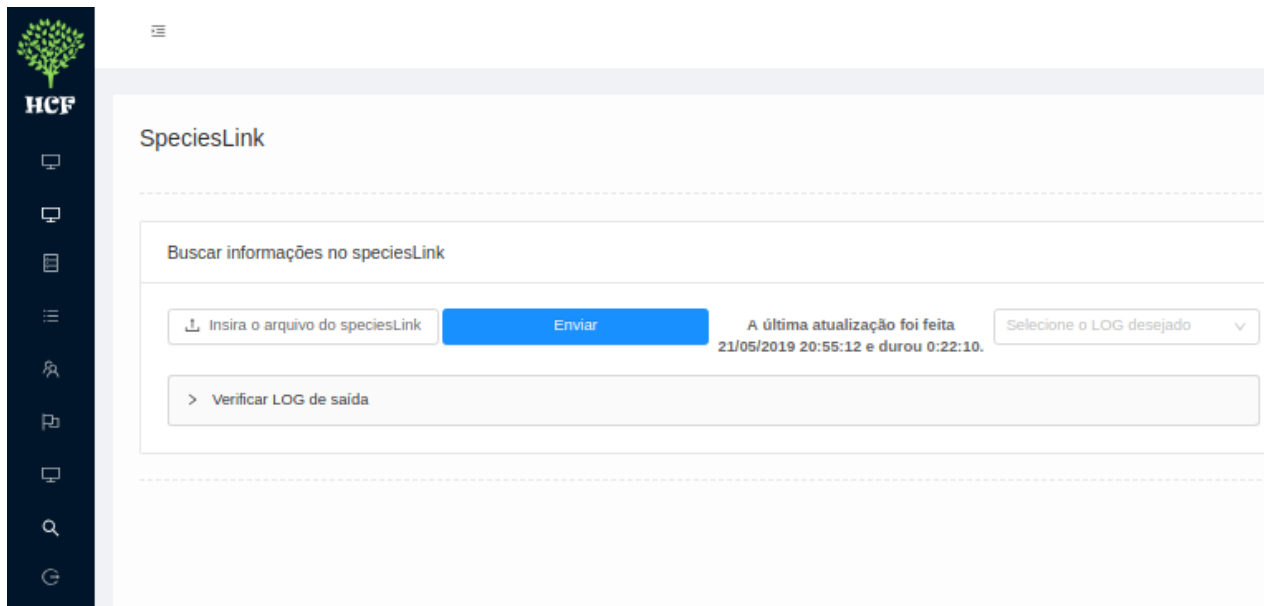


Figura 5.9. Interface da funcionalidade relacionada ao sistema speciesLink, que oferece opções de submissão de arquivos e verificação de logs.

Essas duas novas funcionalidades, implementadas no *HCF-Web*, estão representadas no diagrama de caso de uso apresentado na Figura 5.10.

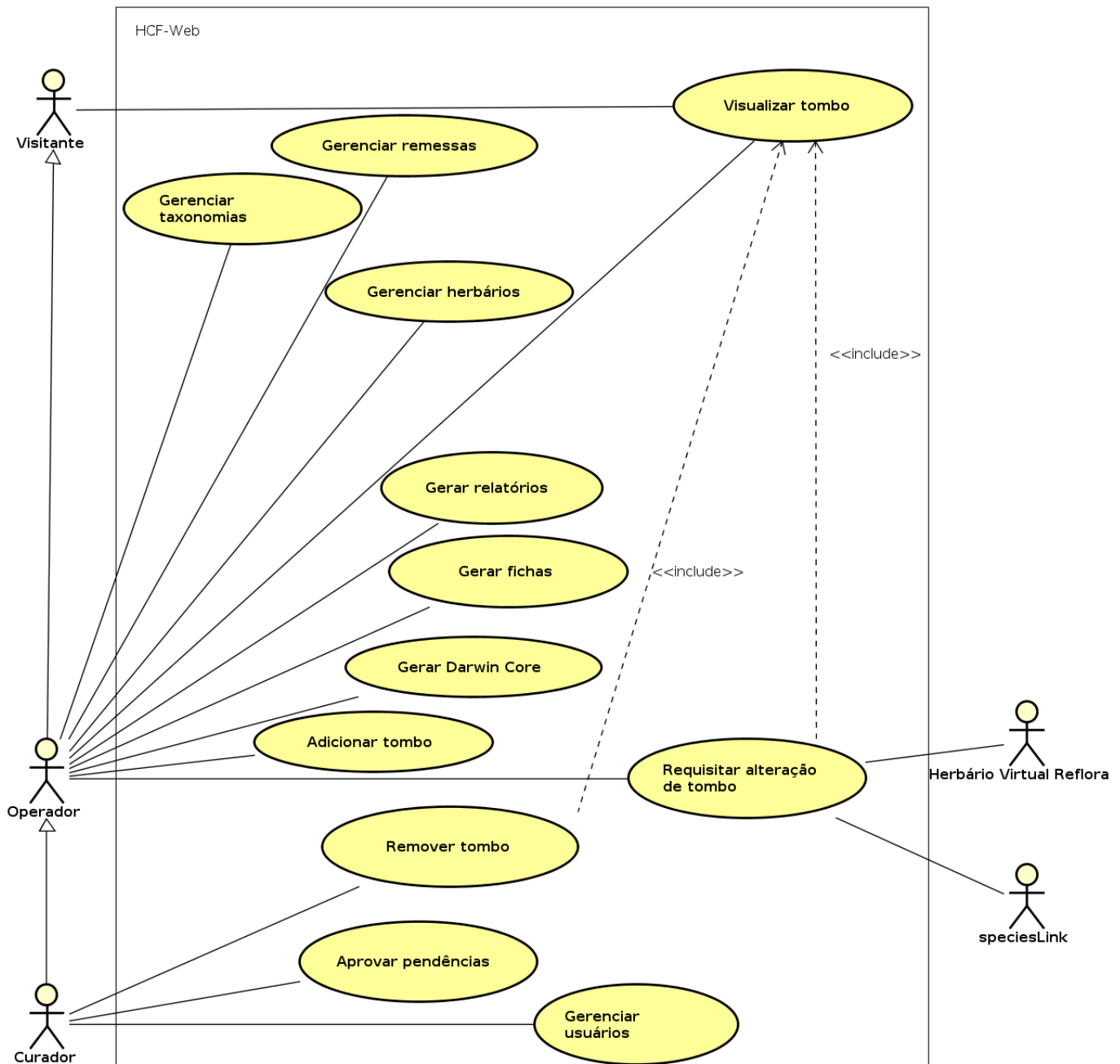


Figura 5.10. Diagrama de caso de uso, com a implementação dos dois serviços, em que dois novos atores foram adicionados.

Na próxima seção são apresentadas as decisões de projeto para a implementação das novas funcionalidades.

