The Encounter of Matthew Flinders and Nicolas Baudin in 1802, by John Ford

Atlas of Living Australia

Documento Técnico Chave

Versão: 1.0.4, 6 Julho 2016

Este documento é uma contribuição do projeto Encounter Bay (ALA node portal internationalization), co-financiado pelos Nós Francês, Espanhol, Português e Belga do GBIF e pelo Secretariado do GBIF através do seu Programa de Suporte ao Aumento de Capacidade.

Autores (por ordem alfabética)

Fabien Cavière, GBIF France, caviere@gbif.fr
Rui Figueira, GBIF Portugal, ruifigueira@isa.ulisboa.pt
André Heughebaert, GBIF Belgium, a.heughebaert@biodiversity.be
Marie-Elise Lecoq, GBIF France, melecoq@gbif.fr
Santiago Martínez de la Riva, GBIF Spain, sama@gbif.es

Tradução: Rui Figueira, GBIF Portugal, ruifigueira@isa.ulisboa.pt

Audiência

Especialistas IT que queiram obter uma primeira compreensão da tecnologia por detrás do portal de dados de biodiversidade Atlas of Living Australia (ALA). É esperado que o leitor deste documento tenha pelo menos um conhecimento médio sobre serviços web, bases de dados e desenvolvimento web em geral.

Objetivos

Este documento destina-se a captar a essência da documentação técnica sobre os componentes do ALA necessários para que os avaliadores/adotantes possam descobrir, avaliar, compreender, instalar, configurar, ajustar e operar as ferramentas utilizadas.

Licença

Este documento é público sob a licença CC-BY-SA 4.0.

Informação de contacto/limitação de responsabilidade

Este documento foi elaborado pelos primeiros a adotar a tecnologia ALA (os nós franceses, portugueses e espanhóis do GBIF), e não pela equipa de desenvolvimento australiana. Assim, a maioria das capturas de ecrã e exemplos não se referem ao portal Atlas of Living Australia, mas aos portais dos primeiros nós a adotar o sistema.

Para mais informação contactar info@gbif.es.

Apoio a este documento pelo Atlas of Living Australia

O Atlas of Living Australia (ALA) tem o prazer de apoiar este documento, que é o resultado do esforço de colaboração de vários Nós GBIF. Consideramos este documento uma instrodução útil às ferramentas para qualquer Nó nacional, ou mesmo para qualquer grupo interessado na criação de um portal temático. Vemos esta colaboração como a primeira de muitas atividades conjuntas em torno das ferramentas do Atlas.

Índice

1 Introdução..... 5
1.1 Como usar este documento..... 5
1.2 Portais do ALA derivados..... 6
2 Arquitetura..... 7
2.1 Fundamentação do design de software ALA..... 7
2.2 Visão geral dos módulos..... 9
2.3 Pesquisa de ocorrências..... 10
2.4 Collectory..... 11
2.4.1 Páginas públicas..... 11
2.4.2 Páginas privadas..... 12
2.5 Página de espécies..... 12
2.6 Hub de dados..... 13
2.7 Regiões..... 17
2.8 Portal Espacial..... 18
2.9 Princípios técnicos básicos..... 20
3 Requisitos e instalação..... 20
3.1 Requisitos do servidor de teste..... 21
3.2 Requisitos do servidor de produção..... 22
3.3 Outros exemplos de Arquitetura do Sistema..... 24
3.3.1 Austrália..... 25
3.3.2 França..... 26
3.4 Processo de instalação(Ambiente de teste)..... 26
3.4.1 Vagrant..... 27
3.4.2 Ansible..... 28
3.4.3 Terminar a MV..... 30
3.5 Processo de Instalação em servidores Remotos (ambiente de Produção) 31
3.5.1 Compreender o conceito..... 31
3.5.2 Passos de instalação..... 31
3.5.3 Como instalar os serviços em diferentes servidores utilizando playbooks Ansible..... 32
4 Configuração..... 33
4.1 Github..... 33
4.2 Armazenamento..... 34
4.2.1 MySQL..... 34
4.2.2 Cassandra..... 34
4.3 Arquitetura dos ficheiros..... 35
4.4 Ficheiro de configuração..... 36
4.5 Definindo o estilo da aplicação..... 37
4.6 Internacionalização..... 41
5 Carregamento de dados..... 42
5.1 Tipos de dados..... 43
5.1.1 Coleção..... 44

5.1.2 Provedor de dados....44
5.1.3 Instituição....44
5.1.4 Hub de dados.... 44
5.1.5 Contacto.... 44
5.2 Dados de configuração.... 45
5.2.1 Código do Provedor....45
5.2.2 Mapa de Provedor....45
5.3 Recurso de dados.... 46
5.3.1 Recurso de dados externo....47
5.3.2 Recursos GBIF.... 47
5.3.3 Usando o IPT.... 47
6 Cópia de segurança e atualização.... 48
6.1 Fazer cópia de segurança das bases de dados....48
6.1.1 Cópia de segurança da base de dados MySQL....48
6.1.2 Cópia de segurança dos dados de Cassandra....49
6.1.3 Fazer um snapshot.... 49
6.1.4 Restaurar um snapshot.... 49
6.1.5 Cópia de segurança do armazenamento de dados SOLR....50
6.2 Ferramentas de atualização....52
7 Funcionalidades.... 53
7.1 Visualização de dados.... 53
7.1.1 Pesquisar ocorrências....55
7.1.2 Pesquisar metadados.... 58
7.2 Funcionalidades de Importação....65
7.3 Funcionalidades de Exportação....66
7.4 Metadatos.... 66
7.4.1 Contactos....66
7.4.2 Códigos de provedor....67
7.4.3 Mapa de provedor....67
7.5 Deteção e relatórios de erros.... 68
8 Leitura adicional....72
8.1 Sobre o software Atlas of Living Australia....72
8.2 Tecnologias e plataformas de desenvolvimento....72
8.3 Gestão de dados.... 73
8.4 Formatos de dados de biodiversidade.... 73
9 Internacionalização.... 74
9.1 Sobre a Interface do Utilizador do portal ALA....74
9.2 Guia para a tradução do UI passo a passo....75
9.3 Configuração da língua....80
9.4 Suporte Multi-língua.... 80
10 Canais de colaboração....83

1 Introdução

O projeto Atlas of Living Australia (ALA) é uma infraestrutura de investigação nacional da Austrália que desenvolveu uma plataforma e ferramentas para permitir aos investigadores e outros utilizadores de informação sobre biodiversidade a pesquisa, o acesso, a combinação com outros dados e a visualização de dados das plantas e animais Australianas.

O software produzido pelo projeto ALA tem sido progressivamente utilizado for a da equipa original do ALA, tendo-se tornado um projeto de software em código aberto (open-source).

Se desejar envolver-se e acompanhar os desenvolvimentos futuros, por favor consulte o capítulo Canais de Colaboração.

Este documento dá uma ideia geral dos princípios por detrás do software, características e possibilidades. O leitor encontrará detalhes mais elaborados na documentação wiki.

1.1 Como usar este documento

Este documento técnico fornece uma visão geral da arquitetura, funcionalidades, e dependências das ferramentas/tecnologia que facilitam a avaliação do software ALA.

Estrutura deste documento:

Capítulo 1: Introdução

Capítulo 2: Arquitetura

Capítulo 3: Requisitos e instalação

Capítulo 4: Configuração

Capítulo 5: Carregamento dos dados

Capítulo 6: Cópia de segurança e atualização

Capítulo 7: Funcionalidades

Capítulo 8: Leituras adicionais

Capítulo 9: Internationalização

Capítulo 10: Canais de colaboração

O documento responde a algumas questões preliminares que pode ter:

- O que é o portal de dados ALA? (leia Functionalidades e Arquitetura)

- Como instalá-lo? (leia Requisitos e Instalação)
- Como configurá-lo? Como alterar a aparência/tema? (leia Configuração)
- Como preenchê-lo com dados? (leia Carregamento de dados)
- Como atualizá-lo e manter o desenvolvimento local sincronizado com o repositório global? (leia Cópia de segurança e atualização)
- Onde obter mais informação detalhada? (leia Leituras adicionais)
- Como traduzir a Interface de Utilizador ALA (leia Internationalização)
- Como juntar-se a esta comunidade? (leia Canais de colaboração)

1.2 Portais do ALA derivados

Em baixo, encontrará uma tabela com informações sobre os Nós que instalaram e configuraram o portal de dados usando módulos do ALA. Pode contactá-los diretamente, se pretender obter mais informações sobre a sua instalação.

Nós	**Sítio Web**	**Contacto**	**Língua**
Atlas of Living Scotland	http://www.als.scot/	info@als.scot	Inglês
GBIF France	http://portail.gbif.fr/	dev@gbif.fr	Francês
GBIF Spain	http://datos.gbif.es/	info@gbif.es	Espanhol
SiBBr	https://portaldabiodiversidade.icmbio.gov.br/portal/		Português
GBIF Portugal	http://dados.gbif.pt	node@gbif.pt	Português

2 Arquitetura

2.1 Fundamentação do design de software ALA.

1. **Orientado aos serviços** - todo o conteúdo disponível para os utilizadores nas páginas web do Atlas está disponível através de um serviço web subjacente. Estes serviços web compõem a API pública do Atlas. Esta serve para promover o acesso a dados abertos e permitir aos parceiros do Atlas beneficiar de destes sistemas bara integrar elementos do Atlas nos seus sítios e ferramentas.
2. **Componentes modulares** - O Atlas é constituído por um conjunto de componentes de micro-serviços, cada qual com uma função específica. Isso ajuda a promover a reutilização e permite flexibilidade na implementação.
3. **Módulos de interface do utilizador re-utilizáveis** - Os componentes da interface do utilizador são baseados numa arquitetura de plugins. Isto destina-se a permitir a reutilização e personalização. Reconhecemos que uma marca personalizado é uma parte importante da produção de um portal para uma determinada comunidade.
4. **Portabilidade** - O Atlas é construído exclusivamente com software opensource. Todas as aplicações do Atlas podem ser instalados utilizando um conjunto de scripts mantidos para ajudar a adoção dos componentes do Atlas.

A plataforma Atlas of Living Australia é dividida em diferentes módulos que podem associados em conjunto, mas também pode trabalhar à parte (por exemplo, em França, a equipa apenas instalou três deles). O diagrama abaixo mostra toda a arquitetura do sistema ALA.

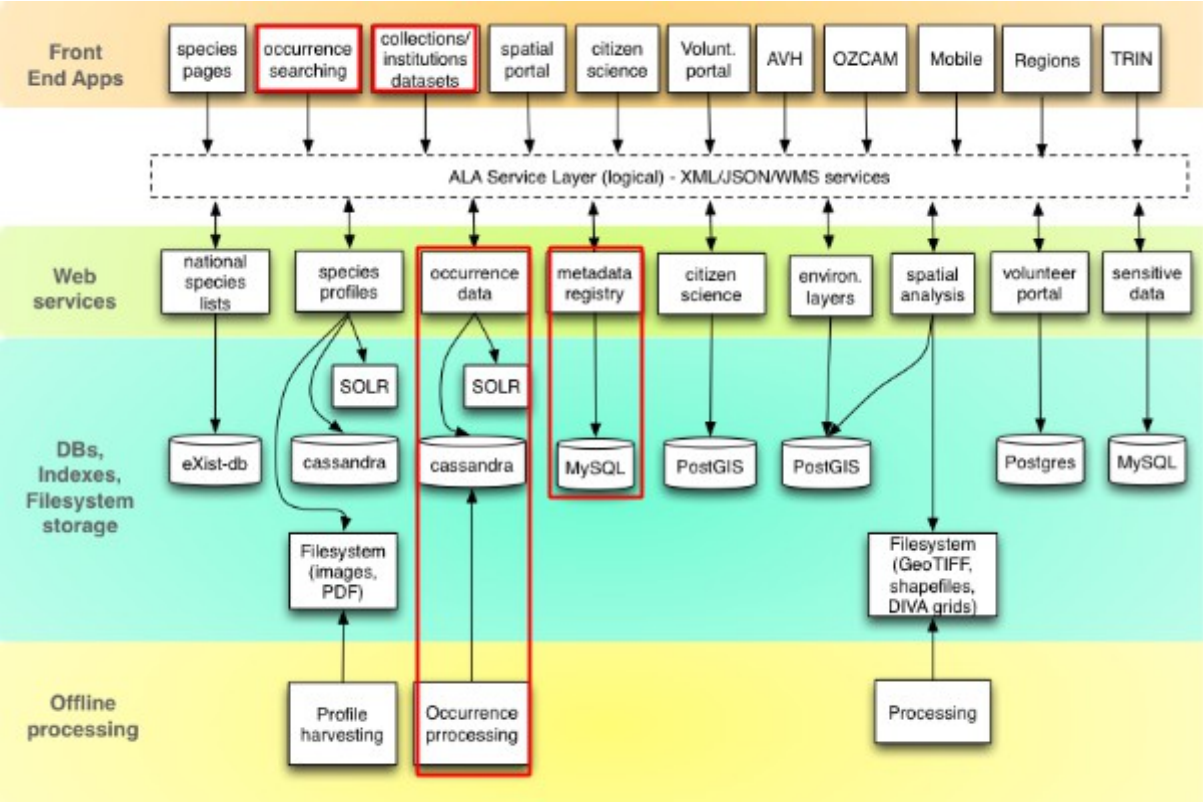


Ilustração 1: Arquitetura ALA

Neste capítulo, focaremos a nossa atenção nos módulos de ocorrências e de coleções (assinalados a vermelho na Ilustração 1), mas faremos também uma descrição curta dos outros componentes que a equipa do projeto ALA desenvolveu. Todos estes itens podem ser instalados em conjunto ou em separado. Iremos também descrever os princípios técnicos básicos (e.g. língua e software utilizado) e aconselhar sobre as necessidades de configuração do servidor com exemplos dos países que já o instalaram.

2.2 Visão geral dos módulos

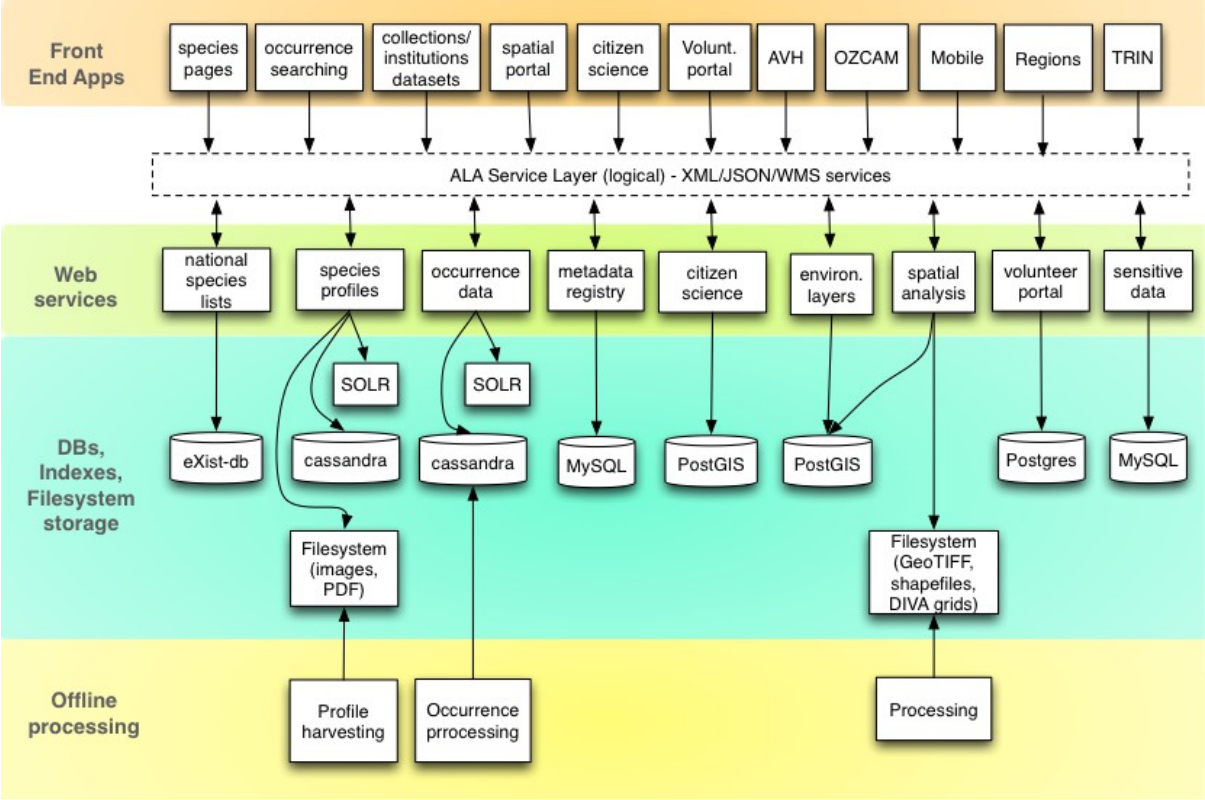


Ilustração 2: Visão geral dos módulos

2.3 Pesquisa de ocorrências

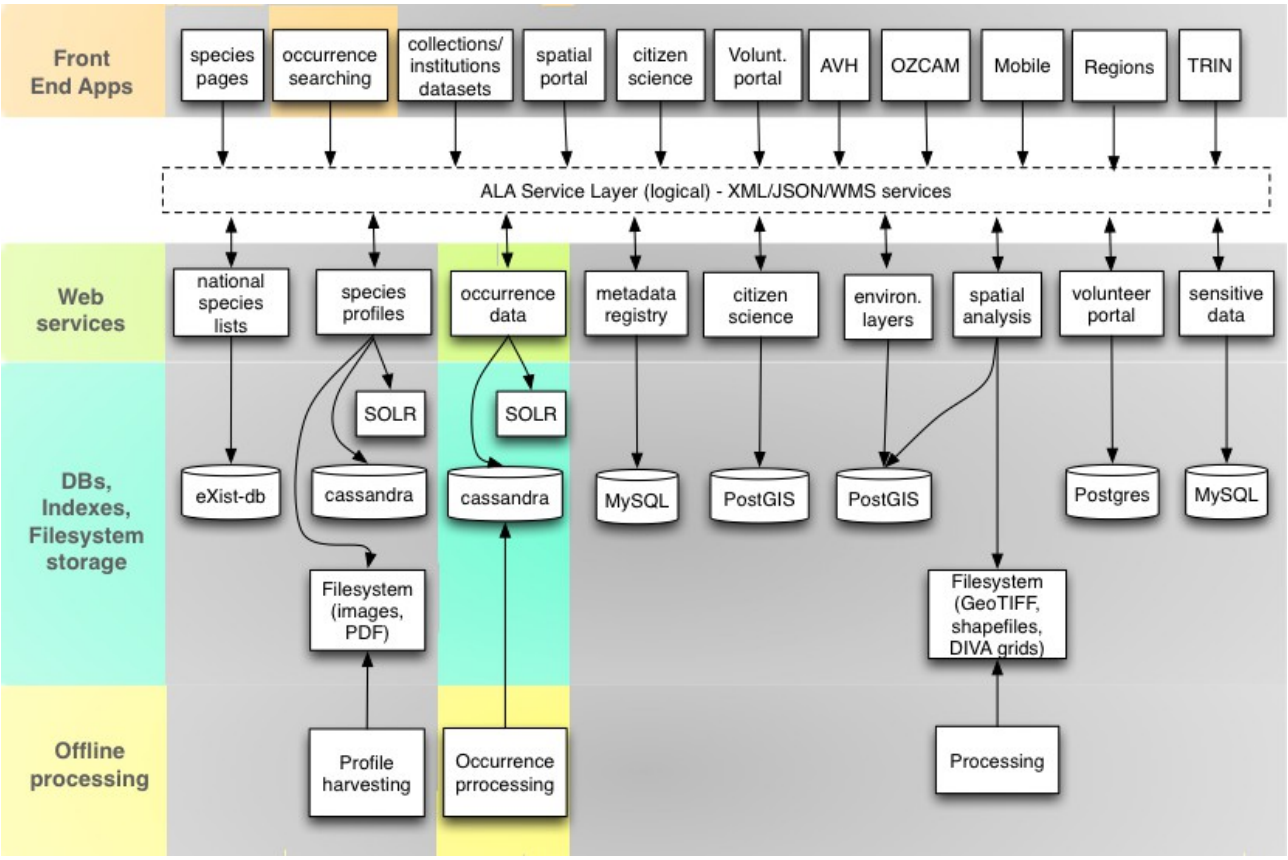


Ilustração 3: Módulos de pesquisa de ocorrências

A pesquisa de ocorrências é o motor de pesquisa do portal. Como irá verificar no capítulo dedicado Às funcionalidades, há diferentes caminhos para fazer uma pesquisa (simples, avançada, por taxon, ou utilizando limites espaciais).

A página de resultados mostra a lista com todos os resultados, diagramas, assim como um mapa com as ocorrências georreferenciadas. Na página de ocorrências. irá encontrar informação sobre estas, após a indexação. Irá também ter a possibilidade de navegar para as páginas da instituição ou coleção (parte do registo de metadados). É importante notar que todos os componentes do software ALA estão ligados.