5.4. Decisões de projeto

Para a implementação das novas funcionalidades, foram utilizados módulos que possibilitassem a identificação de divergências entre os dados do HCF e os seus equivalentes, nos dois herbários virtuais. Na integração do HCF ao Herbário Virtual Reflora, foram utilizados dois módulos:

(i) o `throttled-queue`³ - versão 1.0.5, para definir o número de requisições a serem realizadas em um determinado intervalo de tempo; (ii) e o `request`⁴ - versão 2.88.0, para requisitar a API do Herbário Virtual Refflora. Na integração do HCF ao speciesLink, foi utilizado o módulo `fs`⁵ - versão 0.0.1, para a leitura do arquivo de entrada. Nas duas integrações, foi utilizado o módulo `q`⁶ - versão 1.5.1, para ocorrência de execução síncrona, ou seja, para que a execução de uma função ou linha de código seja inicializada somente após a anterior ter sido finalizada.

A relação `configuracao` foi criada para os serviços do Refflora e do speciesLink, para garantir: (i) que não sejam realizadas múltiplas execuções de um mesmo serviço e, por conseguinte, que não ocorram inconsistências. Apesar disso, os serviços do Refflora e do speciesLink podem ser executados simultaneamente; (ii) que sejam executadas as atualizações programadas pelo usuário; (iii) que somente seja executado o mais recente arquivo retornado pelo speciesLink. Os atributos da relação `configuracao` podem ser visualizados nas Tabelas 5.4 e 5.5.

Tabela 5.4. Atributos da relação `configuracao`.

Atributos	Descrição
<code>id</code>	Identificador dos registros na relação.
<code>hora_inicio</code>	Data e hora de início do processo de atualização dos dados.
<code>hora_fim</code>	Data e hora de término do processo de atualização dos dados.
<code>periodicidade</code>	Atributo utilizado somente para registro relacionado ao serviço do Herbário Virtual Refflora, e para solicitação de atualização programada, a cada 7, 30 ou 60 dias.

³ <https://www.npmjs.com/package/throttled-queue>

⁴ <https://www.npmjs.com/package/request>

⁵ <https://www.npmjs.com/package/fs>

⁶ <https://www.npmjs.com/package/q>

Tabela 5.5. Atributos da relação `configuracao`. (continuação)

Atributos	Descrição
<code>data_proxima_atualizacao</code>	Atributo utilizado somente para registro relacionado ao serviço do Herbário Virtual Reflora, e para solicitação de atualização programada, a cada 7, 30 ou 60 dias.
<code>nome_arquivo</code>	Atributo utilizado somente para registro relacionado ao serviço do sistema speciesLink. Utilizado para armazenar o nome do arquivo que foi submetido, para realização de comparação.
<code>servico</code>	Nome do serviço que está sendo executado.

A inicialização e a finalização de cada processo de comparação de dados são registradas em um arquivo de log, para cada um dos serviços. Adicionalmente, no serviço do Herbário Virtual Reflora, quando uma requisição tem como retorno uma falha, ela é registrada no arquivo de log, juntamente com o código de barra da exsicata em que isso ocorreu. O usuário pode também consultar esse arquivo de log para identificar os códigos de barras que apresentam problemas. Um exemplo de log de execução é apresentado na Figura 5.11.

```
[01/03/2019-21:40:14] Inicializando a aplicação do Reflora.
[02/03/2019-01:47:24] Falha na requisição do código de barra HCF000006041
que foi Error: connect ETIMEDOUT 200.20.168.136:80
[02/03/2019-01:49:36] Falha na requisição do código de barra HCF000006041
que foi Error: connect ETIMEDOUT 200.20.168.136:80
[02/03/2019-01:51:47] Falha na requisição do código de barra HCF000006041
que foi Error: connect ETIMEDOUT 200.20.168.136:80
[02/03/2019-23:58:12] O processo de comparação do Reflora acabou.
```

Figura 5.11. Exemplo de arquivo de log, contendo os registros das falhas que ocorreram durante as três requisições ao Reflora.

Na próxima seção são apresentados os testes realizados para medida do tempo de execução de cada serviço, e os números de divergências e falhas registradas.

5.5. Estimativa de tempo de execução

Os testes apresentados a seguir foram realizados em uma máquina com a seguinte configuração: Intel Core 2 Duo E7500 2.93GHz, com 3GB de Memória de Acesso Randômico (do inglês, *Random Access Memory*) e um disco rígido de 320GB. Foram utilizados: (i) o sistema operacional Linux Mint 19 MATE; (ii) o Node.js - versão 10.15.1; (iii) e o banco de dados MySQL - versão 5.6.40.

Para a validação das duas novas funcionalidades, ou seja, a atualização dos dados do HCF-*Web* em relação às alterações encontradas no Herbário Virtual Refflora e no speciesLink, foi realizada no período de 25 de abril a 15 de junho de 2019. Foram selecionadas, aleatoriamente, conjuntos de amostras da coleção do HCF, compostos por 5 mil, 7,5 mil, 10 mil, 12,5 mil, 15 mil, 17,5 mil, 20 mil, 22,5 mil e 25 mil espécimes⁷. Este último conjunto de amostra selecionado é próximo do acervo total do HCF, composto por cerca de 28 mil plantas. Dessa forma, a quantidade de pendências e a média de tempo de execução das 25 mil exsicatas devem ser muito próximas ao cenário real. Cada amostra foi testada três vezes seguidas: (i) para calcular a média do tempo de execução, em minutos; (ii) e para verificar o número de divergências⁸. Como o Herbário Virtual Refflora oferece uma API, podem ocorrer problemas no servidor ou de conexão da máquina com o servidor ou com a Internet. Por esse motivo, também foram quantificadas as falhas geradas.

Conforme esperado, os resultados da funcionalidade do HCF-*Web*, relacionada ao Refflora, evidenciaram que as médias e os desvios-padrões dos tempos de execução, com exceção da amostra de 22,5 mil exsicatas, variaram de acordo com o tamanho da amostra, como é possível observar na Figura 5.12. Por exemplo, o grau de variação da média do tempo de execução das 5.000 exsicatas (d.p. = 6,9 min) foi muito menor que a das 25.000 exsicatas (d.p. = 40,3 min), porque, provavelmente, quanto maior o tempo necessário para a execução, maiores são as chances de ocorrência de problemas no servidor ou de conexão da máquina com o servidor ou com a Internet.

⁷ Nesse período estava sendo realizada a migração dos dados de 27.045 exsicatas do Herbarium para o HCF-*Web*, por isso não foi possível testar os dados do acervo total do HCF, composto por 28.476 espécimes

⁸ Para cada amostra, antes do início da segunda e terceira execuções, foram apagadas as divergências encontradas na execução anterior.

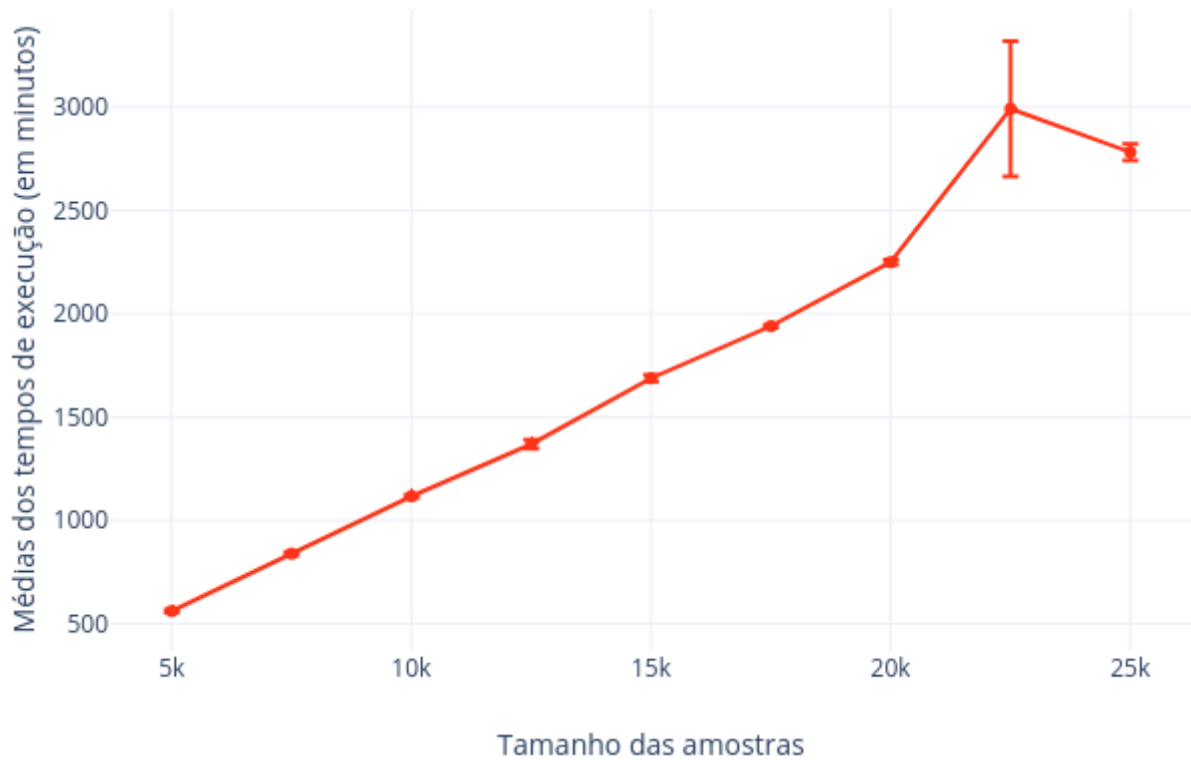


Figura 5.12. Médias e desvios-padrões dos tempos de atualização de cada amostra do HCF-*Web*, em relação aos dados equivalentes no Herbário Virtual Refflora.

Talvez por causa desses problemas, a média do tempo de execução da amostra de 22,5 mil exsicatas tenha sido muito elevada ($\bar{x}=2.992,6$ min). Na primeira execução, o tempo de execução foi de, aproximadamente, 2.530 minutos; na segunda, de 3.207 minutos; e na terceira, de 3.240 minutos. Isso resultou em um grande desvio-padrão (d.p.=327,2 min). Considerando que a média das falhas também foi alta ($\bar{x}=441,6$), é possível que tenham ocorrido problemas de conexão com a Internet nas execuções das 22,5 mil exsicatas, os quais ocasionaram média de tempo elevada.

Em relação às divergências nos dados do HCF-*Web*, relacionadas ao Refflora, também foi observado que, de forma geral, elas variaram de acordo com o tamanho da amostra, como pode ser visto na Tabela 5.6.

Tabela 5.6. Médias das divergências relacionadas ao Herbário Virtual Refflora.

Tamanho da amostra	Média das divergências
5.000	886
7.500	1166
10.000	1527
12.500	1910,6
15.000	2370
17.500	2757,3
20.000	3332,3
22.500	3498
25.000	3896,3

Ainda em relação aos testes com o Herbário Virtual Refflora, o número de falhas variou bastante, como registrado na Tabela 5.7. Isso pode ter ocorrido porque, por exemplo, no período de testagem de determinadas amostras, a Internet tenha funcionado bem.

Tabela 5.7. Médias das falhas relacionadas ao Herbário Virtual Refflora.

Tamanho da amostra	Média das falhas
5.000	11,6
7.500	80,6
10.000	35,6
12.500	184,3
15.000	12,6
17.500	214,3
20.000	110
22.500	441,6
25.000	172,6

Para a realização dos testes da integração do HCF-*Web* ao sistema speciesLink, foi feito o *download* do arquivo .TXT, que continha dados das 28.173 exsicatas disponibilizadas pelo HCF no dia 10 de maio de 2019. O padrão dos resultados, das médias dos tempos de execução em relação aos tamanhos das amostras, foi semelhante ao obtido com o Reflora, ou seja, os tempos de atualização aumentaram à medida em que amostras maiores foram executadas. Dessa forma, a média do tempo de execução de 5 mil exsicatas foi de cerca de 17 minutos, e a de 25 mil, de 64 minutos. As médias dos tempos de atualização de cada amostra do HCF-*Web*, em relação aos dados equivalentes disponíveis no speciesLink, podem ser vistas na Figura 5.13.