Do ponto de vista técnico, temos uma aplicação web chamada biocache-hub que fornece uma interface HTML para o utilizador. Este componente utiliza os serviços web fornecidos pelo componente chamado biocache-service (dados de ocorrência no diagrama). Os dados são armazenados numa base de dados Cassandra, no sistema de ficheiros chamado occ e o ALA utiliza Apache SOLR para indexar os dados de ocorrência.

De modo a processar os dados pelo SOLR, a equipa de desenvolvimento do ALA criou uma ferramenta de indexação chamada Biocache. Esta também funciona com os serviços web biocache-service.

Basicamente, esta aplicação gere o carregamento, amostragem, processamento e indexação dos registos de ocorrência da plataforma ALA. Existem ferramentas adicionais para suporte da deteção de anómalos, duplicados e identificação de outliers extra-limite com base nos polígonos de referência para a distribuição dos taxa.

2.4 Collectory

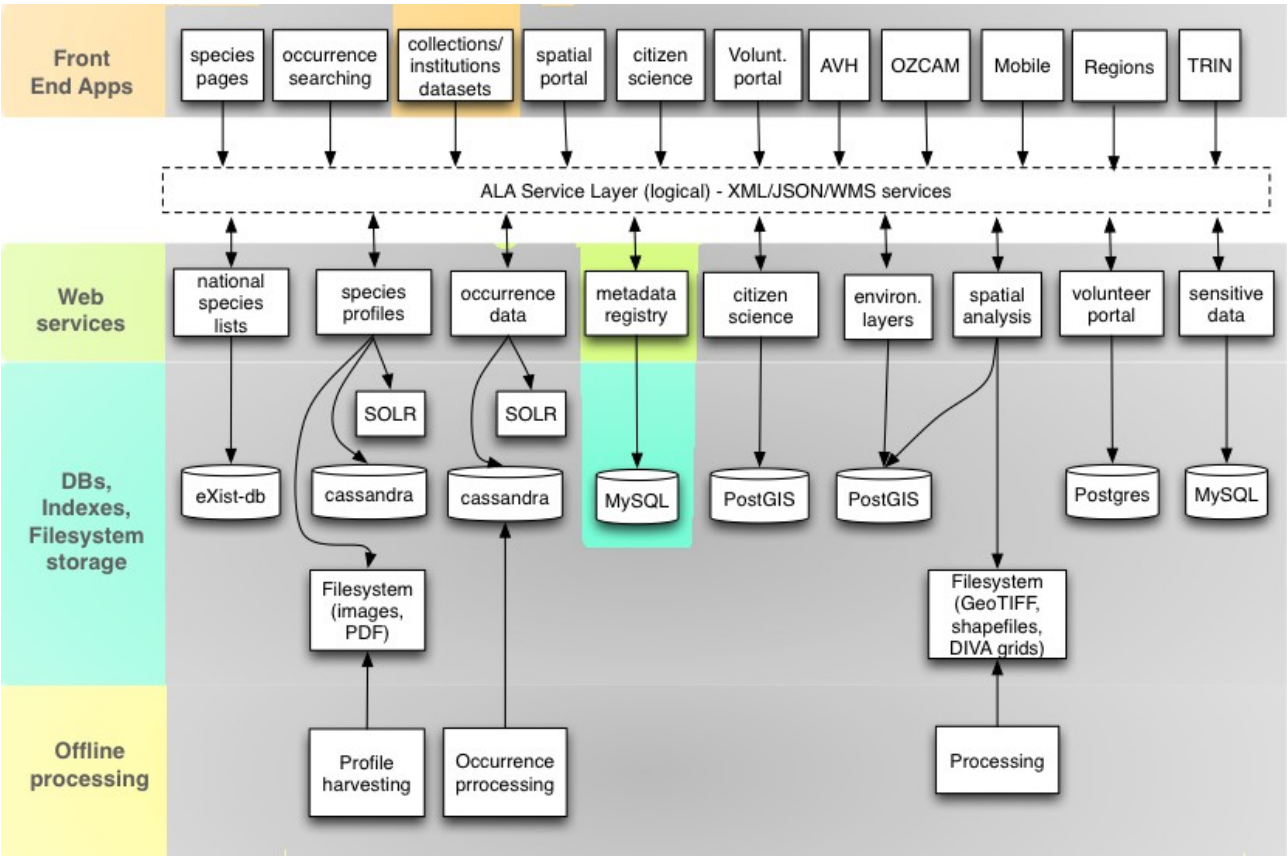


Ilustração 4: Módulos Collectory

O plugin collectory é o módulo que gere o registo dos metadados (informação sobre instituições, coleções, recursos de dados, contactos, etc). Todos os dados são guardados numa base de dados MySQL com o nome collectory. Este módulo está dividido em duas partes: páginas públicas e páginas privadas.

2.4.1 Páginas públicas

É possível aceder a todas as páginas públicas das instituições ou coleções, mas também a um motor de pesquisa com um mapa dedicado às coleções.

2.4.2 Páginas privadas

Esta é a parte de administração, É possível adicionar, modificar ou apagar instituições, coleções, recursos de dados ou contactos. É também possível ver um relatório do seu portal. Se quiser adicionar recursos de dados indexados pelo portal GBIF global GBIF.org, vai encontrar todas as ferramentas necessárias para isso nas páginas privadas.

Esta aplicação pode ser alojada num servidor independente.

Do ponto de vista técnico, existe apenas uma aplicação web e uma instância MySQL. Esta é uma das arquiteturas mais simples de todo o sistema. Pode ver a destacado os elementos utilizados.

2.5 Página de espécies

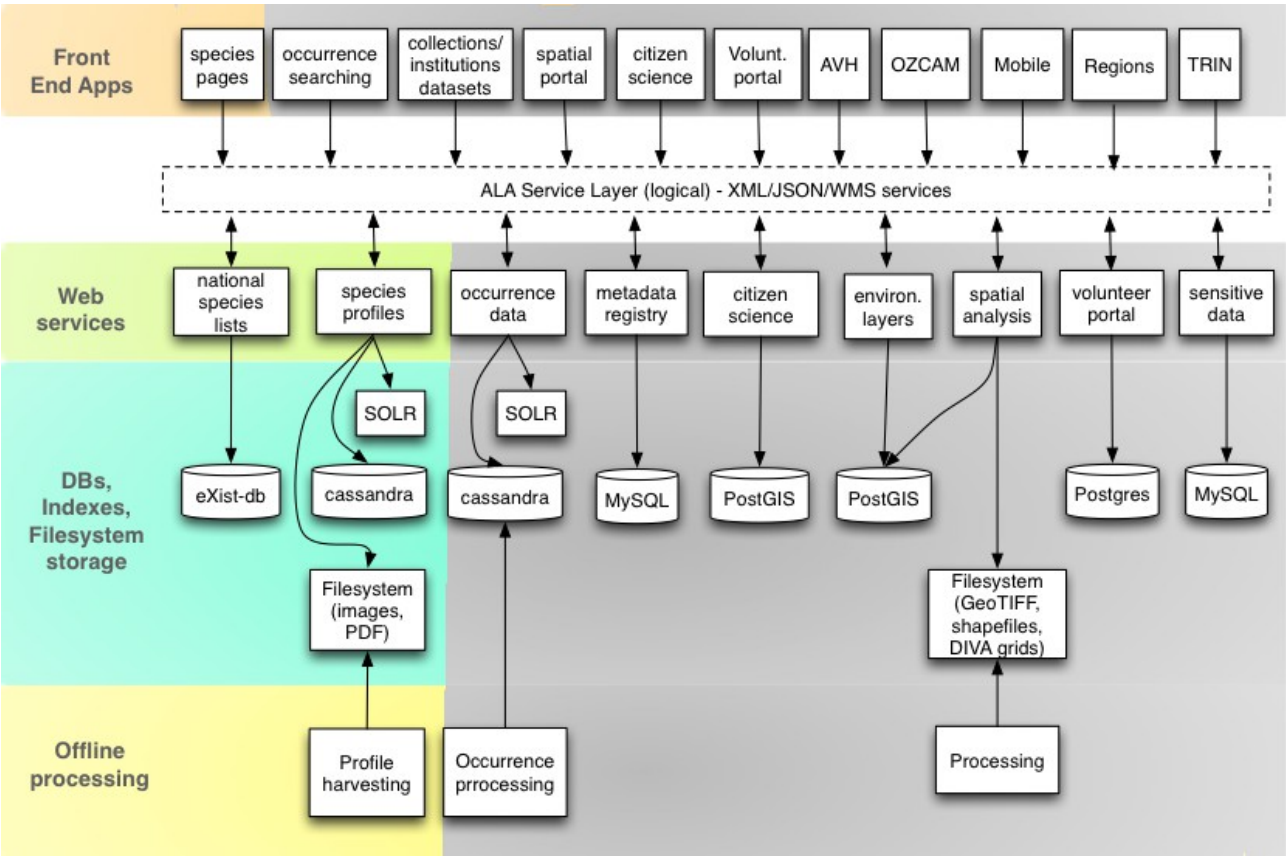


Ilustração 5: Módulo Espécies

Esta página permite que o utilizador navegue pelas espécies existentes que estão alojadas no portal. Pode filtrar a lista de espécies por Secção, Forma de Vida, Grupo Espécies e Estado de Conservação. Para cada espécie o utilizador pode ver a informação sobre o Nome, Classificação, Registos, Literatura e Sequências. O sistema vincula as ocorrências com as páginas das suas

espécies, desta forma permitindo que quando esteja na página de uma espécie possa ver os registos relacionados com esta. Pode encontrar exemplos aqui:

http://bie.ala.org.au/ e
http://bie.ala.org.au/species/Carcharhinus+amblyrhynchos.

Do ponto de vista técnico, está desenvolvido com Grails Framework e utiliza os Serviços Web REST do portal.

2.6 Hub de dados

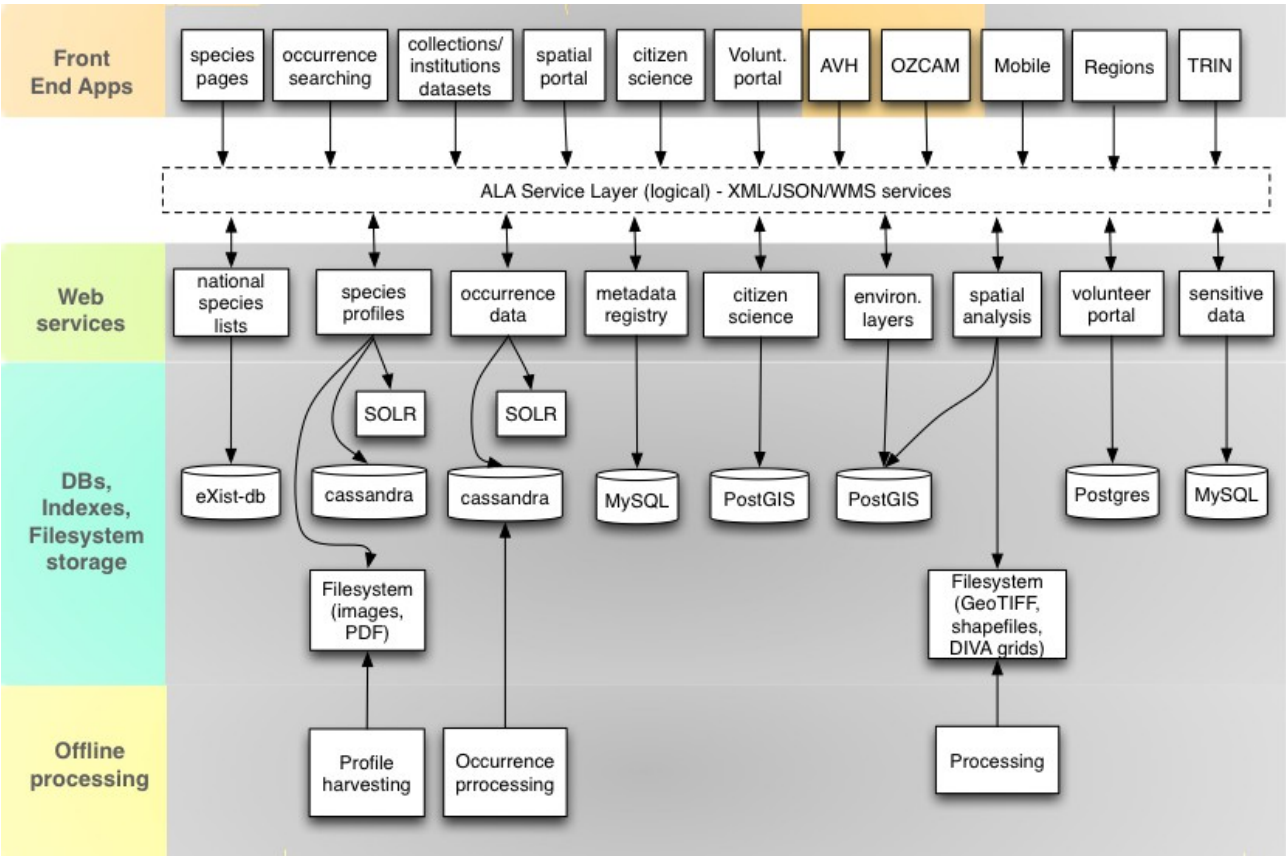


Ilustração 6: Módulos Datahub

A primeira coisa que deve entender é o conceito de Hub de Dados, o que pode ser feito facilmente a partir do seguinte exemplo gráfico.

No Nó do GBIF Espanha, está disponível o portal de dados principal em: http://datos.gbif.es, através do qual os utilizadores podem ver todos os registos que os provedores espanhóis publicam.

O que acontece se um destes provedores quer mostrar um subgrupo destes registos na sua rede institucional?

Para estes casos, o provedor pode utilizar a aplicação web Hub de Dados, que tal como pode ser vista aqui:

http://161.111.171.47:8080/myxo-hub/search#tab_simpleSearch

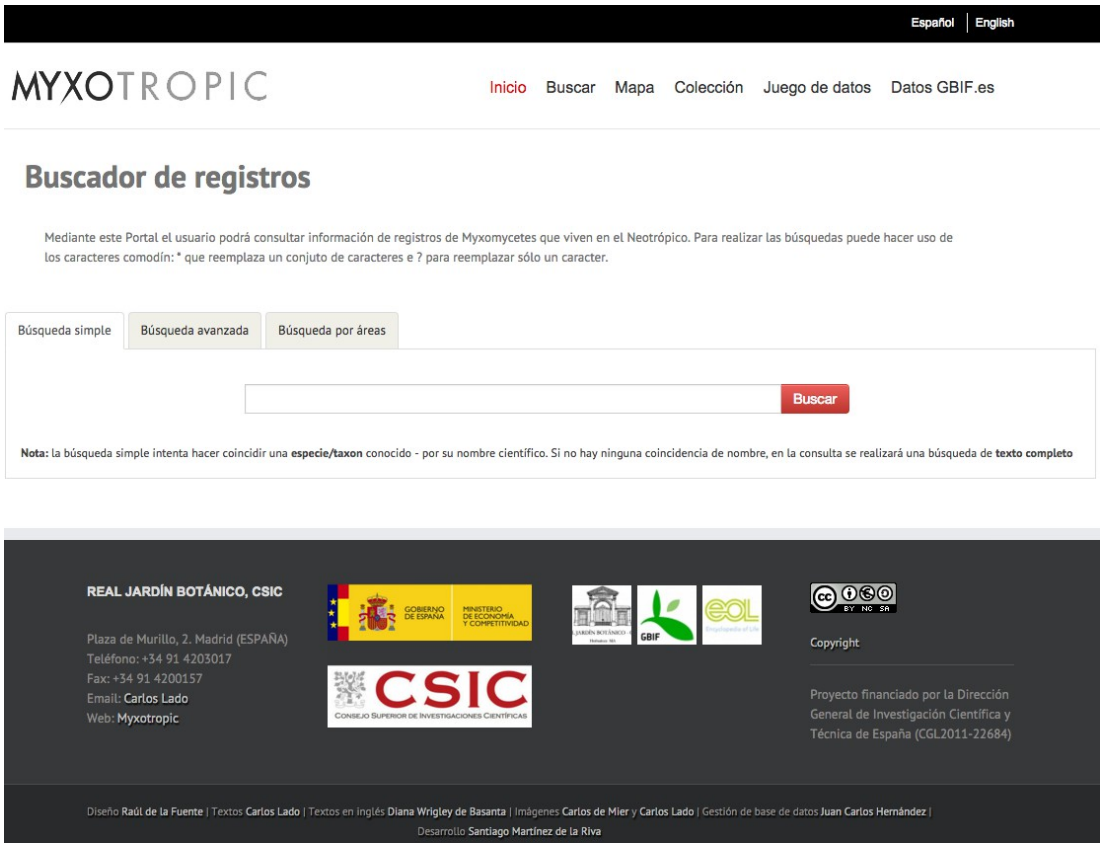


Ilustração 7: Hub de Dados Myxotropic

Neste hub de dados, quando um utilizador faz uma pesquisa, o resultado mostrará apenas os registos que estão no subgrupo definidor pelo provedor, neste caso, o projeto sobre Myxomycetes da região do Neotrópicos, cujo grupo de registos é definido no portal de dados principal, especificamente no recurso de dados abaixo:

http://datos.gbif.es/collectory/public/showDataResource/dr190

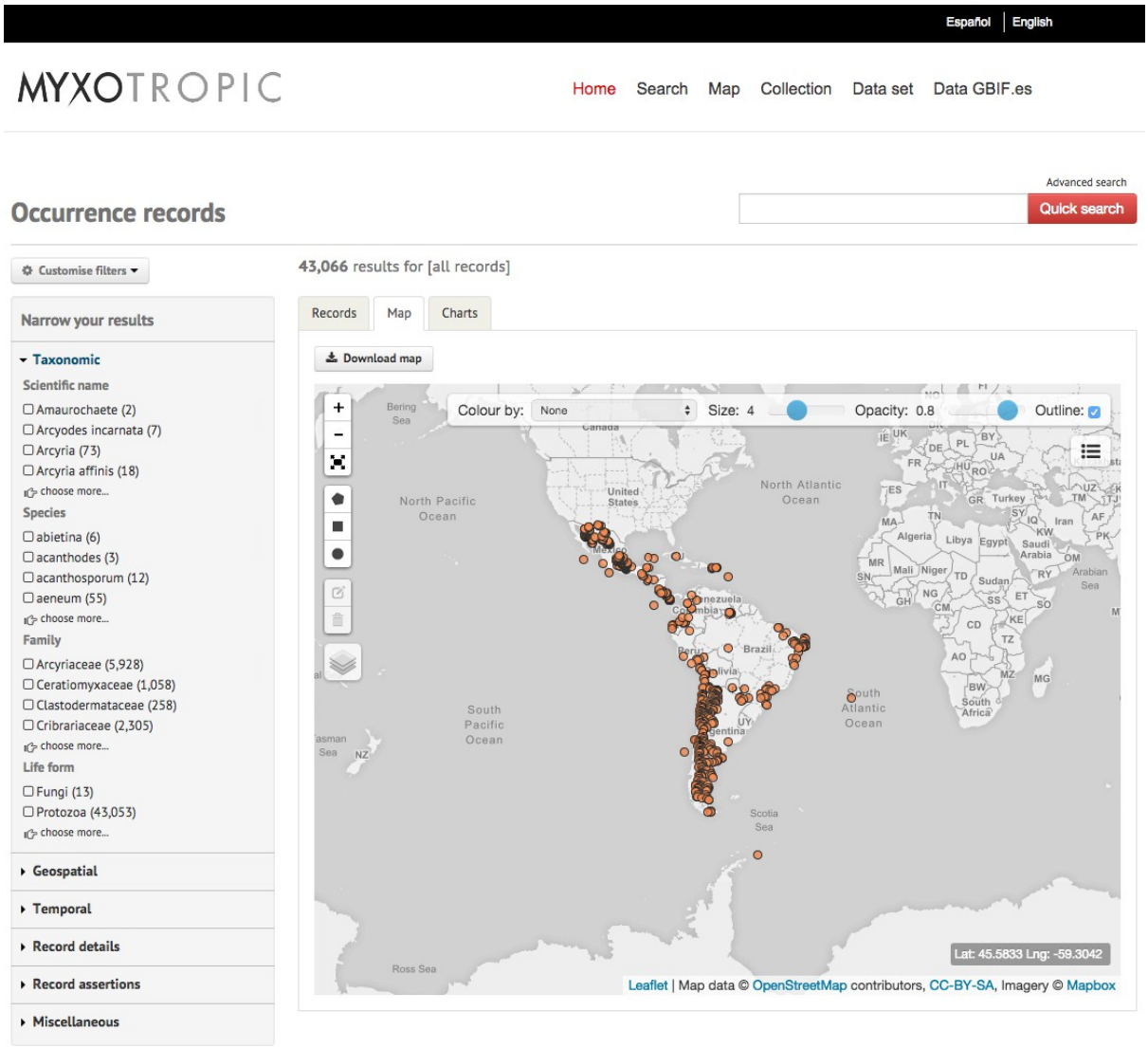


Ilustração 8: Ocorrências Myxotropic

Esta aplicação web pode ser alojada num servidor do provedor. O software que esta aplicação necessita é Java e servidor Tomcat. Naturalmente, a aplicação web deve ser apontada para o subgrupo de registos que o provedor pretende mostrar na sua web.