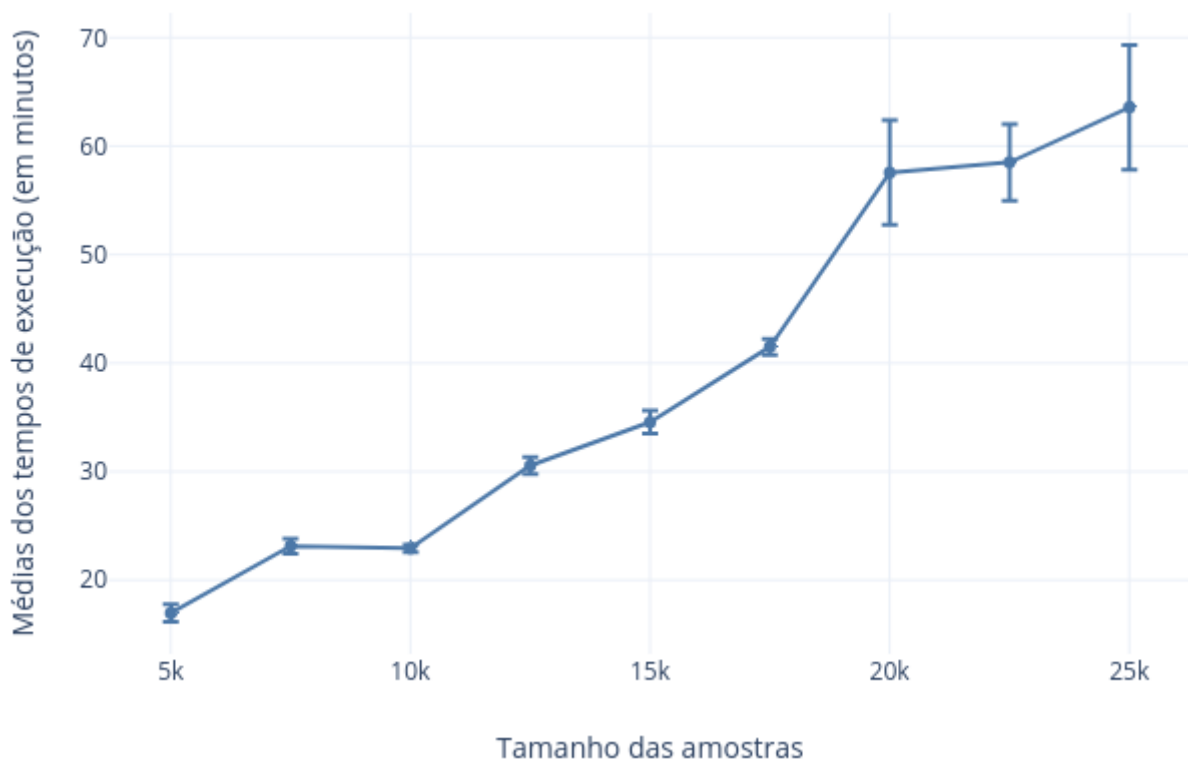


Figura 5.13. Médias e desvios-padrões dos tempos de atualização de cada amostra do HCF-*Web*, em relação aos dados equivalentes no speciesLink.

Cabe registrar a grande diferença entre os tempos de atualização do HCF-*Web*, em relação aos dados disponíveis no Reflora e no speciesLink, ou seja, neste, como não é necessária a utilização da Internet, a execução foi muito mais rápida do que naquele. Por exemplo, a média do tempo de atualização das 5.000 amostras, relacionada ao Reflora, foi de cerca de 561 minutos, e ao speciesLink, de apenas 17 minutos.

Nas amostras maiores - de 20 mil, 22,5 mil e 25 mil exsicatas - o grau de variação das médias dos tempos de execução foi grande, em relação às amostras menores. Por exemplo, na execução das 5 mil exsicatas, o desvio-padrão foi de, aproximadamente, 0,8 minutos, e das 25 mil exsicatas, de 5,7 minutos. É possível que nas três maiores amostras, durante uma ou mais execuções, outro processo estivesse sendo, simultaneamente, realizado pela máquina.

Isso explicaria, por exemplo, que na primeira execução das 22,5 mil exsicatas, o tempo tenha sido de, aproximadamente, 63 minutos, e na segunda e terceira execuções, respectivamente, de 57 minutos e 55 minutos.

Em relação às médias de divergências relacionada ao speciesLink, apresentadas na Tabela 5.8, os valores encontrados foram expressivos, assim como ocorrido com o Refflora.

Tabela 5.8. Médias das divergências relacionadas ao speciesLink.

Tamanho da amostra	Média das divergências
5.000	817,6
7.500	1220
10.000	1645,3
12.500	2328
15.000	2423,6
17.500	2855,3
20.000	3295,6
22.500	3654,3
25.000	4211,3

A seguir, são apresentadas as conclusões referentes à integração do HCF-*Web* ao Herbário Virtual Refflora e ao INCT-HVFF.

Conclusão

O HCF-*Web* disponibiliza o seu acervo no Herbário Virtual Re flora e no INCT-HVFF, que oferecem uma importante funcionalidade: permitem que especialistas cadastrados realizem sugestões de reidentificações das amostras das coleções. Entretanto, até o ano de 2018, as reidentificações das plantas do HCF eram registradas somente nesses herbários virtuais. Para acabar com as inconsistências entre os dados das exsicatas do HCF e os seus equivalentes, no Re flora e no INCT-HVFF, era preciso que a atualização fosse realizada manualmente. Considerando que o acervo do HCF é composto por mais de 28 mil exsicatas, a atividade manual é praticamente inviável, pois demanda tempo e está sujeita a erros. Assim, neste trabalho foi implementada uma extensão para o HCF-*Web*, que permite atualizar seus dados, comparando-os com os equivalentes, nos herbários virtuais, a fim de localizar divergências.

Os 27 testes realizados, para cada um dos novos serviços, demonstraram que a identificação das inconsistências do HCF-*Web*, em relação aos dados do sistema speciesLink, do INCT-HVFF, ocorre mais rapidamente que aos do Herbário Virtual Re flora. Isso se deve ao fato de o speciesLink exigir somente o consumo de um arquivo TXT fornecido por ele, para localizar divergências. Por outro lado, no Herbário Virtual Re flora é preciso utilizar a sua API para solicitar as informações presentes nele para, então, comparar esses dados com os do HCF-*Web*. Apesar de a comparação ser totalmente automática e, portanto, não requerer que o usuário realize *download* para encontrar divergências, o processo no Re flora é mais lento que no speciesLink, pois problemas com o servidor ou de conexão podem ocorrer.

A extensão para o HCF-*Web*, implementada neste estudo: (i) contribuiu para a resolução do problema de inconsistência, entre os dados presentes no Re flora e no INCT-HVFF e os seus equivalentes, no HCF-*Web*; (ii) aumentou a precisão, na identificação de divergências; (iii) e ajudou a otimizar o trabalho dos funcionários do HCF, pois a comparação manual dos dados, que demorava semanas, pode ser realizada, agora, de forma automática e em poucos dias. Dessa forma, as novas funcionalidades do HCF-*Web* ajudaram a resolver o problema da

comparação manual dos dados, que demandava tempo e estava sujeita a erros.

A integração do HCF-*Web* ao Re flora e ao INCT-HCFF concorrerá para a melhoria da qualidade dos dados da coleção do HCF, pois o seu curador, a partir de agora, tem acesso facilitado às sugestões de identificações e reidentificações das plantas, realizadas por especialistas brasileiros e estrangeiros nos dois herbários virtuais. Com a extensão implementada, o desafio de atualizar informações e corrigir erros referentes aos nomes de famílias, gêneros e espécies dos espécimes do HCF poderá ser enfrentado de forma mais ágil, o que contribuirá, certamente, para aprimorar e qualificar, ainda mais, as suas amostras.

Em trabalhos futuros poderão ser adicionadas novas funcionalidades ao HCF-*Web*, como um *dashboard* que exiba informações quantitativas sobre o acervo: (i) espécimes determinados e indeterminados; (ii) famílias, gêneros e espécies das plantas; (iii) pendências existentes; (iv) e locais de coleta de materiais mais frequentes. Além dessa funcionalidade, também poderá ser utilizada uma API, para exibir em um mapa, a partir das coordenadas geográficas fornecidas no registro da amostra, o local de coleta de um espécime. Outra sugestão de trabalho futuro é o desenvolvimento de uma aplicação móvel do HCF-*Web*.