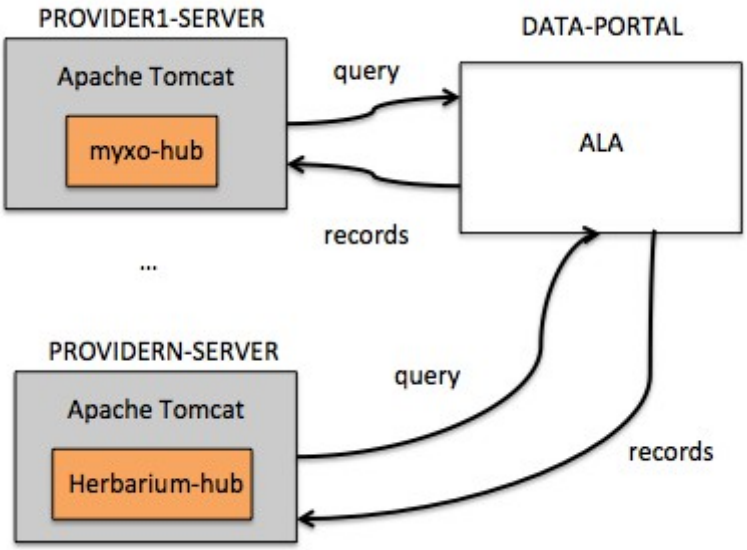


Ilustração 9: Hub de Dados

Pode ver mais sobre como configurar o seu próprio Hub de Dados aqui:

https://github.com/AtlasOfLivingAustralia/documentation/wiki/Configuring-generic-hub

> Nota: Pode aceder ao hub de dados AVH (http://avh.chah.org.au/) e aos hub de dados OZCAM (http://ozcam.ala.org.au/)

2.7 Regiões

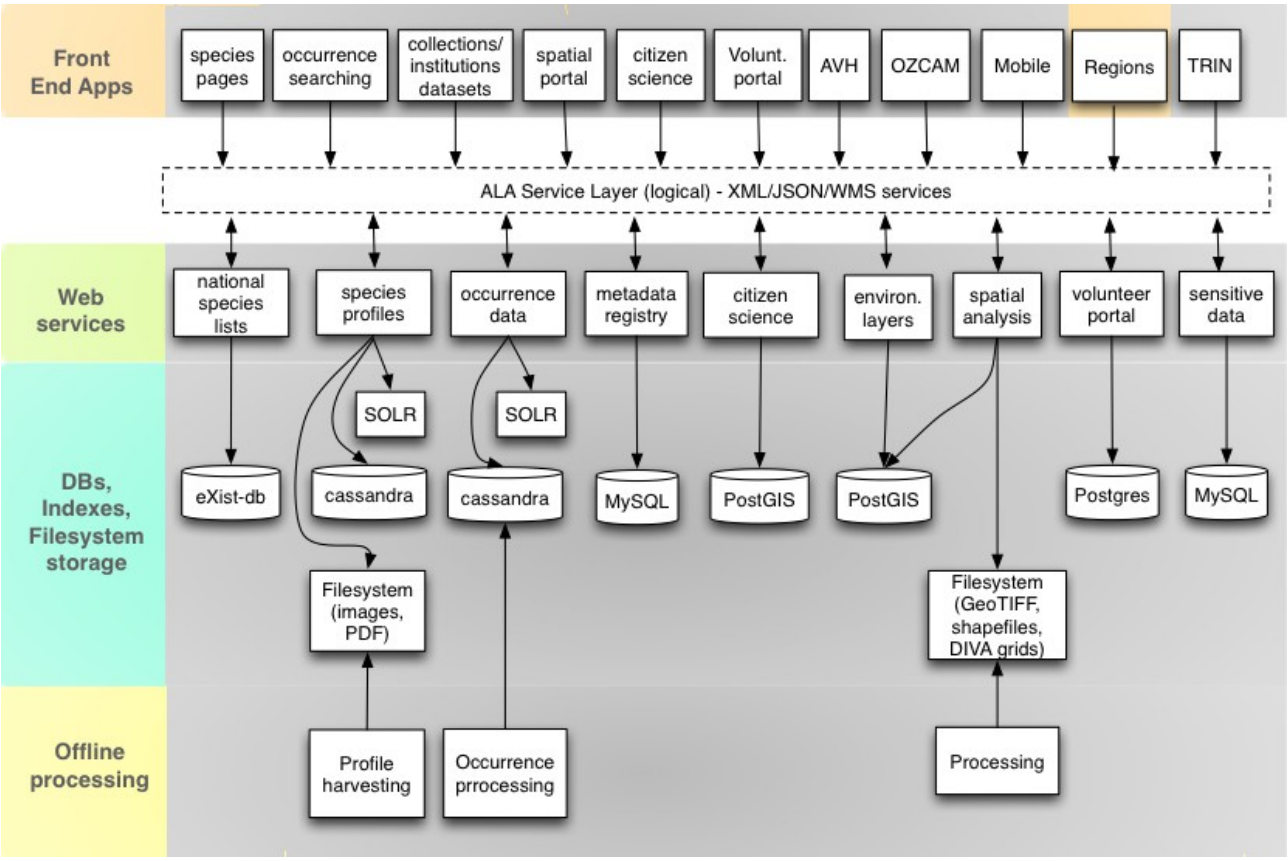


Ilustração 10: Módulo Regiões

Este projeto permite que o utilizador navegue de uma forma conveniente os registos existentes no portal ALA que foram previamente filtrados por ele através da seleção de uma área sobre as regiões mostradas num mapa.

Pode encontrar exemplos aqui:

http://regions.ala.org.au/ e http://regions.als.scot/.

Do ponto de vista técnico, foi desenvolvido com base em Grails Framework e usa Serviços Web REST do seu portal de dados. Pode também utilizar Google Maps Web API e Google Charts.

Esta componente é mais fácil instalar do que o portal espacial. Convidamos a que se junte à comunidade discutindo as potencialidades desta ferramenta, se decidir instalar a ferramenta espacial no seu portal.

2.8 Portal Espacial

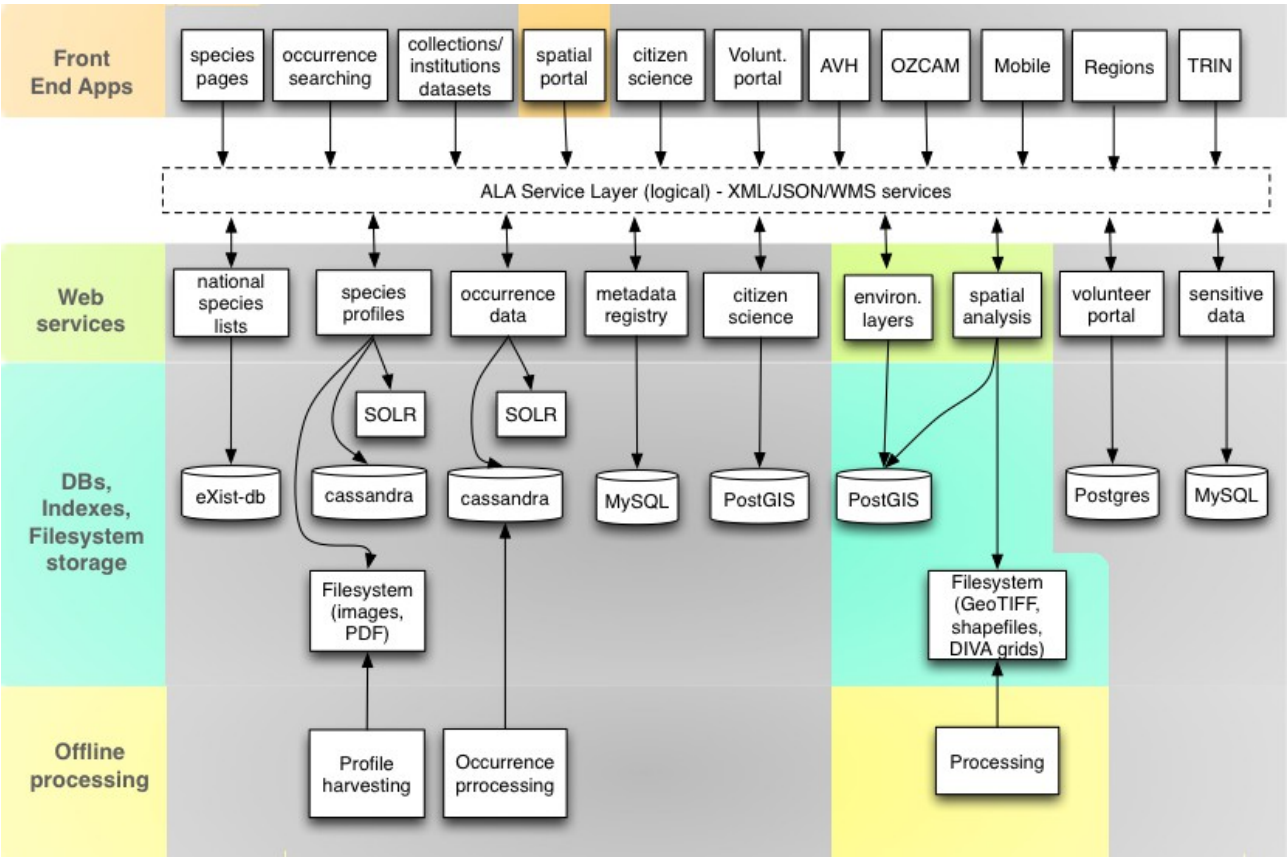


Ilustração 11: Módulos do Portal Espacial

O Portal Espacial ALA é composto por 4 aplicações web. Estas aplicações web estão separadas com base nas suas tarefas e utilizam serviços web para comunicar entre si quando necessário. No entanto, é necessário um apontador para o armazenamento global é necessário.

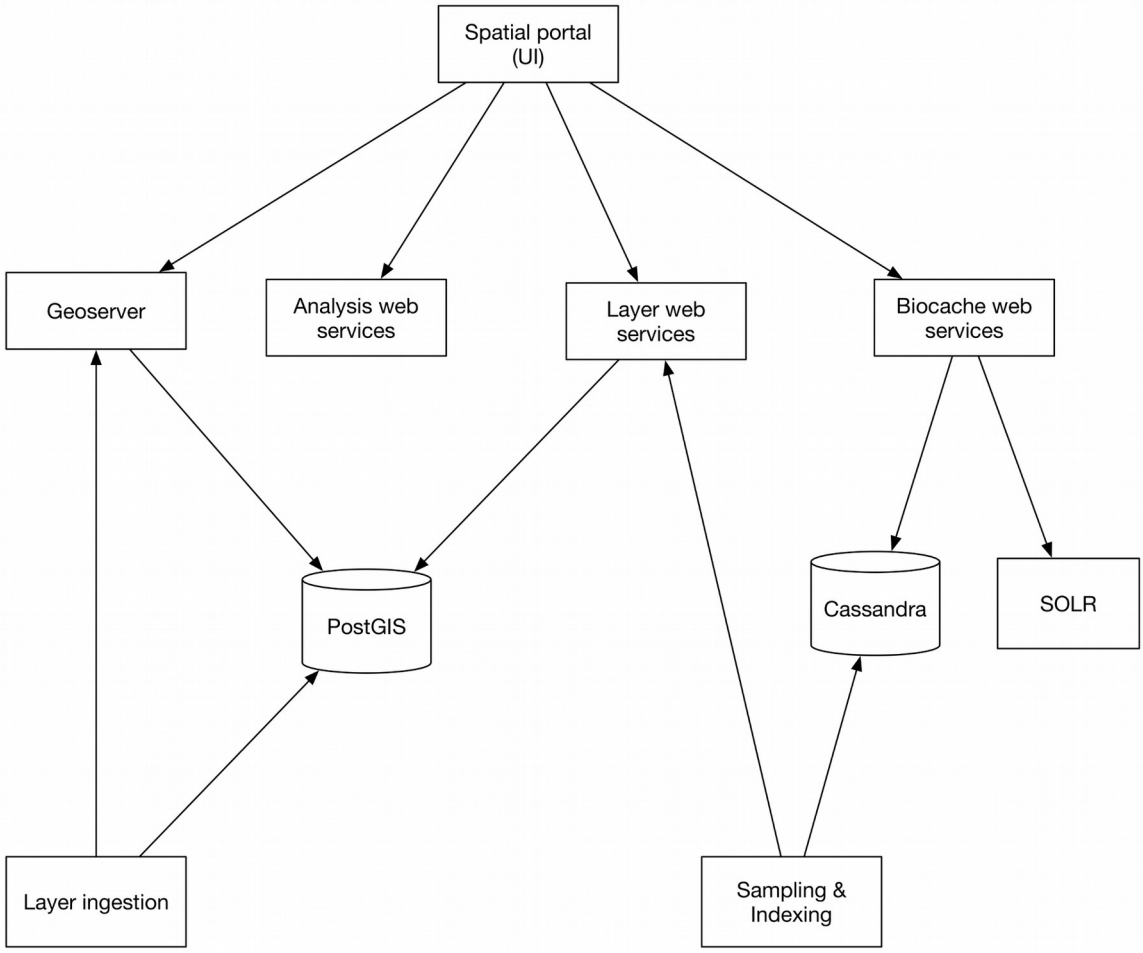


Ilustração 12: Portal Espacial

Os componentes da aplicação web que constituem o total do Portal Espacial são:

- UI Portal Espacial ALA (webportal)
- Serviço Análise Espacial ALA (alaspatial)
- Serviço Camadas Espaciais ALA (layers-service)
- Ações Espaciais ALA (actions)

Existem algumas utilidades que estão construidas para serem partilhadas entre cada um destes componentes:

- Loja Camadas Espaciais ALA Spatial (layers-store)
- Utilidades Espaciais ALA (utilities)

O Portal também utiliza outros serviços web ALA, tais como:

- BIE
- Biocache
- Collectory

O Portal Espacial utiliza a ferramenta padrão Geoserver (versão 2.8+) para fornecer os dados ambientais, de contexto e resultados de análises dinâmicas através de protocolo WMS. Por favor consulte o Geoserver para mais informação em http://www.geoserver.com/. Um serviço WMS adaptado e simplificado foi criado para servir dados de biodiversidade. Este serve os dados de ocorrência como pontos ou células de uma grelha (e.g. 1 grau, latitude/longitude), com algumas distribuições de espécies fornecidas por especialistas servidas como polígonos. Por favor consulte http://spatial.ala.org.au/ws/ para mais informação. Exemplos de uso de visualização de dados de biodiversidade estão disponíveis em http://spatial.ala.org.au/ws/examples/.

Mais informação no repositório Git do projeto ALA (página wiki dedicada ao portal espacial)

2.9 Princípios técnicos básicos

Para o collectory e pesquisa de ocorrências, o ALA utiliza como principal ferramenta o Grails, uma plataforma web. Esta é baseada na linguagem de programação Groovy e funciona com Java e Hibernate, entre outros. Os módulos são Web Application Archive (WAR) e é utilizado Tomcat 7 como servidor web.

A base de dados utilizada pelo módulo collectory é MySQL. A pesquisa de Ocorrências usa Apache SOLR na indexação, sobre a base de dados Apache Cassandra, que assegura o armazenamento. Vai encontrar mais informação nos capítulos sobre Requisitos e instalação e Configuração.

De modo a configurar ou mudar o design da aplicação web utilizado neste projeto, deve ter um nível de conhecimento razoável de HTML e CSS, assim como um conhecimento básico de Grails e Bootstrap.

3 Requisitos e instalação

O Portal ALA requer vários componentes como Java, Tomcat e Cassandra, assim como aplicações web que compõem o Portal ALA. Estes softwares podem ser automaticamente instalados e configurados através de playbook Ansible, uma espécie de scripts, desenvolvidos em simultâneo com o projeto. Assim, se

tiver uma instância de linux Ubuntu disponível e a funcionar, pode utilizar os playbook Ansible do projeto ala-install para instalar e configurar automaticamente um servidor Linux no Portal ALA.

Para além dos playbooks Ansible, se não tiver uma instância Linux disponível, ou se apenas quiser fazer uma instalação limpa do Portal ALA, pode considerar as duas seguintes ferramentas para criar uma instância limpa:

1. Vagrant: Cria e configura uma máquina virtual para alojar o portal ALA;
2. VirtualBox: O software de virtualização para conter máquinas virtuais;

3.1 Requisitos do servidor de teste

- Máquina virtual com Ubuntu 12 ou 14.
- 100GB de armazenamento de disco livre (idealmente SSD), uma vez que a indexação e processamento de dados consome espaço muito rapidamente.
- 32 GB RAM.
- 2 CPUs.

O guia de instalação assume que se pretende criar uma instância Linux e configurar um portal ALA de raiz. Se pretende configurar uma instância de Linux existente (Ubuntu), pode avançar diretamente para a secção Ansible do guia de instalação. Este tutorial pode apenas ser executado em sistemas UNIX (Linux/Mac OS X), uma vez que o Ansible ainda não está disponível para Microsoft Windows, à data de 20 de Maio de 2016. (http://docs.ansible.com/ansible/intro_installation.html)

Utilizando um Macintosh como exemplo, estes são os passos para preparar as ferramentas:

Descarregar Vagrant e instalar o ficheiro descarregado.

Descarregar VirtualBox e instalar o pacote descarregado.

Para instalar o Ansible (versão 2.0.1.0+), a forma mais fácil é instalar via homebrew. É também conveniente ter o Command Line Tools for Mac OS X instalado. Uma vez pronto, os seguintes comandos irão instalar o Ansible:

```
$ brew update
$ brew install ansible
```

No entanto, neste guia, o Vagrantfile irá determinar as configurações do servidor por si. Agora, está pronto para iniciar a instalação do portal ALA.

3.2 Requisitos do servidor de produção

Se quiser implementar o portal ALA numa infraestrutura complexa, pode utilizar várias máquinas virtuais para organizar os diferentes serviços: Tomcat, Solr, Mysql, Cassandra, Image service…

> Nota: Pode ver o capítulo de Instalação para perceber como instalar os playbooks em diferentes servidores.

Na imagem abaixo, pode ver quais podem ser os requisitos que necessita para cada Máquina Virtual (infraestrutura Espanhola):

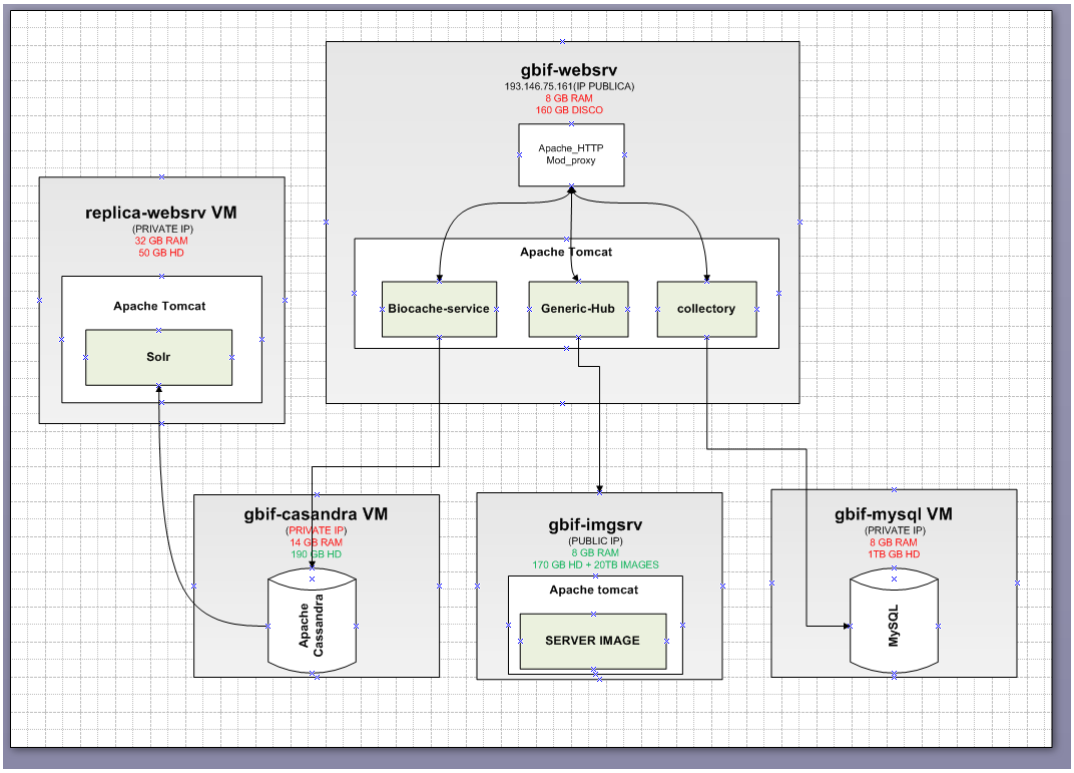


Ilustração 13: Infraestrutura do portal de dados espanhol

Os requisitos estão intimamente relacionados com o número de registos, que pretende publicar, e obviamente vão estar relacionados com o número de visitas que o portal de dados irá ter.