Referências

- AMLER, Vipul; SONPATKI, Prathamesh. *ReactJS by example - Building modern web applications with React*. 1st. ed. Birmingham: Packt Publishing, 2016. ISBN 978-1-78528-964-4.
- ANTONIO, Cassio de Sousa. *Pro react*. New York: Apress, 2015. ISBN 978-1-4842-1260-8.
- ASSOCIATION FOR COMPUTING MACHINERY. *A. M. Turing Award. Sir Tim Berners-Lee*. 2016. Disponível em: <https://amturing.acm.org/award/_winners/berners-lee/_8087960.cfm>.
- BANKS, Alex; PORCELLO, Eve. *Learning react: Functional web development with react and redux*. 1st. ed. Sebastopol: O'Reilly Media, 2017. ISBN 978-1-491-95462-1.
- BRAGA-NETO, Ricardo. *Herbários contribuem para o desenvolvimento da e-taxonomia no Brasil*. 2013. Disponível em: <<http://inct.florabrasil.net/herbarios-contribuem-para-o-desenvolvimento-da-e-taxonomia-no-brasil/>>.
- CARNEIRO, João Luis. *Desenvolvimento de sistema de computador para utilização no herbário da Universidade Tecnológica Federal do Paraná - Câmpus Campo Mourão (HCF)*. 2006. 24 f. Monografia (Tecnologia em Gestão Ambiental) — Universidade Tecnológica Federal do Paraná, Campo Mourão, 2006.
- CENTRO DE REFERÊNCIA EM INFORMAÇÃO AMBIENTAL. *Sistema de informação de coleções de interesse biotecnológico - Manual do usuário*. 2009. Disponível em: <http://sicol.cria.org.br/gestao/manual/_v132/_20091008.pdf>.
- CENTRO DE REFERÊNCIA EM INFORMAÇÃO AMBIENTAL. *speciesBase - Manual de utilização*. 2012. Disponível em: <<http://splink.cria.org.br/files/speciesBase/lastversion/Manual.pdf>>.
- CENTRO DE REFERÊNCIA EM INFORMAÇÃO AMBIENTAL. *Sobre o CRIA*. 2018a. Disponível em: <<http://www.cria.org.br/about/>>.
- CENTRO DE REFERÊNCIA EM INFORMAÇÃO AMBIENTAL. *Projetos - Sistemas de informação*. 2018b. Disponível em: <<http://www.cria.org.br/projetos>>.
- CENTRO DE REFERÊNCIA EM INFORMAÇÃO AMBIENTAL. *Setembro verde 2018*. 2018c. Disponível em: <<http://blog.cria.org.br/2018/09/setembro-verde-2018.html>>.
- CENTRO DE REFERÊNCIA EM INFORMAÇÃO AMBIENTAL. *speciesLink no CDC da RNP. Blog do CRIA*, Campinas, 2019. 30 maio. *Link para "Documento síntese sobre a rede speciesLink, uma rede de informação colaborativa sobre a biodiversidade brasileira"*. Disponível em: <<http://blog.cria.org.br/>>.

CERQUEIRA, Rui; ANTONINI, Yasmine; FERNANDES, Rosan Valter; FONTENELLE, Julio Cesar Rodrigues. Glossário. In: RAMBALDI, Denise Marçal; OLIVEIRA, Daniela América Suárez de (Ed.). *Fragmentação de ecossistemas: causas, efeitos sobre a biodiversidade e recomendações de políticas públicas*. Brasília: Ministério do Meio Ambiente, 2005. p. 486–508. ISBN 87166-48-4. Disponível em: <<https://www2.ib.unicamp.br/profs/fsantos/refer/Glossario.pdf>>.

DOGLIO, Fernando. *Pro REST API development with Node.js*. New York: Apress, 2015. ISBN 978-1-4842-0917-2.

ELMASRI, Ramcz; NAVATHE, Shamkant B. *Sistemas de banco de dados*. 6th. ed. [S.l.]: Pearson Addison Wcsley, 2011. ISBN 978-85-7936-085-5.

EMBARCADERO. *Delphi - Visão geral*. 2018a. Disponível em: <<https://www.embarcadero.com/br/products/delphi>>.

EMBARCADERO. *InterBase® - Overview*. 2018b. Disponível em: <<https://www.embarcadero.com/br/products/interbase>>.

FIELDING, R.; MOGUL, J.; FRYSTYK, H.; MASINTER, L.; LEACH, P.; BERNERS-LEE, T. *Hypertext Transfer Protocol – HTTP/1.1*. 1999. Disponível em: <<https://www.ietf.org/rfc/rfc2616.txt>>.

FREEMAN, Eric Thomas; ROBSON, Elisabeth. *Head first JavaScript programming*. Sebastopol: O’Reilly Media, 2014. ISBN 978-1-449-34013-1.

GADELHA JUNIOR, Luiz M. R.; GUIMARÃES, Pedro; MOURA, Ana Maria; DRUCKER, Debora P.; DALCIN, Eduardo; GALL, Guilherme; TAVARES JUNIOR, Jurandir; PALAZZI, Daniele; POLTOSI, Maira; PORTO, Fabio; MOURA, Francisco; LEO, Wagner Vieira. Sibbr: uma infraestrutura para coleta, integração e análise de dados sobre a biodiversidade brasileira. In: XXXIV CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO – CSBC. [S.l.], 2014.

GLOBAL BIODIVERSITY INFORMATION FACILITY. *IPT: The Integrated Publishing Toolkit*. 2018a. Disponível em: <<https://www.gbif.org/ipt>>.

GLOBAL BIODIVERSITY INFORMATION FACILITY. *HCF - Herbário da Universidade Tecnológica Federal do Paraná - Campus Campo Mourão*. 2018b. Disponível em: <<https://www.gbif.org/dataset/dd31ac22-63d4-4bb4-a987-5eece5ffadb8>>.

GONZALEZ, David. *Developing microservices with Node.js*. 1st. ed. [S.l.]: Packt Publishing, 2016. ISBN 978-1-78588-740-6.

GULGEMIN, Felipe. *BrOffice muda nome oficial para LibreOffice*. 2011. Disponível em: <<https://www.tecmundo.com.br/office/9195-broffice-muda-nome-oficial-para-libreoffice.htm>>.

HERMAN, David. *Effective JavaScript: 68 specific ways to harness the power of JavaScript*. Boston: Addison-Wesley Professional, 2013. ISBN 978-0-321-81218-6.

HUGHES-CROUCHER, Tom; WILSON, Mike. *Node: Up and running*. Sebastopol: O’Reilly Media, 2012. ISBN 978-1-449-39858-3.

INCT - HERBÁRIO VIRTUAL DA FLORA E DOS FUNGOS. *INCT - Herbário Virtual da Flora e dos Fungos – REFLORA e o Serviço Exsiccatae: imagens online de espécimes*. 2018. Disponível em: <<http://inct.florabrasil.net/reflora/>>.

LÍRIO, Elton John de; SARNAGLIA JÚNIOR, Valderes Bento. *E se as plantas não tivessem nome? A importância do Código de Nomenclatura Botânica*. 2014. Disponível em: <<http://www.sambio.org.br/simbioma/simbioma/%20iii/05.pdf>>.

MAIA, Leonor Costa. *CNPq viabiliza o repatriamento de dados históricos da flora brasileira*. 2010. Disponível em: <<http://inct.florabrasil.net/cnpq-viabiliza-o-repatriamento-de-dados-historicos-da-flora-brasileira/>>.

MAIA, Leonor Costa; VIEIRA, Ana Odete Santos; PEIXOTO, Ariane Luna; STEHMANN, João Renato; BARBOSA, Maria Regina de Vasconcellos; MENEZES, Mariângela. *Construindo redes para promover o conhecimento da biodiversidade brasileira: a experiência do INCT - Herbário Virtual*. Recife: Editora Universitária da UFPE, 2017. ISBN 978-85-415-0901-5.

MARDAN, Azat. *Full stack JavaScript: Learn Backbone.js, Node.js and MongoDB*. New York: Apress, 2015. ISBN 978-1-4842-1751-1.

MARDAN, Azat. *React quickly: Painless web apps with react, JSX, redux, and GraphQL*. Shelter Island: Manning Publications, 2017. ISBN 978-1617293344.

MARTINS, Vera Lúcia Campos; VIANA, Regina Célia; SILVA, Martins da; REIS, Josiane dos Santos; SILVA, Maria Jociléia Soares da. *Tipos nomenclaturais de Angiospermae do Herbário IAN da Embrapa Amazônia Oriental: Aristolochiaceae, Asclepiadaceae, Bombacaceae, Compositae, Gesneriaceae, Lecythidaceae, Meliaceae, Moraceae, Rubiaceae e Sapotaceae*. 2006. Disponível em: <<https://ainfo.cnptia.embrapa.br/digital/bitstream/item/28269/1/Doc273.pdf>>.

MCNEILL, J.; BARRIE, F. R.; BUCK, W. R.; DEMOULIN, V.; GREUTER, W.; HAWKSWORTH, D. L.; HERENDEEN, P. S.; KNAPP, K. Marhold S.; PRADO, J.; PRUD'HOMME VAN REINE, W. F.; SMITH, G. F.; WIERSEMA, J. H. *Código Internacional para Nomenclatura de Algas, Fungos e Plantas - Código de Melbourne*. São Paulo: Rima, 2013. ISBN 978-8576562900.

MEHTA, Vijay P. *Pro LINQ object relational mapping with C 2008*. New York: Apress, 2008. ISBN 978-1-4302-0597-5.

MYSQL. *MySQL*. 2018. Disponível em: <<https://www.mysql.com/>>.

NACIONAL CENTER FOR SUPERCOMPUTING APPLICATIONS. *NCSA Mosaic™*. 2018. Disponível em: <<http://www.ncsa.illinois.edu/enabling/mosaic>>.

NODE.JS. *The Node.js event loop, timers, and process.nextTick()*. 2018. Disponível em: <<https://nodejs.org/en/docs/guides/event-loop-timers-and-nexttick/>>.

PEIXOTO, Ariane Luna; BARBOSA, Maria Regina de V.; MENEZES, Mariângela; MAIA, Leonor Costa. *Diretrizes e estratégias para a modernização de coleções biológicas brasileiras e a consolidação de sistemas integrados de informação sobre biodiversidade*. Brasília: Ministério da Ciência e Tecnologia, 2006.

PEIXOTO, Ariane Luna; MAIA, Leonor Costa. *Manual de procedimentos para herbários*. Recife: Editora Universitária UFPE, 2013.

PEIXOTO, Flávio Luna. *O processo de informatização de herbários: estudo de caso*. 2005. 69 f. Monografia (Mestrado em Botânica) — Instituto de Pesquisas Jardim Botânico do Rio de Janeiro/Escola Nacional de Botânica Tropical, Rio de Janeiro, 2005.

PIRES, Paulo de Tarso de Lara; LOPER, Ailson Augusto; MENDES, Carlos José; PETERS, Edson Luiz; MAIA, Gabriela Nicolau; ABREU, Lucas Moura de. *Dicionário de termos florestais*. Curitiba: FUPEF, 2018. ISBN 978-85-60020-10-2. Disponível em: <http://www.apreflorestas.com.br/wp-content/uploads/2018/03/APRE_dicionario_2018_digital.pdf>.

POLISELI, Ana Carla. Herbário da UTFPR tem mais de 12 mil plantas. *Tribuna do Interior*, Campo Mourão, 2013. 13 abr. Disponível em: <<https://www.tribunadointerior.com.br/noticia/herbario-da-utfpr-Interior>>.

REACT. *Reconciliation*. 2018. Disponível em: <<https://reactjs.org/docs/reconciliation.html>>.

REFLORA. *Herbário Virtual*. 2018a. Disponível em: <<http://reflora.jbrj.gov.br/reflora/herbarioVirtual/ConsultaPublicoHVUC/ConsultaPublicoHVUC.do>>.

REFLORA. *Estatísticas do Herbário Virtual. Exsicatas. Número total de exsicatas por herbário*. 2018b. Disponível em: <<http://reflora.jbrj.gov.br/reflora/herbarioVirtual/ExibeEstatisticaHVUC/ExibeEstatisticaHVUC.do?modoOperacaoExibirEstatisticaHV=quantidadeExsicatas>>>

SANGALI, Elaine Cecilio. *HCF-Web: um sistema web para o Herbário HCF da UTFPR-CM*. 2019. 41 f. Monografia (Bacharelado em Ciência da Computação) — Universidade Tecnológica Federal do Paraná, Campo Mourão, 2019.

SATHEESH, Mithun; D'MELLO, Bruno Joseph; KROL, Jason. *Web development with MongoDB and NodeJS*. 2nd. ed. Birmingham: Packt Publishing, 2015. ISBN 978-1-78528-752-7.

SILVA, Luís Alexandre Estevão da; ROCHA, Mônica Sousa da; BELLON, Ernani; RIBEIRO, Rafael da Silva; OLIVEIRA, Felipe Alves de; CLEMENTE, Leonardo da Silva; MAGDALENA, Ulises Rodrigo; MEDEIROS, Erika von Sohsten; FORZZA, Rafaela Campostrini. Jabot - sistema de gerenciamento de coleções botânicas: a experiência de uma década de desenvolvimento e avanços. *Rodriguésia*, Rio de Janeiro, v. 68, n. 2, p. 391–410, 2017. ISSN 0370-6583.

SISTEMA DE INFORMAÇÃO SOBRE A BIODIVERSIDADE BRASILEIRA. *Termos Darwin Core*. 2018a. Disponível em: <<http://sibbr.gov.br/areas/index.php?area=publicar&subarea=termos-dwc>>.