No caso do nó espanhol, que no momento em que se escreve este texto, tem 12 milhões de registos, 189 conjuntos de dados e 89 instituições, a principal máquina virtual é "gbif-webserv" que contém o serviço Apache e Tomcat.

Este servidor mostra-nos a página inicial http://datos.gbif.es e redireciona as chamadas ao serviço tomcat para o porto :80.

O serviço Tomcat contém as seguintes aplicações web do portal ALA:

- generic-hub: http://datos.gbif.es/generic-hub/search#tab_simpleSearch
- collectory: http://datos.gbif.es/collectory/public/showDataResource/dr190
- biocahe-service (serviço API):
- http://datos.gbif.es/biocache-service/occurrences/collections/co11

O "gbif-websrv" precisa de uma boa capacidade de armazenamento, porque os conjuntos de dados que serão publicados no portal de dados são armazenados neste servidor.

Deve ter em mente que todos os utilizadores que visitam o portal de dados estão a usar os recursos da máquina virtual, pelo que se quiser adicionar, atualizar ou excluir registos no sistema através do processo biocache, irá utilizar parte desses recursos.

Assim, se precisar de mais recursos, porque está a receber muitos acessos ou tem um grande número de registos, pode adicionar RAM a este Máquina Virtual, e pode também adicionar outra Máquina Virtual para alojar, por exemplo, o módulo collectory, que mostra apenas a metainformação de instituições, coleções e recursos de dados, que são armazenado no servidor "gbif-mysql".

Neste caso o servidor “gbif-mysql” tem 8 GB RAM e 1 TB de armazenamento. Abaixo estão 3 links onde pode ver o tipo de informação que é armazenada no servidor mysql.

- Recurso: http://datos.gbif.es/collectory/public/showDataResource/dr190
- Coleção: http://datos.gbif.es/collectory/public/show/co188
- Instituição: http://datos.gbif.es/collectory/public/show/in61

“gbif-websrv” faz chamadas à Máquina Virtual “replica-websrv”, porque neste servidor pode estar alojado o serviço SOLR, na qual alguns campos dos registos que são mostrados no portal de dados são indexados com este sistema Desta forma, quando utilizador pesquisar esta informação indexada, esta é encontrada mais rapidamente.

Este serviço é usado em cada consulta que é feita pelo utilizador, pelo que é necessária uma quantidade considerável de RAM, não só para executar as consultas, também para indexar o número de registos que existem. e.g:

http://datos.gbif.es/generic-hub/occurrences/searchq=data_resource_uid:dr190#tab_mapView

http://datos.gbif.es/generic-hub/occurrences/search?q=data_resource_uid%3Adr190&fq=country%3A%22Chile%22&fq=taxon_name:%22Arcyodes+incarnata%22

Para compreendê-lo melhor com estes exemplos, dataresource_uid, país e taxon_name são facetas/campos que são indexados no serviço Solr. Para cada uma destas urls, estará a usar recursos do servidor "replica-webbsrv".

Por outro lado, se quiser ver a informação de um registo específico,

http://datos.gbif.es/generic-hub/occurrences/92fad279-47cc-4bf3-9d3d-f0644ad8fe79

Neste caso, a informação destes campos está alojada no MV gbif-cassandra, que é um servidor No-SQL, para a qual é suficiente ter 8 GB de RAM para a gerir. Novamente estes dados estão intimamente relacionados com a quantidade de informação que existe para armazenar e com o número de pedidos ao servidor.

Em cada caso, deve estimar a quantidade de memória que precisaria de acordo com o número de utilizadores e as informações que será gerida no portal de dados.

3.3 Outros exemplos de Arquitetura do Sistema

Por agora, a plataforma Atlas of Living Australia foi instalada e configurada e, vários países, tais como Argentina, Brasil, França, Portugal, Escócia e Espanha. Irá encontrar abaixo alguns exemplos da arquitetura do sistema instalado por estes países. Vai encontrar mais informação no capítulo Configuração.

3.3.1 Austrália

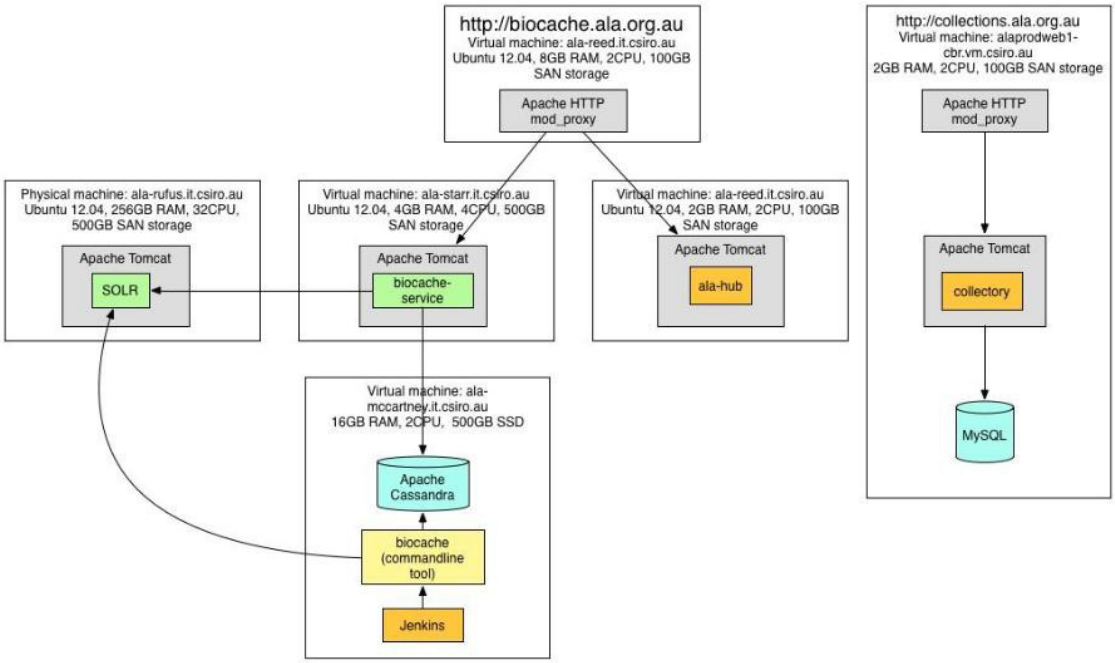


Ilustração 14: Infraestrutura Australiana

3.3.2 França

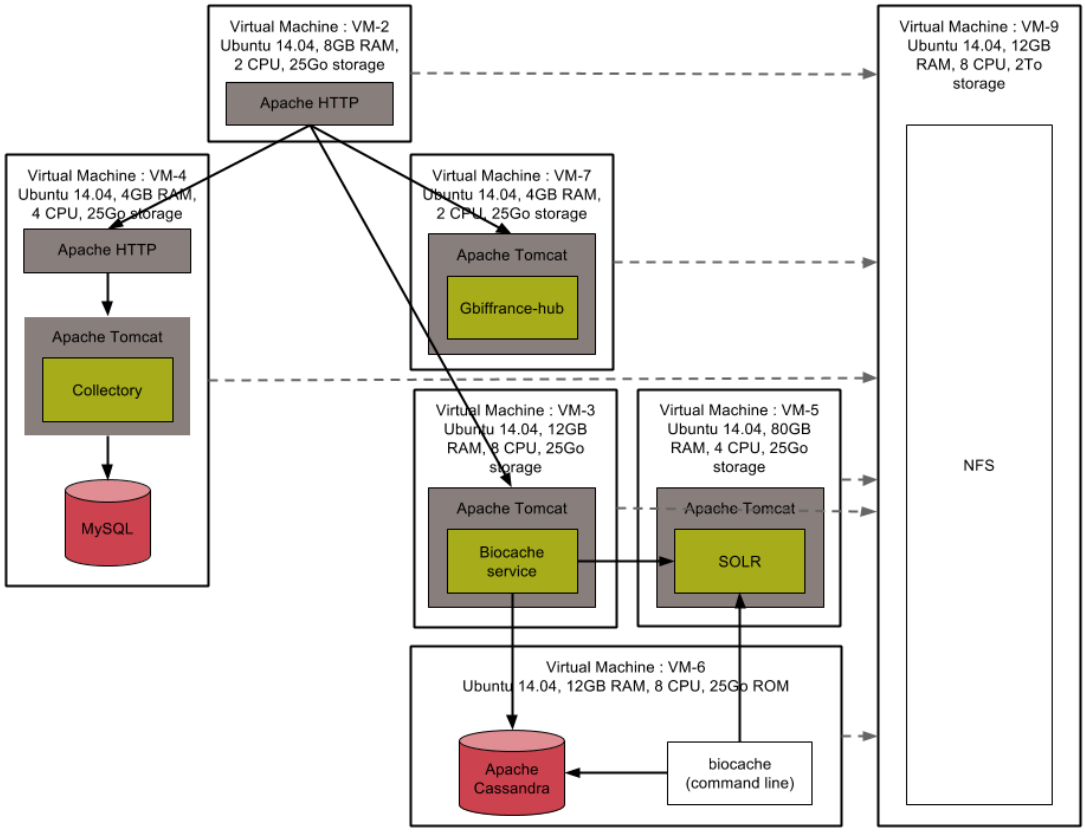


Ilustração 15: Infraestrutura Francesa

Para mais informações sobre o sistema, convida-mo-lo a contactar diretamente a equipa de TI de cada país. Uma vez que pode haver alterações destas equipas ao longo do tempo, poderá encontrar as informações e contactos mais recentes sobre as pessoas responsáveis diretamente no site dos países:

- Australia: http://www.ala.org.au/
- France: http://www.gbif.fr
- Scotland: http://www.als.scot/
- Spain: http://www.gbif.es

3.4 Processo de instalação(Ambiente de teste)

Tem de ter um ficheiro Vagrant e um playbook Ansible para poder guiar o Vagrant e Ansible nos procedimentos automáticos.

A instalação completa, num INTEL i7 2.66Ghz MacBook Pro (final de 2010) com SSD, demora cerca de 30 minutos a acabar. Isto inclui o tempo de transferência dos ficheiros war através da internet.

Para obter os ficheiros Vagrant e Ansible, clone o repositório GitHub ALA-install em:

https://github.com/AtlasOfLivingAustralia/ala-install

```
$ git clone https://github.com/AtlasOfLivingAustralia/ala-install.git
$ cd ala-install
```

3.4.1 Vagrant

```
$ cd vagrant/ubuntu/
$ vagrant up
```

Neste momento, deve ver uma instância ubuntu ativa e a correr se abrir o VirtualBox:

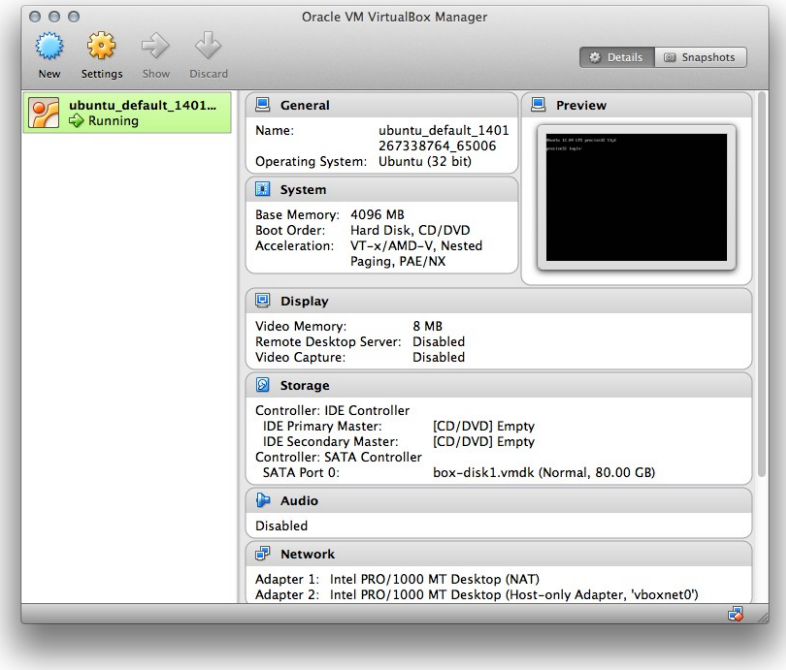


Ilustração 16: Ubuntu

À data de 28 de Maio de 2014, poderá ver "default: stdin: is not a tty" a vermelho. Este erro não afeta a instalação se fizer:

```
$ vagrant ssh
```

Pode fazer login a instância ubuntu que acabou de instalar:

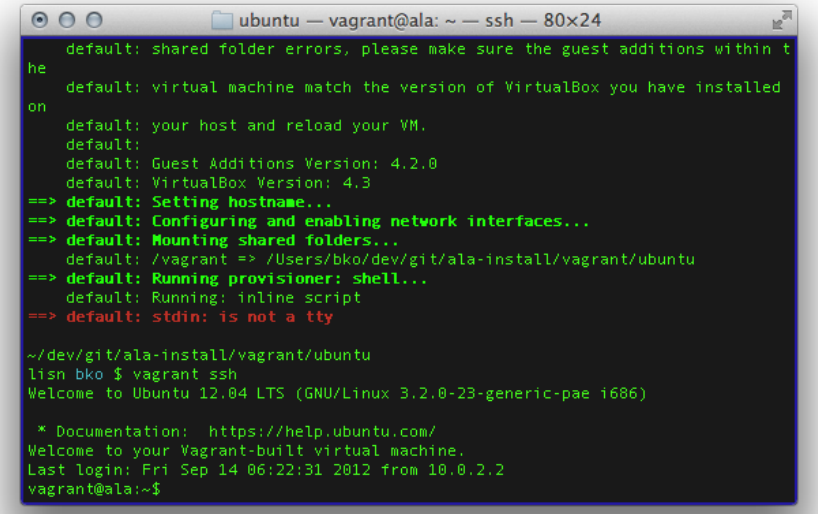


Ilustração 17: Login Ubuntu

3.4.2 Ansible

Neste momento está em condições de configurar o servidor ubuntu com o Ansible. Para correr o playbook Ansible:

```
$ cd ../../ansible/
$ ansible-playbook -i inventories/vagrant/demo-vagrant ala-demo.yml
--private-key ~/.vagrant.d/insecure_private_key -u vagrant -s
```

O playbook deve acabar. Podem existir alguns erros menores não abortivos que não afetam o processo. Se for este o caso, por favor comunique o problema aos programadores, que irão investigá-lo. Obrigado!

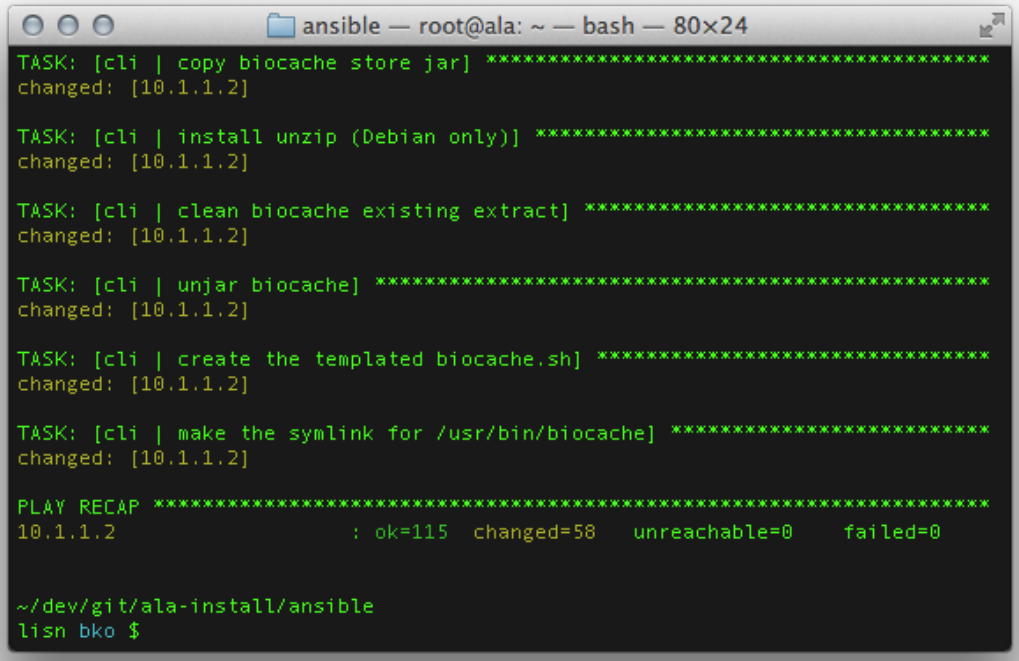


Ilustração 18: Ansible

O portal de demonstração do ALA deve agora estar acessível. Para conveniência, uma vez que em Vagrantfile o hostname está definido como ala.vagrant.dev e tem o endereço IP

```
10.1.1.2, adding a line in /etc/hosts:
10.1.1.2  ala.vagrant.dev
```

…ira permitir visitar o portal demo ALA a partir do servidor hóspede.

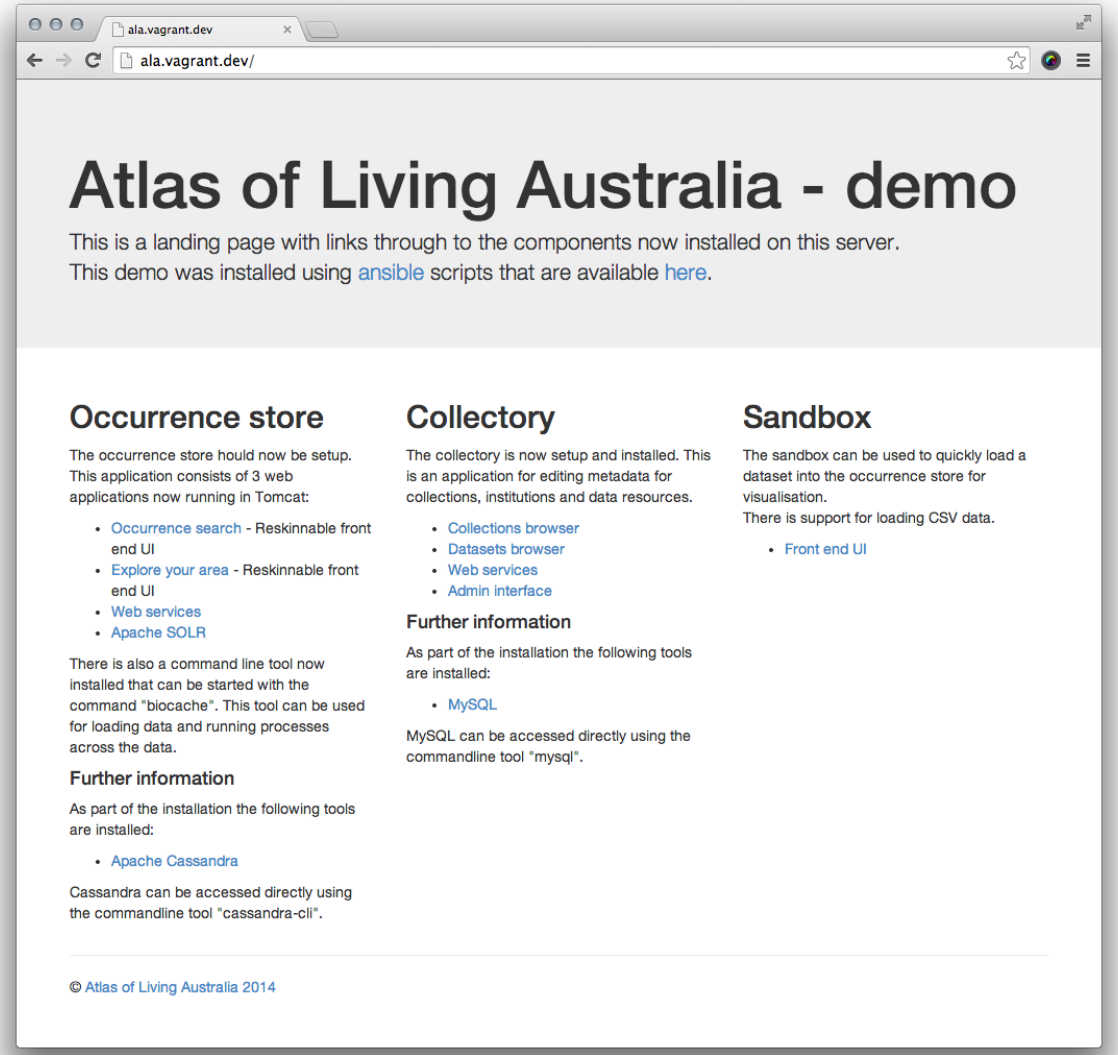


Ilustração 19: Demo ALA

Parabéns! O seu portal demo ALA está ativo e a correr.

3.4.3 Terminar a MV

Quando estiver satisfeito com a instalação de teste, possivelmente terá desativar, tendo para tal de parar a máquina virtual. Para fazer isso:

```
$ cd ../vagrant/ubuntu/
$ vagrant halt
```

Este comando desliga a máquina que o Vagrant está a gerir. Se quiser remover esta instância de teste, faça, na mesma diretoria:

```
$ vagrant destroy
```

Isto irá limpar a máquina virtual e todas as configurações feitas com o Ansible.