SISTEMA DE INFORMAÇÃO SOBRE A BIODIVERSIDADE BRASILEIRA. *O que é o SiBBr*. 2018b. Disponível em: <<http://www.sibbr.gov.br/internal/?area=osibbr>>.

SISTEMA DE INFORMAÇÃO SOBRE A BIODIVERSIDADE BRASILEIRA. *Softwares para gestão de coleções*. 2018c. Disponível em: <<http://www.sibbr.gov.br/areas/index.php?area=colecoes&subarea=softwares-para-gestao-de-colecoes>>.

SPECIESLINK. *HCF - Herbário da Universidade Tecnológica Federal do Paraná - Câmpus Campo Mourão*. 2018a. Disponível em: <<http://splink.cria.org.br/manager/detail?resource=HCF>>.

SPECIESLINK. *Resumo e projeto de pesquisa*. 2018b. Disponível em: <<http://splink.cria.org.br/description?criaLANG=pt>>.

SPECIESLINK. *Coleções participantes*. 2018c. Disponível em: <<http://splink.cria.org.br/manager/index?setlang=pt>>.

SPECIESLINK. *O projeto: como participar*. 2018d. Disponível em: <http://splink.cria.org.br/how_to_participate?criaLANG=pt>.

SPECIESLINK. *Dados e ferramentas*. 2018e. Disponível em: <<http://splink.cria.org.br/tools?criaLANG=pt>>.

SPECIESLINK. *SpLinker*. 2018f. Disponível em: <<http://splink.cria.org.br/splinker?criaLANG=pt>>.

SPECIESLINK. *Programas para gerenciamento de coleções*. 2018g. Disponível em: <<http://splink.cria.org.br/specify?criaLANG=pt>>.

SPECIFY. *About Specify*. 2018. Disponível em: <<http://www.sustain.specifysoftware.org/about/>>.

TANENBAUM, Andrew S.; WETHERALL, David. *Redes de computadores*. São Paulo: Pearson Prentice Hall, 2013. ISBN 978-85-7605-924-0.

TEIXEIRA, Pedro. *Professional Node.js: Building Javascript based scalable software*. Birmingham: Wrox, 2012. ISBN 978-1-118-18546-9.

UNIVERSIDADE ESTADUAL DO RIO DE JANEIRO. *Herbários na UERJ - Câmpus Maracanã*. 2009. Disponível em: <<http://www.ibrag.uerj.br/index.php/noticias/289-herbarios-na-uerj-campus-maracana.html>>.

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ. *Herbário - Universidade Tecnológica Federal do Paraná*. 2015. Disponível em: <<http://www.utfpr.edu.br/campomourao/o-campus/herbario>>.

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ. *Herbário do câmpus Campo Mourão é a sexta maior coleção botânica do Paraná*. 2017. Disponível em: <<http://portal.utfpr.edu.br/noticias/reitoria/herbario-do-campus-campo-mourao-e-a-sexta-maior-colecao-botanica-do-parana>>.

UNIVERSITY OF KANSAS. *Specify - Welcome to the Specify 6.5 help*. 2013. Disponível em: <http://www.sustain.specifysoftware.org/wp-content/uploads/2017/03/complete_specify_help.pdf>.

UNIVERSITY OF OXFORD. *BRAHMS Project. BRAHMS 7 documentation*. 2013. Disponível em: <<https://herbaria.plants.ox.ac.uk/bol/brahms/support/documentation>>.

UNIVERSITY OF OXFORD. *Oxford University Herbaria. BRAHMS project: Botanical Research and Herbarium Management System*. 2018a. Disponível em: <<https://herbaria.plants.ox.ac.uk/bol/oxford/brahms>>.

UNIVERSITY OF OXFORD. *BRAHMS Database. BRAHMS software*. 2018b. Disponível em: <<https://herbaria.plants.ox.ac.uk/bol/brahms/softwarev7>>.

VALOIS, Afonso Celso Candeira; SALOMÃO, Antonieta Nassif; ALLEM, Antonio Costa. *Glossário de recursos genéticos vegetais*. Brasília: EMBRAPA, 1999. ISBN 0102-0110.

WILSON, Jim R. *Node.js the right way: Practical, server-side JavaScript that scales*. Dallas: Pragmatic Bookshelf, 2013. ISBN 978-1-937785-73-4.

YOUNG, Alex R.; HARTER, Marc. *Node.js in practice*. Shelter Island: Manning Publications, 2014. ISBN 978-1617290930.

Apêndices

Atributos das relações do HCF-*Web*

Nas tabelas abaixo, são apresentadas as 31 relações do HCF *Web*, e descritos apenas os atributos cuja redação não é clara o suficiente para identificá-los.

Tabela A.1. Atributos da relação *tombo*s.

Atributos	Descrição
hcf	-
data_tombo	-
data_coleta_dia	-
observacao	-
nomes_populares	-
numero_coleta	-
latitude	-
longitude	-
altitude	-
entidade_id	Herbário onde está depositado o espécime relacionado a um determinado tombo.
local_coleta_id	-
variedade_id	-
tipo_id	-

Tabela A.2. Atributos da relação `tombos`. (continuação)

Atributos	Descrição
<code>situacao</code>	Indica a situação de uma amostra: regular, isto é, no acervo do herbário; doada; emprestada; ou permutada.
<code>especie_id</code>	-
<code>genero_id</code>	-
<code>familia_id</code>	-
<code>sub_familia_id</code>	-
<code>sub_especie_id</code>	-
<code>nome_cientifico</code>	-
<code>colecacao_anexa_id</code>	-
<code>cor</code>	As cores vermelho, azul e verde indicam se a amostra foi coletada, respectivamente, no Paraná, Brasil ou exterior.
<code>data_coleta_mes</code>	-
<code>data_coleta_ano</code>	-
<code>created_at</code>	-
<code>updated_at</code>	-
<code>ativo</code>	-
<code>taxon</code>	-
<code>rascunho</code>	Usado quando o tombo ainda não foi aprovado.

Tabela A.3. Atributos da relação `tombos_fotos`.

Atributos	Descrição
<code>id</code>	-
<code>tombo_hcf</code>	-
<code>codigo_barra</code>	-
<code>num_barra</code>	-
<code>caminho_foto</code>	-
<code>em_vivo</code>	Indica que o espécime não passou pelo processo de prensagem e secagem.
<code>created_at</code>	-
<code>updated_at</code>	-
<code>sequencia</code>	Sequência de fotos que correspondem a um tombo.
<code>ativo</code>	-

Tabela A.4. Atributos da relação `colecoes_anexas`.

Atributos	Descrição
<code>tipo</code>	Outras coleções botânicas, que não sejam de plantas e fungos, por exemplo, carpoteca (coleção de frutos) e xiloteca (coleção de pedaços de madeira desidratados).
<code>observacoes</code>	-
<code>id</code>	-
<code>created_at</code>	-
<code>updated_at</code>	-

Tabela A.5. Atributos da relação `alteracoes`.