3.5 Processo de Instalação em servidores Remotos (ambiente de Produção)

3.5.1 Compreender o conceito

Este processo é muito semelhante ao precedente, pelo que as diferenças neste caso estão em que não necessita de uma máquina Virtual, nem do software Vagrant, apenas tem de correr o ansible que aponte para o servidor remoto, como é explicado na secção da documentação do wiki do ala-install:

https://github.com/AtlasOfLivingAustralia/ala-install#installing-the-ala-demo-on-ec2-or-other-cloud-providers

Fluxo de informação:

1. Na sua máquina local, deve instalar o git para clonar o repositório remoto do ala-install
2. Deve criar um inventory particular
3. Finalmente deve correr o playbook ansible no seu servidor.

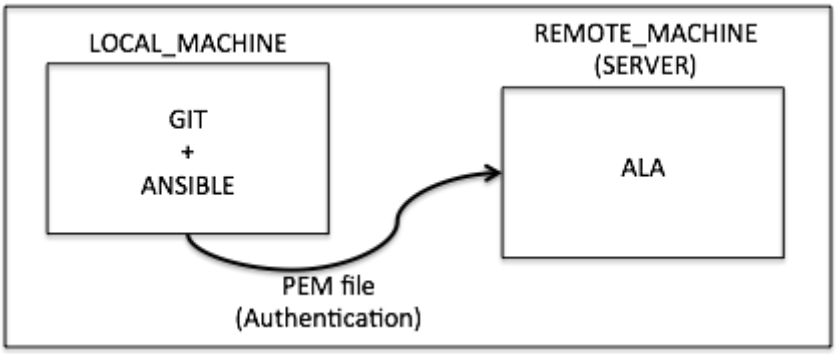


Ilustração 20: Fluxo de informação

Como pode executar estes comandos? Muito simples, siga apenas os seguintes passos de instalação.

3.5.2 Passos de instalação

Uma vez compreendido este processo para instalar o software do ALA, pode fazer a cópia do ala-install para a sua máquina local utilizando git:

```
$ git clone https://github.com/AtlasOfLivingAustralia/ala-install.git
$ cd ala-install
```

Agora tem de criar um novo inventário (inventory), um inventário seu, para ajustar a instalação ao seu servidor. Para fazer isso, deve copiar o inventário demo-ec2 e substituir o nome.

```
$ cd ansible/inventories
$ cp demo-ec2 ala-server
```

Depois disso, tem de editar o novo inventário e substituir as direções IP pelo IP que tem o seu servidor. Pode fazê-lo utilizando o editor Vim e estes comandos:

```
%s/old_direction/new_direction/
$ vim ala-server
```

e neste caso particular, deve usar o seguinte comando:

```
:/%s/ala-demo.org/myserver.com/
```

Neste instante, pode gravar o ficheiro e pode utilizá-lo para correr o ansible. Execute-o na diretoria ansible:

```
$cd ..
```

Mas antes disso, tem de fazer o seguinte: copiar a authorized_key para a diretoria /root/.ssh/:

```
server$ cp /home/ubuntu/.ssh/authorized_keys /root/.ssh/
```

Finalmente, pode instalar o seu playbook particular apontando para o seu servidor remoto:

```
$ansible-playbook -i inventories/ala-server ala-demo.yml –private -key
/tmp/xx.pem -u root
```

3.5.3 Como instalar os serviços em diferentes servidores utilizando playbooks Ansible

Com esta ideia, pode modificar outros inventários para instalar cada serviço que queira em cada servidor específico. Por exemplo:

```
$../ansible-playbook - i inventories/cassandra-server cas-standalone.yml
--private-key /tmp/yy.pem -v root
```

o inventário cassandra-server deve conter os comandos necessários para instalar os seguintes papeis (roles):

- common
- java
- tomcat

- mysql
- apache
- cas

A única coisa que tem de saber é que se quiser instalar serviços noutros servidores, tem de modificar o ficheiro config do biocache-hub ou collectory para apontar as seguintes variáveis (cassandra_hosts, solr_home, list.tool.ulr, layer.service.url, service.bie.url, datasource.url...) para estes serviços nos novos servidores.

Neste momento, pode criar uma infraestrutura mais complexa, instalando os serviços em diferentes máquinas virtuais e tentando criar um Portal de Dados otimizado.

4 Configuração

4.1 Github

Todos os módulos do software estão depositados na conta Github do Atlas of Living Australia. Vamos apresentar abaixo uma tabela com o nome dos módulos e a diretoria github onde pode encontrar o código dos elementos específicos.

Necessita seguir os seguintes passos quando inicia a instalação e configuração de um novo item:

- Ir para a diretoria específica na conta Github do ALA
- Fazer fork da diretoria para a sua conta
- Alterar o nome (por exemplo, generic-hub passou a ser gbiffrance-hub em França)
- Clone a nova diretoria para a sua máquina local
- Desfrute configurando e instalando o novo componente.

Módulo	Diretoria	Link github's
Pesquisa de ocorrências	Generic-hub	https://github.com/AtlasOfLivingAustralia/generic-hub
Collectory	Generic-collectory	https://github.com/AtlasOfLivingAustralia/generic-collectory
Hub de Dados	Generic-hub	https://github.com/AtlasOfLivingAustralia/generic-hub

Portal Espacial	Spatial-portal	https://github.com/AtlasOfLivingAustralia/spatial-portal
Lista de Espécies	specieslist-webapps	https://github.com/AtlasOfLivingAustralia/specieslist-webapp
Página de Espécies	generic-bie	https://github.com/AtlasOfLivingAustralia/generic-bie

> Nota: Este capítulo incidirá sobre a configuração e mudança de estilo do collectory e biocache-hub.

4.2 Armazenamento

4.2.1 MySQL

O nome da base de dados é "collectory" e pode encontrar abaixo o esquema da mesma. tenha cuidado com a definição do encoding, em França, houve problemas na ligação entre a variável local do nosso servidor e o encoding da base de dados. Irá necessitar configurá-lo como UTF-8 se não quiser ter erros.

Para aceder à base de dados, irá necessitar fazer a ligação ao seu servidor (como utilizador sudo) que aloja o MySQL, e depois depois fazer o seguinte:

```
$ mysql
[] use collectory; //see the database in order to query it
[] show tables; // see all tables of the database
```

Pode descarregar o DIAGRAM ER a partir deste link: https://github.com/AtlasOfLivingAustralia/collectory-plugin/raw/master/Collectory_schema_20140916.pdf

> Nota: No ficheiro de configuração `collectory-config.properties` do collectory (/`data/collectory/config/`), verifique se a ligação MySQL está bem definida em encoding como UTF-8

```
dataSource.url=jdbc:mysql://localhost:3306/collectory?
autoReconnect=true&connectTimeout=0&useUnicode=true&characterEncoding=UTF-8
```

4.2.2 Cassandra

Para aceder ao Cassandra, necessita fazer a ligação ao servidor que aloja o Cassandra e em seguida executar as seguintes linhas de comando abaixo:

```
//where occurrence data is stored
$ cd /data/cassandra/data/occ

//connect to cassandra
$ cassandra-cli
```

```
//see elements stored
[default@unknown] use occ;
// see one elements (below)
[default@occ] list occ limit 1;
// It's also possible if you want to print out a specific record
[default@occ] get occ where uuid = 'e47e0e31-ff9c-4f31-b598-34f452cb023f';
```

Na última linha um UUID é gerado automaticamente pelo Cassandra. Desde que um registo contido pelo mesmo recurso de dados determinado pelo seu druid, o UUID será estável. Ira encontrar mais informação sobre isto no site web oficial.

4.3 Arquitetura dos ficheiros

O Generic-hub e Generic-collectory são ambos baseados nos seus próprios plugins (biocache-hub e collectory). De modo a manter todo o seu trabalho quando haja uma atualização do plugin, eles criaram um sistema que ira a sua modificação antes da modificação do plugin. Pode modificar qualquer ficheiro (e.g. controladores, serviços, etc.) mas necessita seguir os seguintes passos.

Mantenha a mesma visualização de árvore tal como na do plugin.

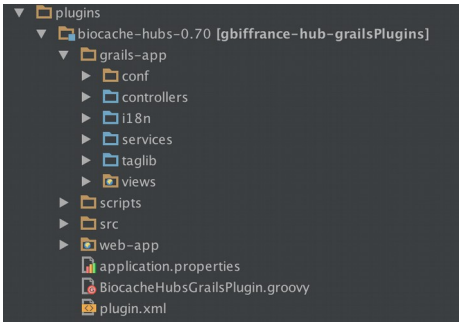


Ilustração 21: Vista de árvore do Plugin

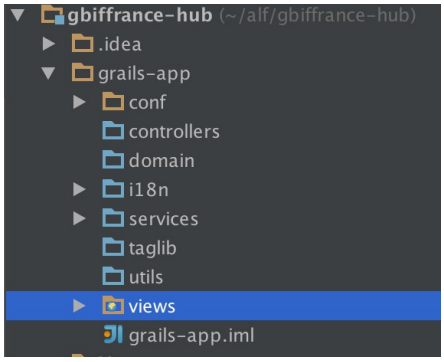


Ilustração 22: Vista de árvore GBIFrance-hub

Se quiser modificar o ficheiro `list.gsp` localizado em `grails-app/views/occurrence`, necessita de copiar este ficheiro para a diretoria correta na vista de árvore do gbiffrance-hub's grails-app. Irá seguir a mesma lógica para cada ficheiro que modifique (vista, controlador, domínio, etc.).

> Nota: Não precisa de copiar todos os ficheiros mas apenas os que deseja modificar. Se uma diretoria estives vazia, pode apagá-la da sua vista de árvore.

> Nota: se encontrar um erro (bug), pode corrigi-lo no plugin e fazer um *pull request* de modo a que a sua correção ajude toda a comunidade.

4.4 Ficheiro de configuração

Existem vários ficheiros de configuração que pode modificar de modo a alterar o seu portal:

Nome	Local	Descrição

`basisOfRecord.txt`	`/data/biocache/vocab`	Vocabulários utilizados durante a indexação do recurso de dados.
		Se tiver alguns erros de qualidade durante a indexação, pode precisar de ver este ficheiro para ter a certeza que não existe informação em falta
		
		e. g.
		`FossilRecord FossilSpecimen`
		
		`Quando, durante a indexação, o sistema encontra FossilRecord, este deve interpretar como FossilSpecimen`

`stateProvince2Countries.txt`	`/data/biocache/vocab`	Lista de estados ou províncias com o nome do país
`stateProvince`	`/data/biocache/vocab`	Lista do estado ou província
`biocache-config.properties`	`/data/biocache/config`	Ficheiro de configuração do biocache
`collectory-config.properties`	`/data/collectory/config`	Ficheiro de configuração do collectory
`generic-hub-config.properties`	`/data/yourOrg-hub/config`	Ficheiro de configuração do seu Org-hub

4.5 Definindo o estilo da aplicação

Nota: Esta é a mesma lógica do collectory e do generic-hub. Iremos focar na descrição do generic-hub

Os passos básicos são:

Obter o código a partir do GitHub

1. Fazer fork do projeto generic-hub e renomeá-lo para refletir a sua organização (e.g. : gbiffrance-hub).
2. Ir à página do generic-hub e clicar no botão "Fork" no canto superior direito da página.
3. Selecionar o seu repositório e clicar em OK.
4. Para renomear um projeto, clique no icon "Settings" no canto direito e edite o nome do projeto (e.g. yourOrg-hub).
5. Depois clone/ faça checkout ao projeto para o seu PC local.

Criar o seu layout específico

Copie o ficheiro de layout generic.gsp (`grails-app/views/layouts/generic.gsp`) para um novo ficheiro (e.g. `yourOrg.gsp`) na mesma localização

Diretamente no seu IDE (e.g.: IntelliJ)

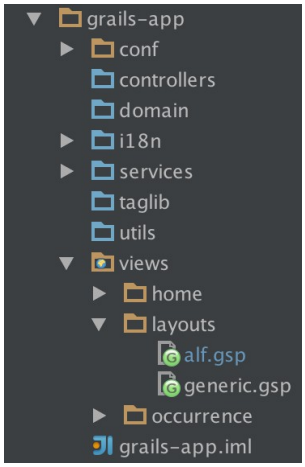


Ilustração 23: Criar o seu layout específico

ou utilizando a linha de comandos:

```
$ cd yourOrg-hub/grails-app/views/layout
$ cp generic.gsp yourOrg.gsp
```

Editar Config.groovy para utilizar o novo ficheiro de layout

Existem dois mecanismos para a configuração da app web:

- ficheiro config externo (ficheiro de propriedades) – localização por defeito:
  `/data/appName/conf/appName-config.properties`
- ficheiro config interno do Grails em:
  `/grails-app/conf/Config.groovy`

> Nota : Os valores do config externo irão ter precedência sobre os valores em Config.groovy. O ala-demo utiliza o ficheiro config externo e é o modo recomendado para definir os valores de configuração.

Altere as seguintes variáveis (ou no ficheiro externo ou no ficheiro interno):

```
skin.layout = 'generic'
skin.orgNameLong = 'Generic Data Portal'
```

faça isto:

```
skin.layout = 'yourOrg'
skin.orgNameLong = 'Your Org Name'
```

Exemplo de um ficheiro interno:

```
/*****************************************************************************
 *  SKINNING
 \*****************************************************************************/
skin.layout = 'alf'
skin.orgNameLong = "Atlas of Living France"
skin.orgNameShort = "Occurrences"
// whether crumb trail should include a home link that is external to this webat
skin.includeBaseUrl = true
//skin.headerUrl = "classpath:resources/generic-header.jsp" // can be external U
//skin.footerUrl = "classpath:resources/generic-footer.jsp" // can be external U
skin.fluidLayout = true // true or false
chartsBgColour = "#FFFFFF"
// 3rd part WMS layer to show on maps
map.overlay.url = ""
map.overlay.name = ""
biocache.baseUrl="http://recherche.gbif.fr/biocache-service"
```

Ilustração 24: Ficheiro interno

> Nota : pode querer apontar a app para uma versão local do biocache-service através da adição/edição da linha biocache.baseUrl

```
biocache.baseUrl = "http://yourOrg.org/biocache-service"
```

> Nota : se utilizar um ficheiro de propriedades externo, então omita as aspas à volta dos valores string.

Editar o ficheiro de layout yourOrg.gsp

O ficheiro de layout é um ficheiro GSP, que é semelhante ao formato de um ficheiro Java JSP, com algumas diferenças menores (ver documentação do Grails). O Grails usa a biblioteca SiteMesh para fornecer os componentes HTML comuns para as páginas.

> Nota : O generic-hub (~ biocache-hubs plugin) usa o framework Bootstrap CSS, pelo que existem alguns elementos HTML que necessitam de estar presentes em todas as páginas para estas serem representadas corretamente.

Criar um ficheiro CSS (e JS opcionais) alterado

O Grails fornece um mecanismo para gerir os recursos estáticos (CSS, JS e imagens), chamado Resources plugin, e recomendamos que se utilize esta caraterística.

Os recursos são declarados no ficheiro:

```
yourOrg-hub/grails-app/conf/ApplicationResources.groovy
```

com os ficheiros relacionados sendo geridos como módulos. Os módulos podem depender de (*dependOn*) outros módulos, que fornecem um modo de assegurar que estes ficheiros dependentes aparecem antes de ser chamado o ficheiro (e.g. um plugin jQuery ira dependOn jQuery e então jQuery será mostrado na página antes do plugin).

O ficheiro de layout contém a etiqueta (*tag)* requeridos que declara quais os módulos que são requeridos para todas as páginas que utilizam este layout particular. E.g.

```
<r:require modules="bootstrap2, hubCore" />
```

Para adicionar um novo módulo, simplesmente adicione a referência a esse módulo:

```
<r:require modules="bootstrap2, hubCore, youOrg" />
```

Ilustração 25: Adicione um novo módulo

e depois adicione o módulo ao ApplicationResources.groovy:

```
yourOrg {
      dependsOn 'bootstrap2', 'hubCore' //
      resource url: [dir:'css', file:'yourOrg.css']
      resource url: [dir:'js', file:'yourOrg.js']
      resource url: [dir:'images', file:'twitter.png']
      resource url: [dir:'css/images', file:'arrow_stat_grey_right.pgn']
}
```

Como exemplo, veja como a página institucional foi alterada:

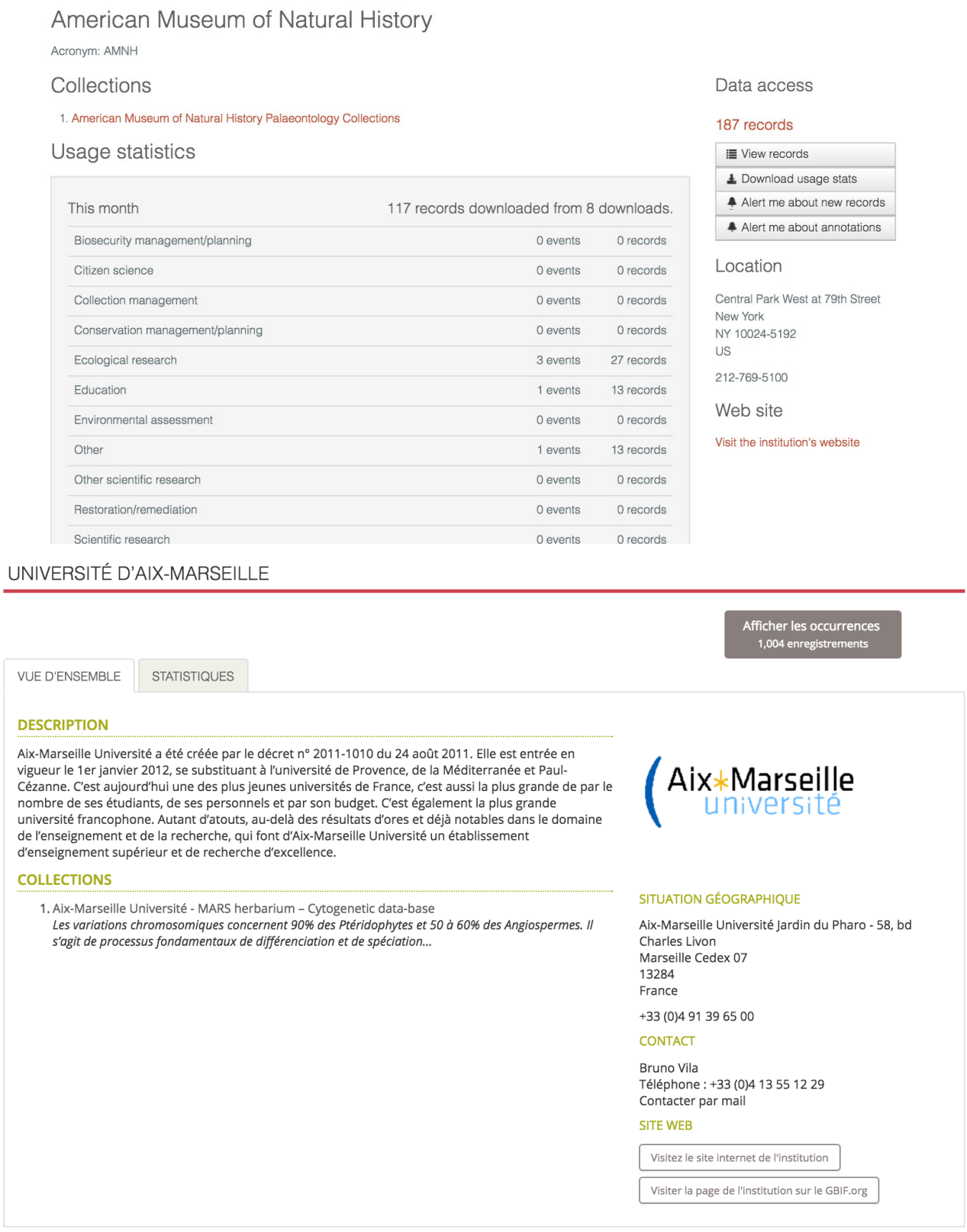


Ilustração 27: Página institucional (França)

Outro exemplo mostra a configuração da página de resultados:

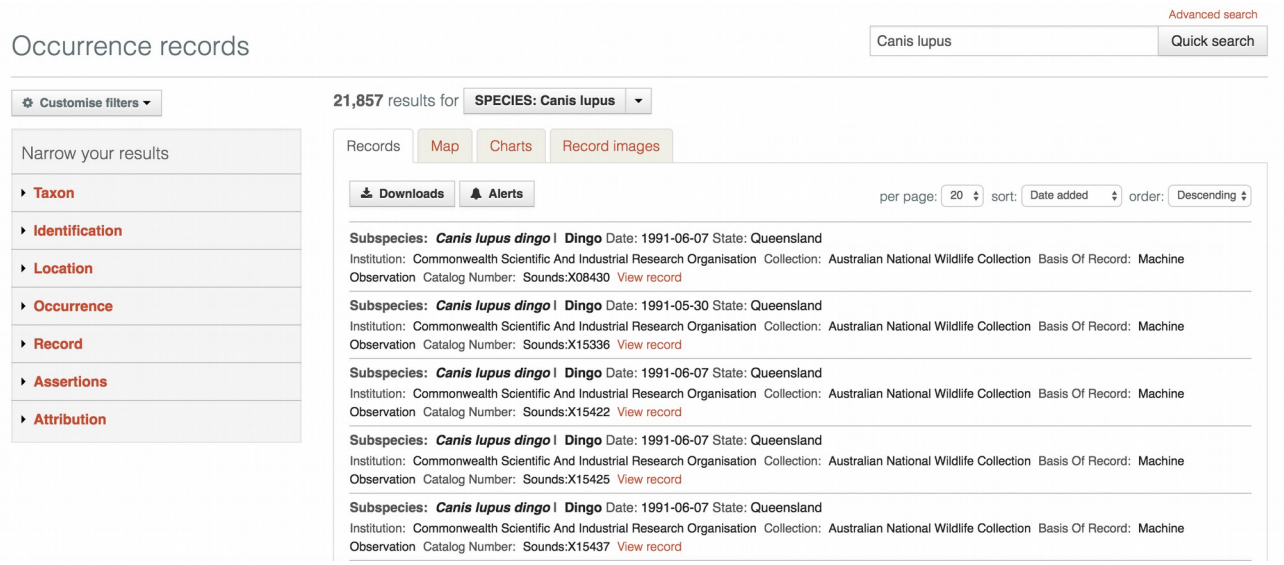


Ilustração 28: Página de resultados (Austrália)

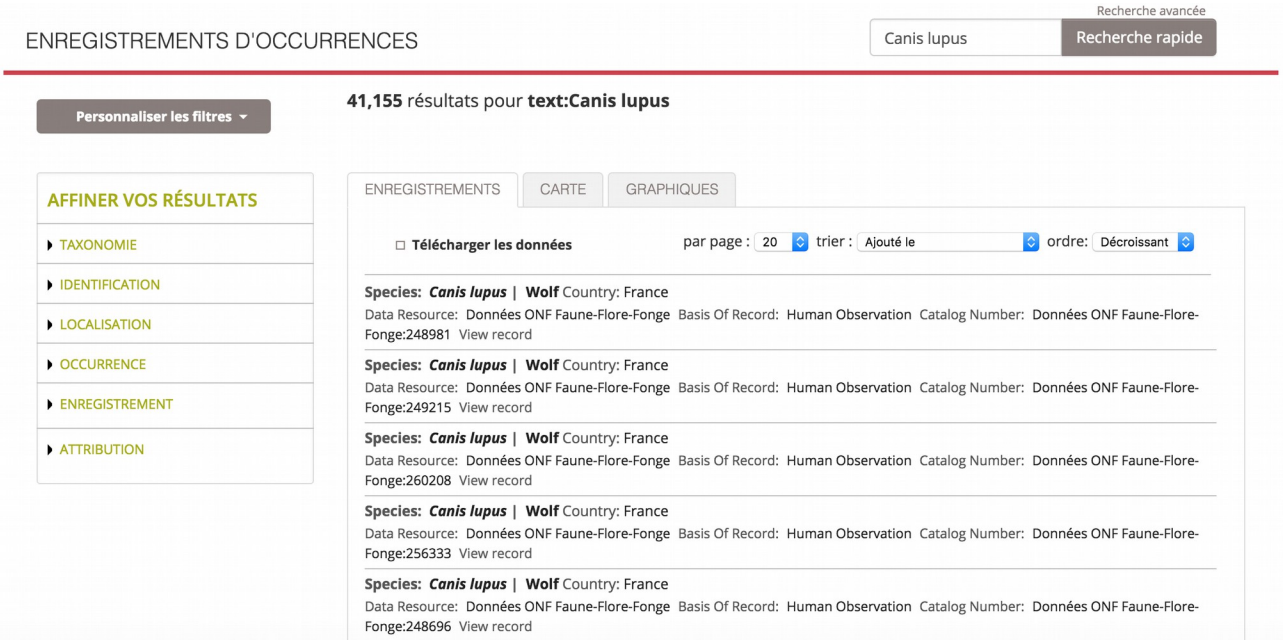


Ilustração 29: Página de resultados (França)

4.6 Internacionalização

Convidamos a ler o capítulo Internationalização para mais informação sobre este tópico. Existem alguns módulos já preparados para serem traduzidos, como o Generic-collectory e Generic-hub mas isto é possível para outros que

ainda não estão preparados, porque o código necessita de ser internacionalizável.

5 Carregamento de dados

Entre os dados manipulados, alguns são adicionados pelo administrador, outros são objetos criados pelo GBIF.org, tal como os que estão alojados nos IPTs.

O carregamento dos dados no portal ALA é uma das permissões para administradores, sendo os dados de diferentes tipos:

- Metadados: coleções, instituições, provedores de dados, hubs de dados, contactos, recursos de dados
- Dados de configuração: providerCode, providerMaps
- Ocorrência: Recursos (incluindo recursos GBIF)

Ira encontrar abaixo um esquema mostrando como os tipos de dados se relacionam.

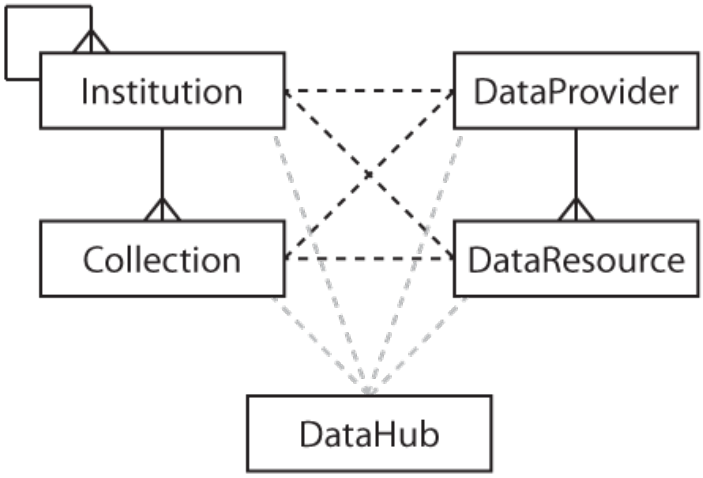


Ilustração 30: Diagrama de relações das entidades

1. Uma Instituição pode ter várias Coleções.
2. Uma Coleção pode ter apenas uma Instituição.
3. Um Provedor de Dados pode ter muitos Recursos de Dados, muitas Coleções ou muitas Instituições.

4. Um Recurso de Dados pode ter um ou mais Instituições e um Provedor de Dados, mas pode estar ligado a muitas Coleções e muitas Instituições.
5. Um Hub de Dados pode ter muitas Coleções, Instituições e Recursos de Dados.

5.1 Tipos de dados

A adição de objetos no portal é levada a cabo através das páginas de administração, pelo que o primeiro passo é fazer login para poder vê-la. A figura abaixo é a alteração feita pelo GBIF França, mas pode ver o mesmo menu que no software original.

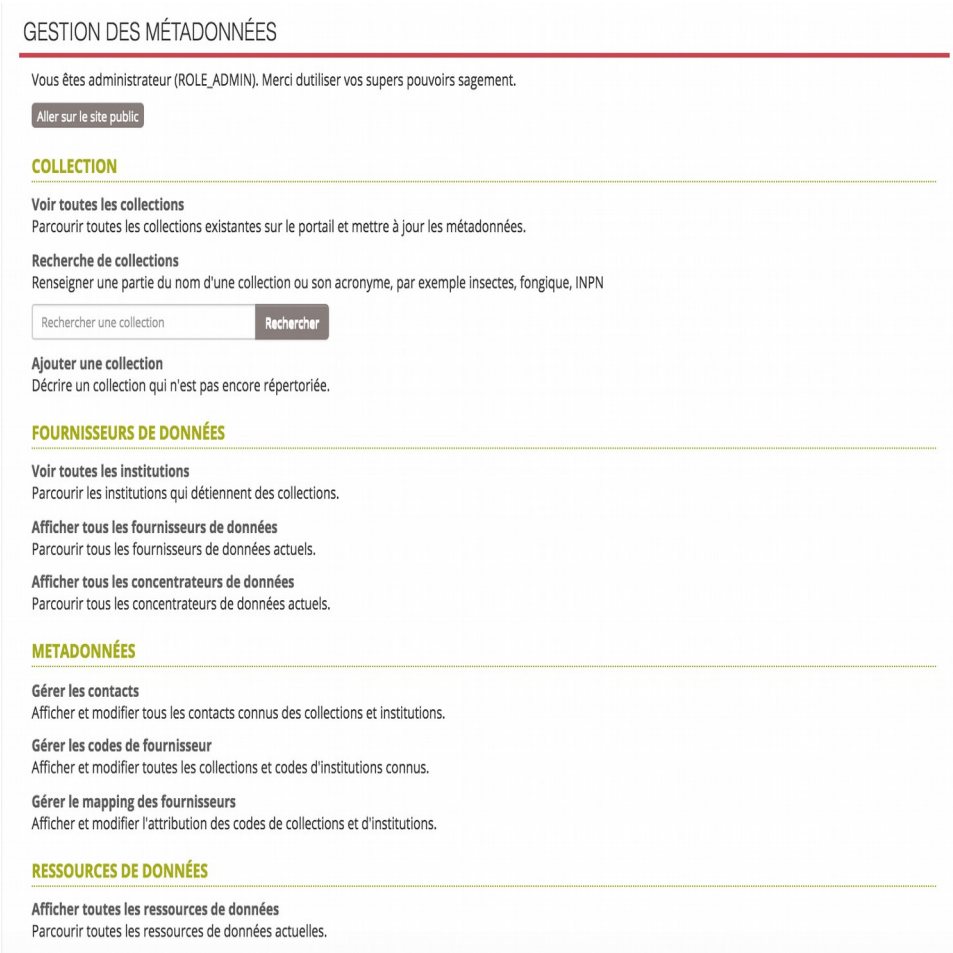


Ilustração 31: GBIF França

Pode adicionar uma coleção diretamente nesta página utilizando o link adicionar coleção. Para os outros objetos, terá que ir através da página "Ver todos [nome_Object]". Um link "Adicionar [name_Object]" irá ser mostrado e levá-lo-á ao formulário de adição do objeto.

5.1.1 Coleção

Para criar uma nova coleção, pode utilizar quer o link "Adicionar uma coleção" na página de administração ou na página de visualização de coleções; ou pode criar uma coleção quando cria um novo provedor de dados.

5.1.2 Provedor de dados

Um provedor de dados pode ser uma instituição ou um hub de dados; uma instituição é um museu, com uma coleção privada ou pública; um hub de dados é uma estrutura que tem múltiplos dados.

5.1.3 Instituição

Para criar uma nova instituição, será solicitada informação tal como: sítio da coleção, descrição, tipo, coleções, logótipo, localização, provedores de dados relacionados, contactos, etc.

5.1.4 Hub de dados

Para criar um hub de dados, é necessário adicionar informações básicas como nome, url, etc., mas será também necessário especificar as coleções, instituições e recursos de dados associados a este Hub. Para saber sobre mais configurações e como funciona, convidamos a ler o capítulo sobre a arquitetura.

5.1.5 Contacto

Pode criar contactos diretamente no menu dedicado a estes, mas também pode criar o contacto quando ao criar uma instituição, uma coleção, um provedor de dados ou um hub de dados.

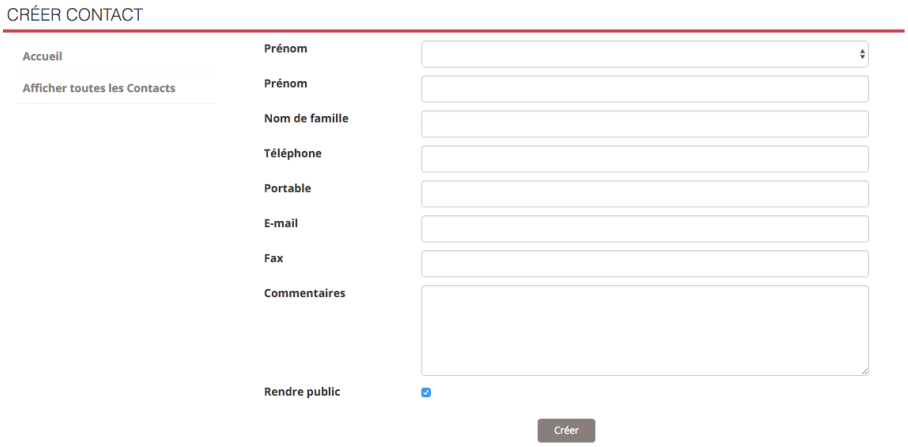


Ilustração 32: Formulário de contacto

Pode também adicionar um papel a um contacto, uma vez que esta é uma variável abstrata à qual pode adicionar conceito que quiser e pode atribuir responsabilidades, por exemplo de gerir um recurso, ser informado sobre qualquer anotação feita por qualquer ocorrência, ou se este é o contacto principal do recurso. Se sim, o seu nome estará visível na página pública do recurso.

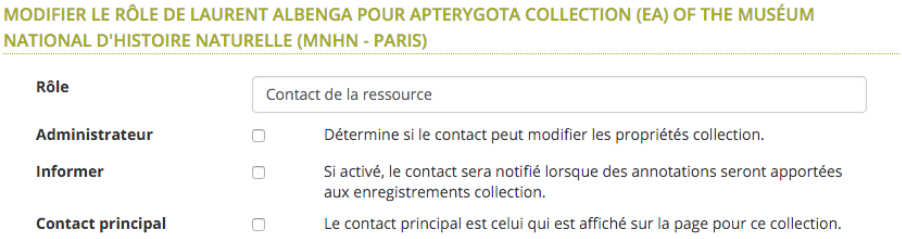


Ilustração 33: Contactos do recurso

5.2 Dados de configuração

5.2.1 Código do Provedor

Preencher aqui o Código do Provedor com o Código da Instituição e o Código da Coleção que irá encontrar nas ocorrências do seu recurso de dados.

5.2.2 Mapa de Provedor

De modo a ligar o recurso de dados à coleção e/ou instituição, necessita criar o mapa de provedor antes de realizar a indexação das ocorrências.

Id	7
Collection	Aix-Marseille Université - MARS herbarium – Cytogenetic data
Codes de l'institution	• Aix-Marseille Université
Codes de la collection	• MARS Herbarium Cytogenetic Database
Exact	Oui
Avertissement	
Correspond à n'importe quel Code de Collection	Non
Dernière mise à jour	2015-05-22 18:09:51.0
Date de création	2015-05-22 18:09:51.0

Editer Supprimer

Ilustração 34: Mapa de Provedor 1

Basicamente, precisa de informar o nome da instituição, nome da coleção (não obrigatório), código da instituição e código da coleção, e seguidamente, pode criar o mapa de provedor.

Institution ---- sélectionner une institution -----

Collection ---- sélectionner une collection -----

Codes de l'institution
38
agoralogie
AIX
Aix-Marseille Université
Ananas

Codes de la collection
38
agoralogie
AIX
Aix-Marseille Université
Ananas

Exact ☑

Avertissement

Correspond à n'importe quel Code de Collection ☐

Créer

Ilustração 35: Mapa de Provedor 2

5.3 Recurso de dados

Existem duas formas de adicionar recursos de dados ao portal: através de um ficheiro de dados ou através de um conjunto de dados disponível online num outro nó da rede GBIF.

5.3.1 Recurso de dados externo

Os recursos de dados de ocorrências é carregado como um ficheiro em formato Darwin Core Archive; este ficheiro deve conter apenas um recurso.

Adicionalmente, a este recurso é atribuído um UID (Unique Identification Number) para o identificar entre os outros recursos do portal e na página em que é mostrada a informação do ficheiro.

De modo a esta página ser mostrada no portal, os dados devem ser integrados (ingest).

Isto é feito através do terminal da instância da máquina virtual de Ubuntu, enquanto administrador, ativando a consola de da aplicação biocache, de modo a fazer finalmente o 'ingest' do recurso. Para fazê-lo, deve executar em modo sudo.

```
$ sudo su
$ biocache
biocache > ingest -dr UID_of_resource
```

Se o recurso de dados for demasiado grande, deve dividir o processo da ferramenta ingest em três comandos diferentes:

```
biocache > load  UID_of_resource
biocache > process -dr UID_of_resource
biocache > index -dr UID_of_resource
```

5.3.2 Recursos GBIF

O recurso de um ponto nodal do GBIF pode também ser acedido e assimilado pelo portal.

Os passos iniciais são diferentes do processo anterior: na página de admin do portal é necessário preencher o formulário (utilizando as credenciais de login da sua conta gbif.org), selecionar o país e o número de recursos necessário.

Uma vez submetido o formulário, o processo de descarregamento dos recursos inicia imediatamente.

As instituições, organizações ou indivíduos que detêm os dados e dão acesso a estes são os provedores de dados.

5.3.3 Usando o IPT

Existe uma terceira opção para importar os recursos de dados hospedados por um IPT para o portal: pode recolhê-los diretamente do IPT do provedor.

Para isso, deve fazer o passo inicial de criar um Provedor de Dados.

Na área de administração, vá a “Ver todos os provedores de dados” e seguidamente a "Adicionar um novo Provedor de Dados".

O nome e URL do IPT será solicitado na secção "insira o nome"; seguidamente, clique em "Selecione endpoint" para a integração, na secção IPT.

Para completar a criação do novo Provedor de Dados, atualize os recursos de dados ("Atualize recursos de dados").

Uma vez concluído, o IPT e o portal ALA estão ligados. O passo seguinte é fazer o ‘ingest’ dos dados.

Para isso, abra um terminal (disponível em sistemas Unix OS) e escreva o seguinte linha de comandos:

```
$ sudo biocache
```

para tornar-se administrador na máquina biocache.

```
Biocache > ingest -a
```

para fazer que a máquina biocache integre todos os dados.

6 Cópia de segurança e atualização

6.1 Fazer cópia de segurança das bases de dados

6.1.1 Cópia de segurança da base de dados MySQL

A cópia de segurança e a reposição da base de dados mysql é bastante simples de ser realizada. Pode utilizar o comando mysqldump como indicado abaixo:

```
$ mysqldump --lock-all-tables collectory > dump_mysql.sql --default-
character-set=UTF-8

// collectory = name of the database
// dump_mysql.sql = output file
```

Para restaurar a base de dados usando o MySQL é muito simples. Deve apenas seguir este passo

```
$ mysql < dump_mysql.sql
```

Nota: não esqueça de especificar a base de dados que deseja utilizar!

6.1.2 Cópia de segurança dos dados de Cassandra

Fazer uma cópia de segurança de Cassandra é essencialmente efetuar um snapshot do armazenamento de dados. O Cassandra vem com alguns comando que podemos usar para esta tarefa: cassandra-cli e nodetool.

Ira encontrar mais informação sobre o Cassandra no capítulo 2 sobre a arquitetura.

6.1.3 Fazer um snapshot

Por defeito, esta diretoria requer autorização de root, pelo que necessita de sudo.

```
// Assuming it's the first time we make a snapshot,
// this directory should be empty before we do
$ cd /data/cassandra/data/occ/occ && ls

// Make a snapshot of occ, which we store occurrence data
$ nodetool snapshot occ
```

A diretoria `1406163740504` é criada em `/data/cassandra/data/occ/occ/snapshots`. Se fizer a lista de ficheiros na diretoria `1406163740504`, irá notar que tem os mesmos ficheiros que em `/data/cassandra/data/occ/occ`. `1406163740504` é onde realizou a cópia de segurança das ocorrências e `/data/cassandra/data/occ/occ` é onde os dados de ocorrências serão restaurados.

6.1.4 Restaurar um snapshot

Para utilizar este snapshot numa nova máquina, seguir estas instruções:

1. Desligar o cassandra:

```
$ service cassandra stop
```

2. Limpar todos os ficheiros na diretoria commitlog directory:

```
$ rm /data/cassandra/commitlog/*
```

3. Mover os dados do snapshot

```
$ mv /data/cassandra/data/occ/attr/occ-migration-20160113/*
/data/cassandra/data/occ/attr/
$ mv /data/cassandra/data/occ/dellog/occ-migration-20160113/*
/data/cassandra/data/occ/dellog/
$ mv /data/cassandra/data/occ/distribution_outliers/occ-migration-
20160113/* /data/cassandra/data/occ/distribution_outliers/
$ mv /data/cassandra/data/occ/duplicates/occ-migration-20160113/*
/data/cassandra/data/occ/duplicates/
$ mv /data/cassandra/data/occ/loc/occ-migration-20160113/*
/data/cassandra/data/occ/loc/
```

```
$ mv /data/cassandra/data/occ/occ/occ-migration-20160113/*
/data/cassandra/data/occ/occ/
$ mv /data/cassandra/data/occ/occ_duplicates/occ-migration-20160113/*
/data/cassandra/data/occ/occ_duplicates/
$ mv /data/cassandra/data/occ/occ_outliers/occ-migration-20160113/*
/data/cassandra/data/occ/occ_outliers/
$ mv /data/cassandra/data/occ/outliers/occ-migration-20160113/*
/data/cassandra/data/occ/outliers/
$ mv /data/cassandra/data/occ/qa/occ-migration-20160113/*
/data/cassandra/data/occ/qa/
$ mv /data/cassandra/data/occ/qid/occ-migration-20160113/*
/data/cassandra/data/occ/qid/
$ mv /data/cassandra/data/occ/queryassert/occ-migration-20160113/*
/data/cassandra/data/occ/queryassert/
$ mv /data/cassandra/data/occ/taxon/occ-migration-20160113/*
/data/cassandra/data/occ/taxon/
$ mv /data/cassandra/data/occ/upload/occ-migration-20160113/*
/data/cassandra/data/occ/upload/
```

4. Reiniciar o cassandra

```
$ service cassandra restart
```

5. Correr nodetool repair.

```
$ sudo nodetool repair occ attr  dellog  distribution_outliers  duplicates
loc  occ  occ_duplicates  occ_outliers  outliers  qa  qid  queryassert  taxon
upload
```

6.1.5 Cópia de segurança do armazenamento de dados SOLR

Fazer uma cópia de segurança do SOLR é essencialmente fazer uma cópia da diretoria de dados. Irá encontrar mais informação sobre o SOLR no capítulo 2 sobre a arquitetura.

```
// SOLR index emplacement
$ ls /data/solr/biocache/data
index tlog
```

Analisando o interior da diretoria `data` podemos encontrar as diretorias `index` e `tlog`. É necessário fazer a cópia de segurança de toda a diretoria `data`.

```
// Making a copy of Solr index
$ cd /data/solr/biocache
$ sudo mkdir solr-index-backup
$ sudo chown tomcat7:tomcat7 solr-index-backup
```

Neste ponto, para `index` e `tlog` dentro de `solr-index-backup`, pode copiá-los de `/data/solr/biocache/data` no localhost ou de um host remoto. Uma vez que estes conteúdos estejam copiados. confirme que estes têm definidos como owner e group a entidade tomcat7, que é o user/group no Ubuntu por defeito que corre o Tomcat.

Neste momento, irá criar um núcleo Solr que utilize esta cópia de segurança, a qual pode ser trocada mais tarde. Para isso, navegue no browser até à página de administração do SOLR em `~/solr/#/~cores/biocache`, clique 'Add core' e insira os valores tal como indicado na imagem abaixo:

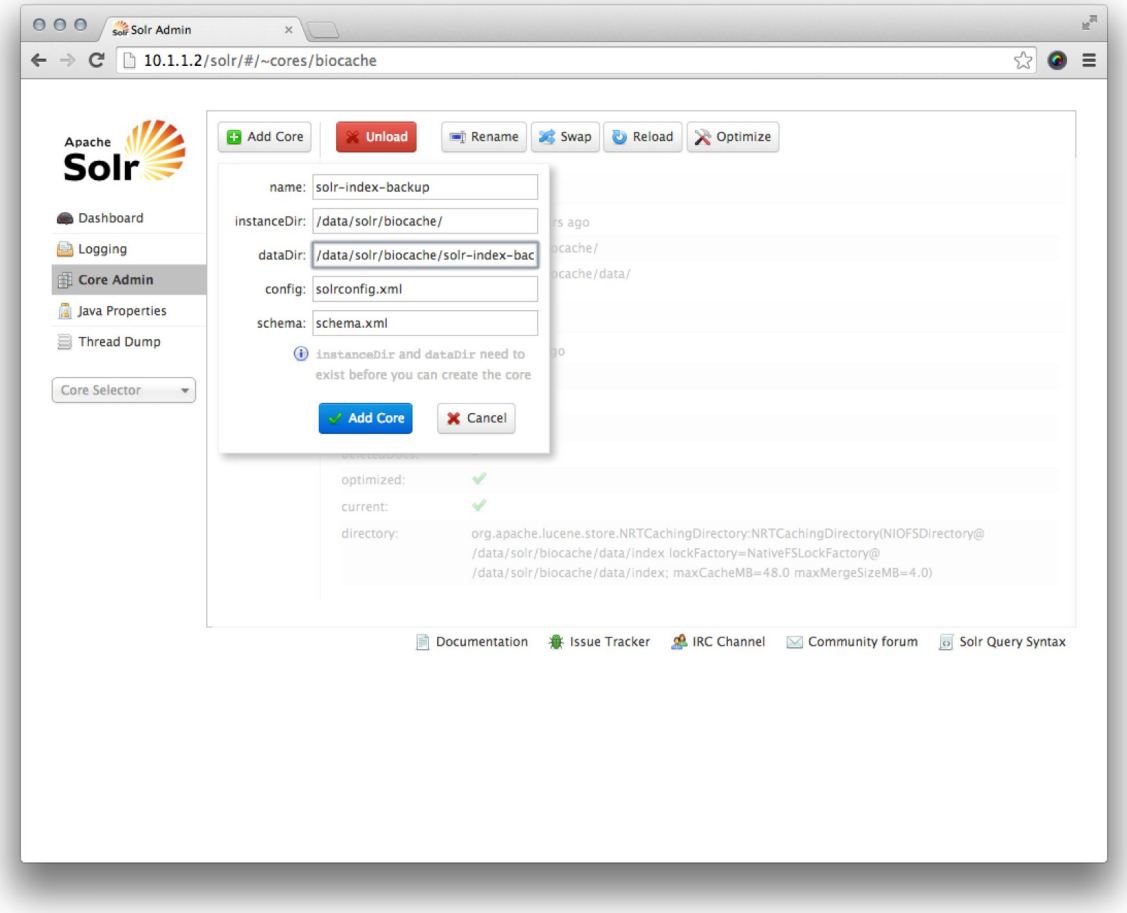


Ilustração 36: Solr

Uma vez que o núcleo tenha sido criado com sucesso, clique no novo núcleo e veja se todos os detalhes das secções 'Core' e 'Index' são os mesmos exceto para as diretorias dos ficheiros:

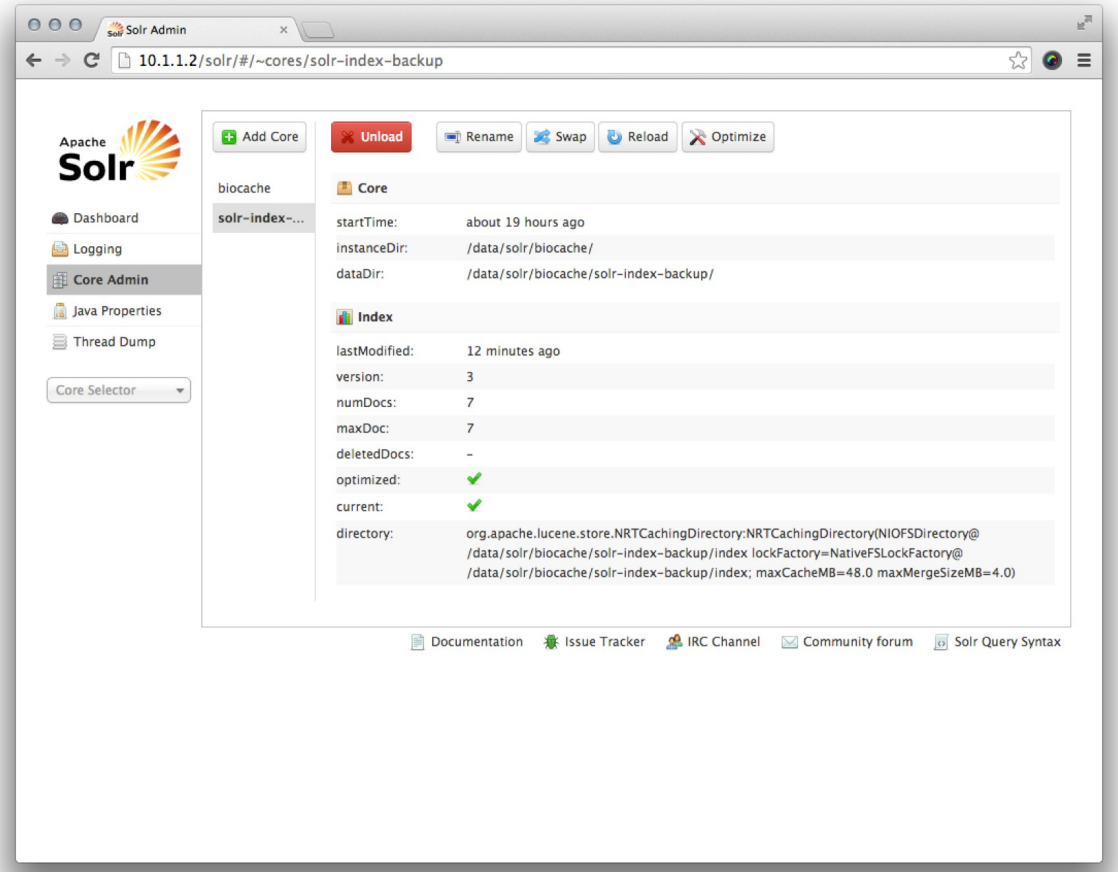


Ilustração 37: Cópia de segurança do índice Solr

Quando estiver pronto, pode usar o 'Swap' para utilizar a cópia de segurança.

6.2 Ferramentas de atualização

Graças aos scripts de Ansible, que quiser fazer a atualização de um ou vários componentes, apenas tem que voltar a correr o script de configuração. O maior valor do Ansible é a possibilidade de correr o script tantas quantas vezes quisermos. A parte de atualização da arquitetura não será difícil.

Por exemplo, se quiser atualizar os ficheiros de configuração do backend do biocache, tem apenas de correr o seguinte comando no terminal:

```
ansible-playbook -i inventories/vagrant/gbiffrance-biocache-backend
biocache-backend.yml -u ubuntu -s --tags biocache-properties
```

Pode utilizar várias características associadas ao ansible-playbook como:

```
--skip-tags
the list of tags will be skip during the running
```

```
--tags
only this list of tags will be executed
```

Se quiser atualizar ferramentas como o Tomcat ou o Java, convidamos a consultar o Github do Atlas of Living Australia ou diretamente a mailing list, para ver se os módulos são compatíveis com a nova versão destes software, mas não esqueça de fazer as suas modificações num fork do projeto e mudar o nome deste (ver o capítulo configuração e modificação do estilo). Se for este o caso, pode usar os scripts Ansible porque pode executá-los tantas vezes quantas quiser.

Para mais informação sobre a atualização, convidamos a ler o capítulos de instalação. Pode utilizar esta caraterística das tags como explicado antes.

7 Funcionalidades

As funcionalidades-chave do portal ALA podem ser agrupadas como se segue:

- Visualização de dados
- Caraterísticas de importação
- Caraterísticas de exportação
- Gestão de metadados
- Deteção de relatórios de erros

  Nota: Por favor consulte o Glossário Funcional para uma lista de entidades do ALA

7.1 Visualização de dados

A visualização diz respeito a dados e vários objetos presentes no portal como: coleções (Collections), instituições (Institutions), provedores de dados (DataProvider), recursos de dados (DataResource), relatórios, eventos de auditoria. Veja Diagrama da Relação entre Entidades.

O Provedor de Dados representa a organização global que pode ser composta por várias instituições, e.g: em Espanha, o CSIC é composto por vários centros ou instituições, cada Instituição pode publicar várias Coleções, e estas coleções contêm a metainformação relacionada com o dataset global que é partilhado através de um ou vários Recursos de Dados.

Finalmente, temos uma caraterística no portal de dados que nos permite criar um Hub de Dados (DataHub): pode criar o seu próprio portal com um subconjunto de registos provenientes de vários Recursos de Dados. Estes podem estar relacionados com várias Instituições ou diferentes Provedores de Dados.

Para gerir e utilizar esta informação, existem dois papeis no portal de dados: Administradores e Visitantes, que obviamente, não têm as mesmas permissões e possibilidades de visualização:

- Os Administradores podem criar, modificar, e apagar informação relacionada com os Provedores de Dados, Instituições, Coleções, Recursos de Dados e Contactos.
- Visitantes podem ver os dados de:
    - Do **portal collectory** (dedicado aos metadados de conjuntos de dados): Os utilizadores podem ver os metadados das coleções e instituições (nome, contactos, localização, etc.). A partir destas páginas terá acesso aos dados de ocorrência filtrados por esta instituição ou coleção.
    - A **máquina de pesquisa de ocorrências**: pode fazer uma pesquisa específica e ver a lista de resultados de dados de ocorrência. Através das páginas das ocorrências, pode aceder à instituição, provedor de dados e páginas de coleções.
    - A **máquina de pesquisa de conjuntos de dados**: será capaz de ver os metadados dos conjuntos de dados (nome, contactos, localização, etc.). A partir destas páginas, será possível aceder aos dados do conjunto de dados específico.

Ilustração 38: Páginas de boas-vindas

7.1.1 Pesquisar ocorrências

A pesquisa de ocorrências pode ser feita de duas formas: pesquisa básica ou pesquisa avançada.

A pesquisa básica permite inserir na caixa de pesquisa o nome científico ou nome comum da espécie, de modo a obter as ocorrências da mesma.

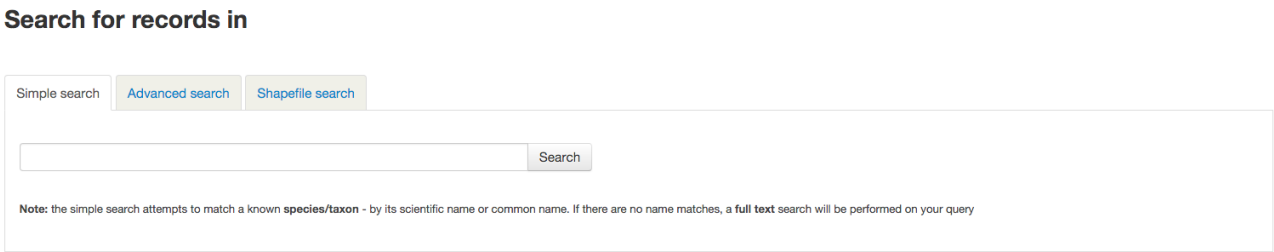


Ilustração 39: Pesquisa básica

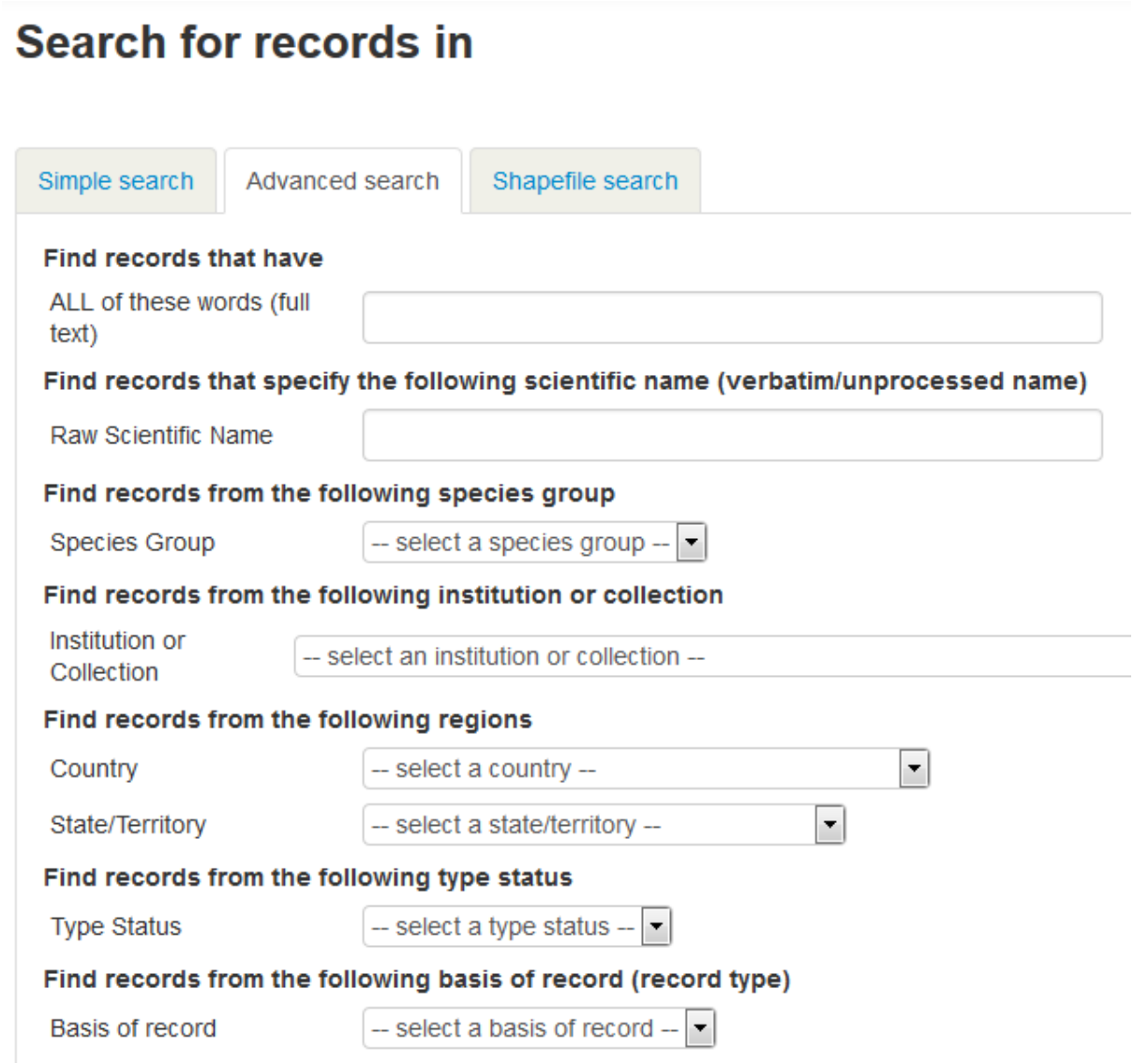


Ilustração 40: Pesquisa avançada

A Pesquisa Avançada permite adicionar critérios e filtros para limitar os resultados da pesquisa.

Podemos ver os resultados de uma pesquisa de três modos de visualização:

a vista por defeito é um mapa centrado na área contendo os resultados. Os pontos representam as ocorrências que correspondem aos critérios da pesquisa.

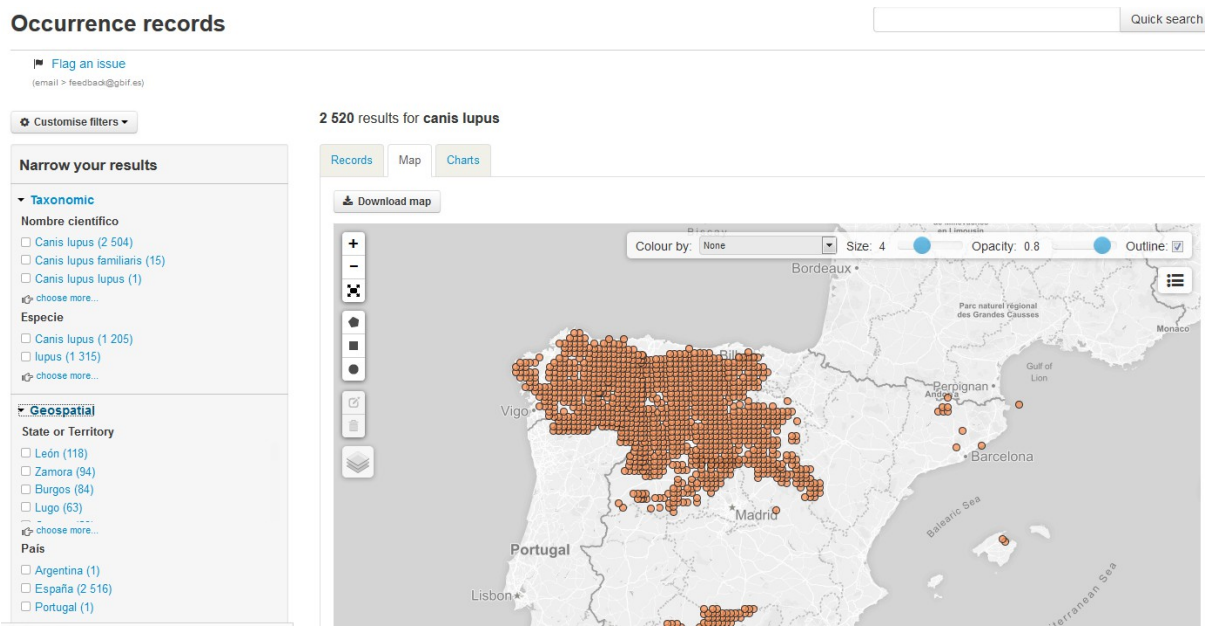


Ilustração 41: Página de resultados de uma pesquisa (Mapa)

A segunda vista é uma lista dos resultados encontrados na pesquisa. Cada elemento inclui o nome, tipo de registos, número de catálogo e uma ligação ao registo original. O administrador pode configurar esta lista de parâmetros para incluir outros itens, e.g: nome científico, nível taxonómico, data, país, nome do conjunto de dados…

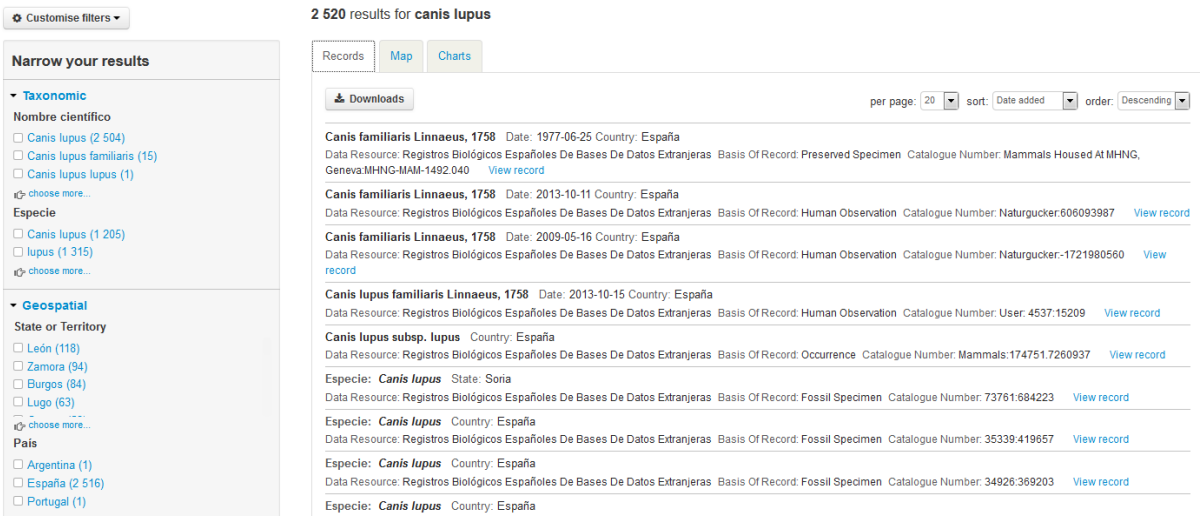


Ilustração 42: Página de resultados da pesquisa (vista de lista de registos)

A última vista mostra diagramas ilustrando a análise dos dados resultantes da pesquisa utilizando diferentes critérios: espécies, data, etc.