Atributos	Descrição
<code>id</code>	-
<code>usuario_id</code>	-
<code>status</code>	-
<code>observacao</code>	Observação feita pelo curador.
<code>ativo</code>	-
<code>created_at</code>	-
<code>updated_at</code>	-
<code>tombo_hcf</code>	-
<code>tombo_json</code>	Sugestão feita por um usuário.
<code>identificacao</code>	Utilizado quando os táxons do espécime foram determinados (identificados).
<code>data_identificacao_dia</code>	-
<code>data_identificacao_mes</code>	-
<code>data_identificacao_ano</code>	-

Tabela A.6. Atributos da relação `tombo_alteracoes_antigas`.

Atributos	Descrição
<code>sequencia</code>	-
<code>data</code>	-
<code>descricao</code>	Alterações antigas, realizadas nos tomos.
<code>tombo_hcf</code>	-

Tabela A.7. Atributos da relação `retirada_exsiccta_tombos`.

Atributos	Descrição
<code>retirada_exsiccta_id</code>	-
<code>tombo_hcf</code>	-
<code>tipo</code>	Indica se a exsicata foi emprestada, permutada ou doada.
<code>data_vencimento</code>	-
<code>created_at</code>	-
<code>updated_at</code>	-
<code>devolvido</code>	-

Tabela A.8. Atributos da relação `remessas`.

Atributos	Descrição
<code>id</code>	-
<code>observacao</code>	Anotações sobre permutas, doações e empréstimos realizados.
<code>data_envio</code>	-
<code>entidade_destino_id</code>	-
<code>herbario_id</code>	-
<code>created_at</code>	-
<code>updated_at</code>	-

Tabela A.9. Atributos da relação `usuarios`.

Atributos	Descrição
<code>id</code>	-
<code>nome</code>	-
<code>ra</code>	-
<code>email</code>	-
<code>senha</code>	-
<code>ativo</code>	-
<code>tipo_usuario_id</code>	-
<code>telefone</code>	-
<code>herbario_id</code>	-
<code>created_at</code>	-
<code>updated_at</code>	-

Tabela A.10. Atributos da relação `tipos_usuarios`.

Atributos	Descrição
<code>id</code>	-
<code>tipo</code>	-
<code>created_at</code>	-
<code>updated_at</code>	-

Tabela A.11. Atributos da relação `historico_acessos`.

Atributos	Descrição
<code>id</code>	-
<code>data_criacao</code>	-
<code>usuario_id</code>	-

Tabela A.12. Atributos da relação `tombos_coletores`.

Atributos	Descrição
<code>tombo_hcf</code>	-
<code>coletor_id</code>	-
<code>created_at</code>	-
<code>updated_at</code>	-
<code>principal</code>	1 indica que a amostra foi coletada pelo coletor principal; e 0 , por um coletor secundário.

Tabela A.13. Atributos da relação `coletores`.

Atributos	Descrição
<code>id</code>	-
<code>nome</code>	-
<code>email</code>	-
<code>numero</code>	Número de identificação do coletor.
<code>data_criacao</code>	-
<code>ativo</code>	-
<code>created_at</code>	-
<code>updated_at</code>	-

Tabela A.14. Atributos da relação locais_coleta.

Atributos	Descrição
id	-
descricao	-
solo_id	-
relevo_id	-
vegetacao_id	-
cidade_id	-
fase_sucessional_id	-
complemento	Anotações complementares do local da coleta.
created_at	-
updated_at	-
fase_numero	-

Tabela A.15. Atributos da relação **estados**.

Atributos	Descrição
id	-
nome	-
codigo_telefone	-
pais_id	-
created_at	-
updated_at	-

Tabela A.16. Atributos da relação **países**.

Atributos	Descrição
id	-
nome	-
sigla	-
created_at	-
updated_at	-

Tabela A.17. Atributos da relação `solos`.

Atributos	Descrição
<code>id</code>	-
<code>nome</code>	-
<code>created_at</code>	-
<code>updated_at</code>	-

Tabela A.18. Atributos da relação `relevos`.

Atributos	Descrição
<code>id</code>	-
<code>nome</code>	-
<code>created_at</code>	-
<code>updated_at</code>	-

Tabela A.19. Atributos da relação `fase_sucessional`.

Atributos	Descrição
<code>numero</code>	-
<code>nome</code>	Estágio sucessional da amostra coletada, em relação à formação florestal.
<code>created_at</code>	-
<code>updated_at</code>	-

Tabela A.20. Atributos da relação *vegetacoes*.

Atributo	Descrição
id	-
nome	-
created_at	-
updated_at	-

Tabela A.21. Atributos da relação *autores*.

Atributos	Descrição
id	-
nome	Primeira pessoa que descreveu a espécie.
iniciais	-
created_at	-
updated_at	-
ativo	-
nomeAntigo	-

Tabela A.22. Atributos da relação *generos*.

Atributos	Descrição
id	-
nome	-
familia_id	-
created_at	-
updated_at	-
ativo	-

Tabela A.23. Atributos da relação *familias*.

Atributos	Descrição
id	-
nome	-
created_at	-
updated_at	-
ativo	-

Tabela A.24. Atributos da relação *sub_familias*.

Atributos	Descrição
id	-
nome	-
familia_id	-
autor_id	-
created_at	-
updated_at	-
ativo	-

Tabela A.25. Atributos da relação *especies*.

Atributos	Descrição
id	-
nome	-
autor_id	-
genero_id	-
familia_id	-
created_at	-
updated_at	-
ativo	-

Tabela A.26. Atributos da relação *sub_especies*.

Atributos	Descrição
id	-
nome	-
especie_id	-
genero_id	-
familia_id	-
autor_id	-
created_at	-
updated_at	-
ativo	-

Tabela A.27. Atributos da relação variedades.

Atributos	Descrição
id	-
nome	Nome da variedade a que pertence o espécime.
autor_id	-
especie_id	-
genero_id	-
familia_id	-
created_at	-
updated_at	-
ativo	-

Tabela A.28. Atributos da relação tipos.

Atributos	Descrição
id	-
nome	Identificação do tipo nomenclatural do material coletado. Por exemplo, a amostra pode ser do tipo holótipo, parátipo, isótipo etc.
created_at	-
updated_at	-

Tabela A.29. Atributos da relação herbarios.

Atributos	Descrição
id	-
nome	-
caminho_logotipo	-
sigla	-
email	-
ativo	-
endereco_id	-
created_at	-
updated_at	-

Tabela A.30. Atributos da relação telefones.

Atributos	Descrição
id	-
numero	-
herbario_id	-
ativo	-

Tabela A.31. Atributos da relação `enderecos`.

Atributos	Descrição
<code>id</code>	-
<code>logradouro</code>	Corresponde ao endereço.
<code>numero</code>	-
<code>complemento</code>	Corresponde a um ponto de referência.
<code>cidade_id</code>	-
<code>created_at</code>	-
<code>updated_at</code>	-