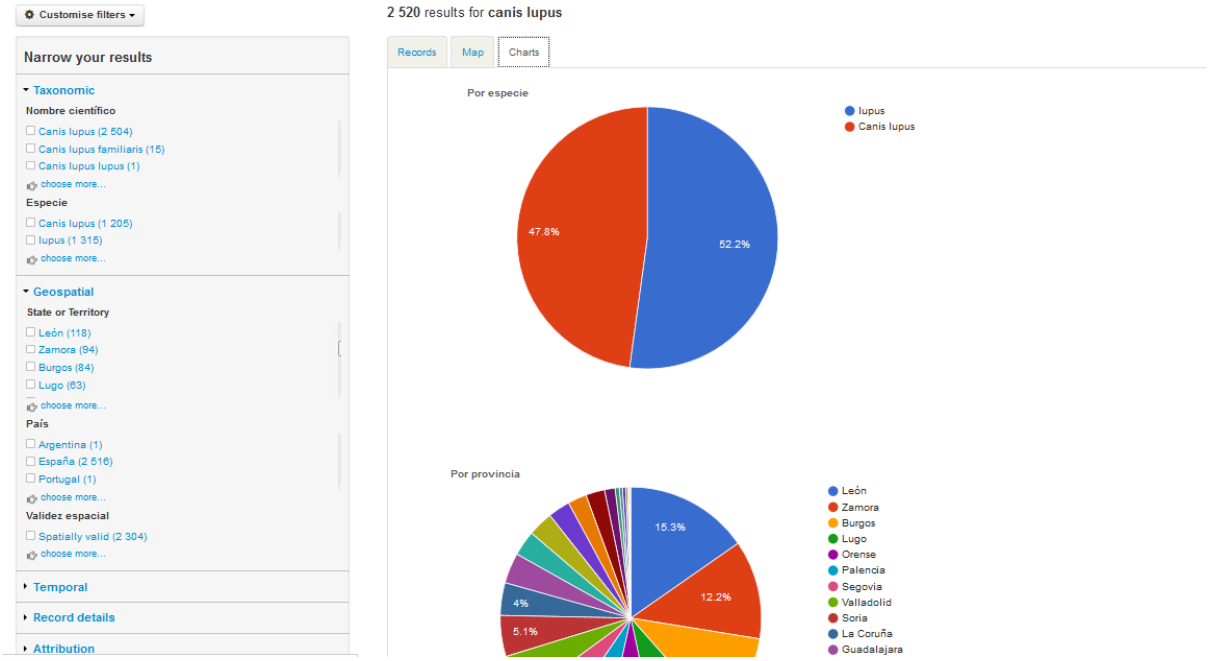


Ilustração 43: Página de resultados da pesquisa (Gráficos)

Independentemente da vista escolhida, os filtros (taxonomia, espaço, data, etc.) estão disponíveis no lado esquerdo da página para permitir ao utilizador refinar a sua pesquisa.

7.1.2 Pesquisar metadados

Coleções

Debaixo desta categoria, o visitante pode ver dados das coleções, instituições e projetos que estão relacionados.

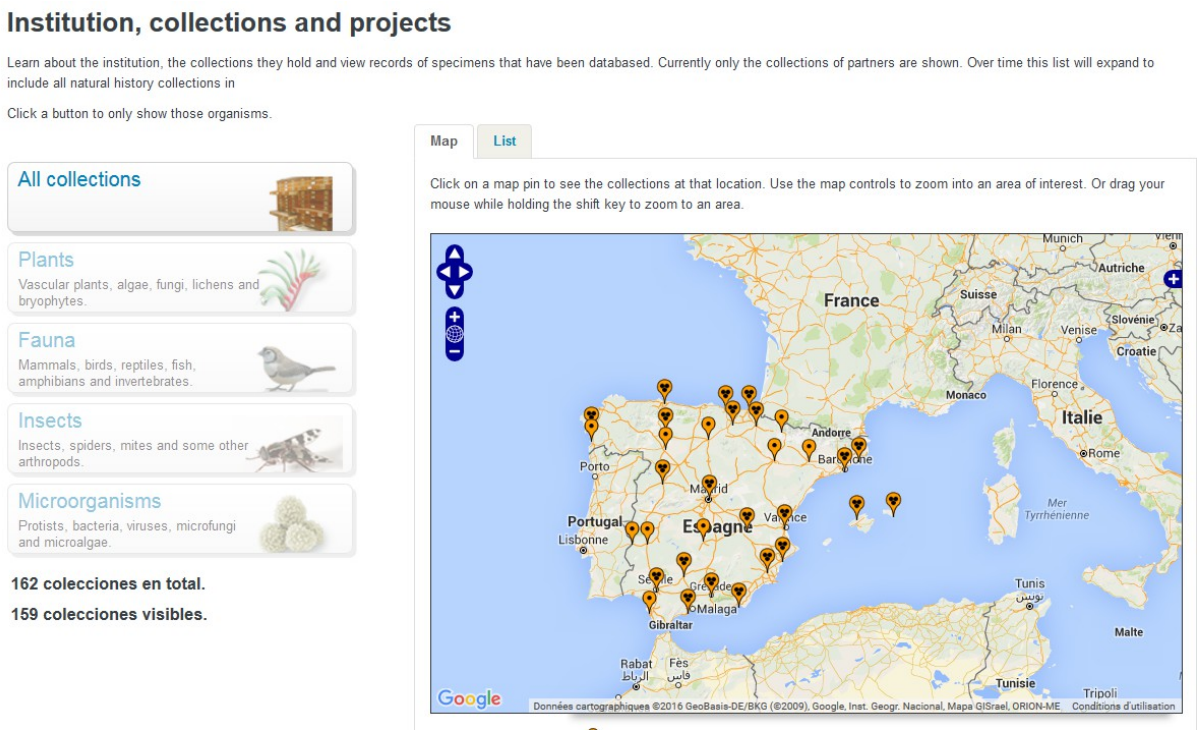


Ilustração 44: Página das Coleções (Mapas)

É mostrado um mapa global com a localização das coleções ligadas. Note-se uma vista por lista também está disponível.

Esta visualização pode ser refinada aplicando filtros com base no tipo de organismo: plantas, animais, fungos (cogumelos), micro-organismos (vírus, bactérias, etc.) e paleontologia (fosseis).

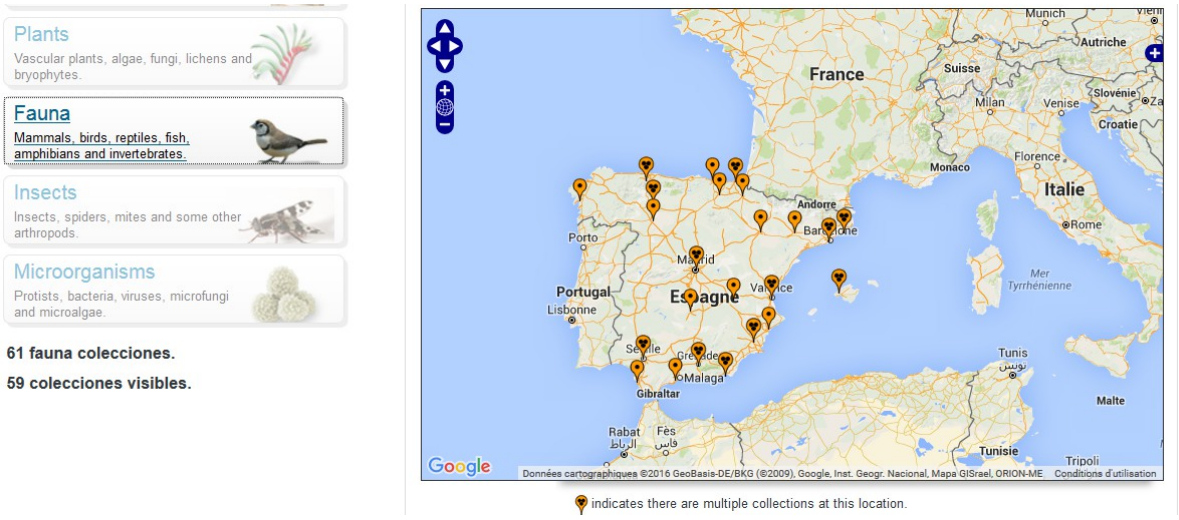


Ilustração 45: Página de filtro das coleções de fauna

Estes filtros são exclusivos uns dos outros (não cumulativos).

Conjuntos de Dados

Ao contrário das duas opções anteriores, a pesquisa por conjuntos de dados não utiliza qualquer mapa, mas sem caixas de texto ou listas de seleção para definir os critérios de pesquisa.

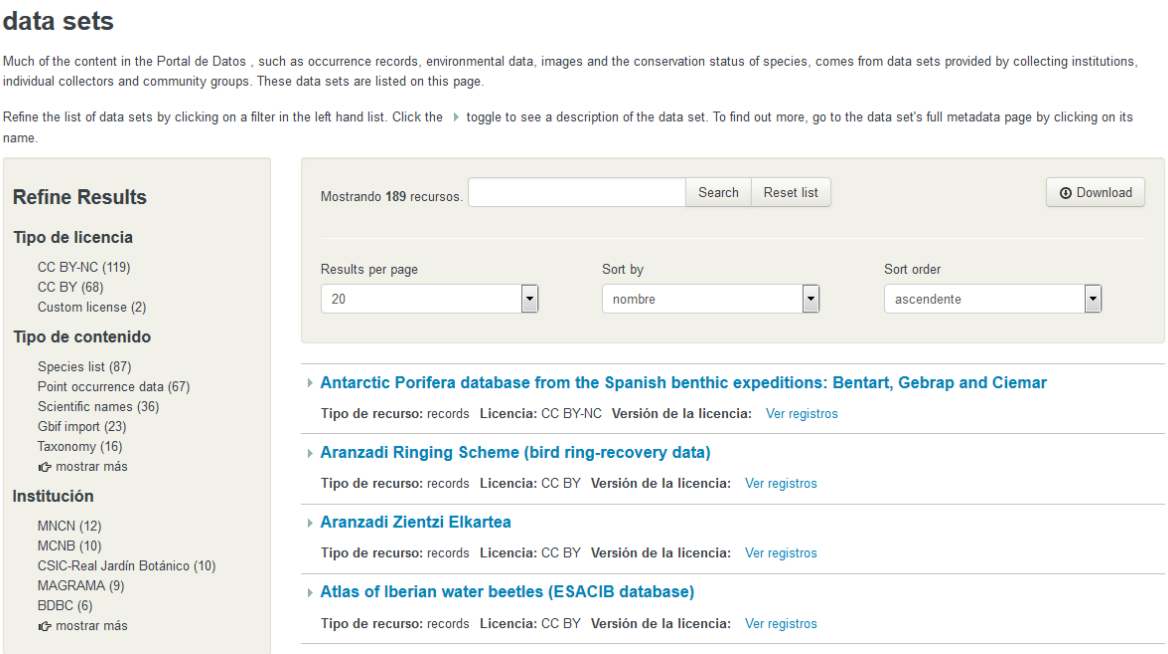
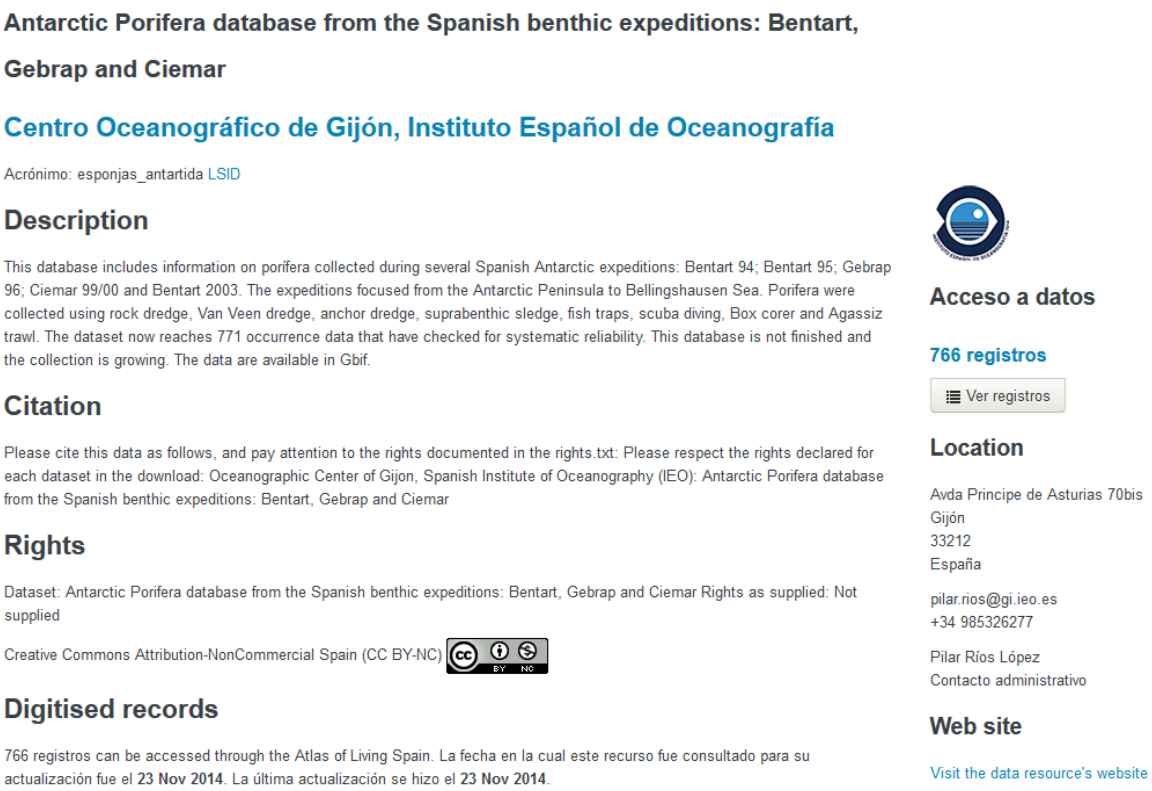


Ilustração 46: Página de Conjuntos de Dados

Os resultados são mostrados como lista dos diferentes conjuntos de dados, incluindo a informação correspondente: o tipo de recursos incluídos, a licença e uma ligação para mostrar os registos incluídos em cada um.

Cada conjunto de dados mostrado inclui um link que leva à página que mostra a informação detalhada disponível sobre este: descrição, tipo de conteúdo, lista de registos, etc.



Antarctic Porifera database from the Spanish benthic expeditions: Bentart, Gebrap and Ciemar

Centro Oceanográfico de Gijón, Instituto Español de Oceanografía

Acrónimo: esponjas_antartida LSID

Description

This database includes information on porifera collected during several Spanish Antarctic expeditions: Bentart 94; Bentart 95; Gebrap 96; Ciemar 99/00 and Bentart 2003. The expeditions focused from the Antarctic Peninsula to Bellingshausen Sea. Porifera were collected using rock dredge, Van Veen dredge, anchor dredge, suprabenthic sledge, fish traps, scuba diving, Box corer and Agassiz trawl. The dataset now reaches 771 occurrence data that have checked for systematic reliability. This database is not finished and the collection is growing. The data are available in Gbif.

Citation

Please cite this data as follows, and pay attention to the rights documented in the rights.txt: Please respect the rights declared for each dataset in the download: Oceanographic Center of Gijon, Spanish Institute of Oceanography (IEO): Antarctic Porifera database from the Spanish benthic expeditions: Bentart, Gebrap and Ciemar

Rights

Dataset: Antarctic Porifera database from the Spanish benthic expeditions: Bentart, Gebrap and Ciemar Rights as supplied: Not supplied

Creative Commons Attribution-NonCommercial Spain (CC BY-NC)

Digitised records

766 registros can be accessed through the Atlas of Living Spain. La fecha en la cual este recurso fue consultado para su actualización fue el **23 Nov 2014**. La última actualización se hizo el **23 Nov 2014**.

Acceso a datos

766 registros

Ver registros

Location

Avda Principe de Asturias 70bis
Gijón
33212
España

pilar.rios@gi.ieo.es
+34 985326277

Pilar Ríos López
Contacto administrativo

Web site

Visit the data resource's website

Ilustração 47: Página de detalhe dos conjuntos de dados

A análise dos registos e metadados associados ao conjunto de dados é também apresentada através de gráficos utilizando diferentes critérios (data, localização, etc.) para melhor visualização do conteúdo do conjunto de dados.

É também possível aceder diretamente aos registos incluídos no conjunto de dados.

Dados georreferenciados

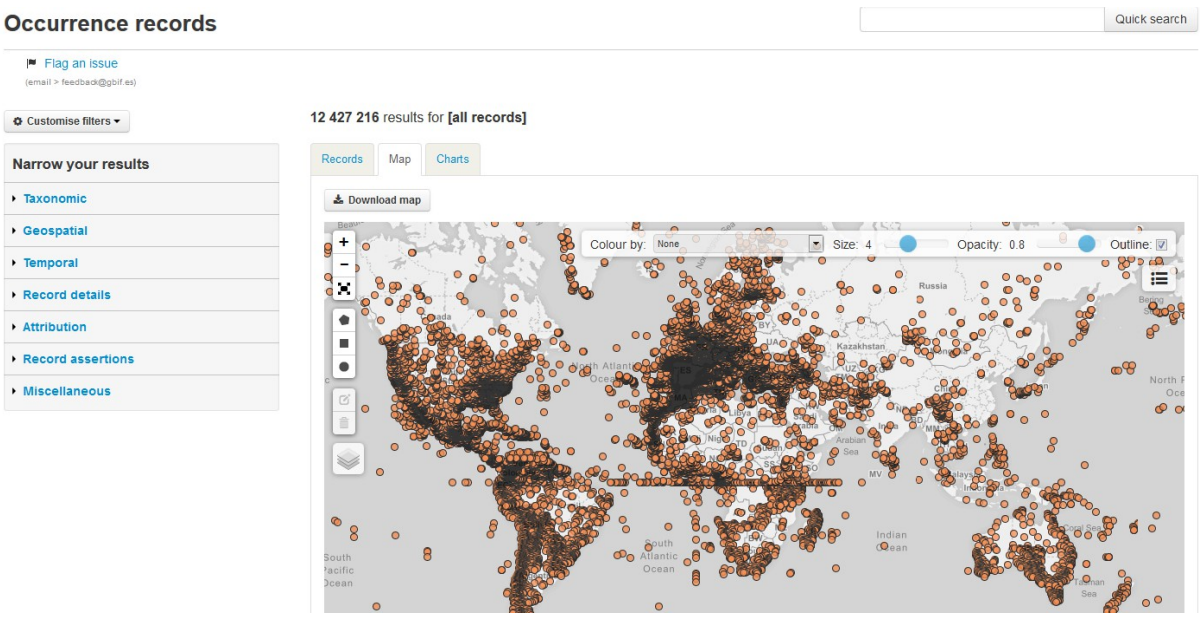


Ilustração 48: Dados georreferenciados

Este é similar à pesquisa de ocorrências, exceto que todas as ocorrências são mostradas num mapa.

Administração

Os administradores têm outras formas de ver e editar os metadados a partir da página de administração.

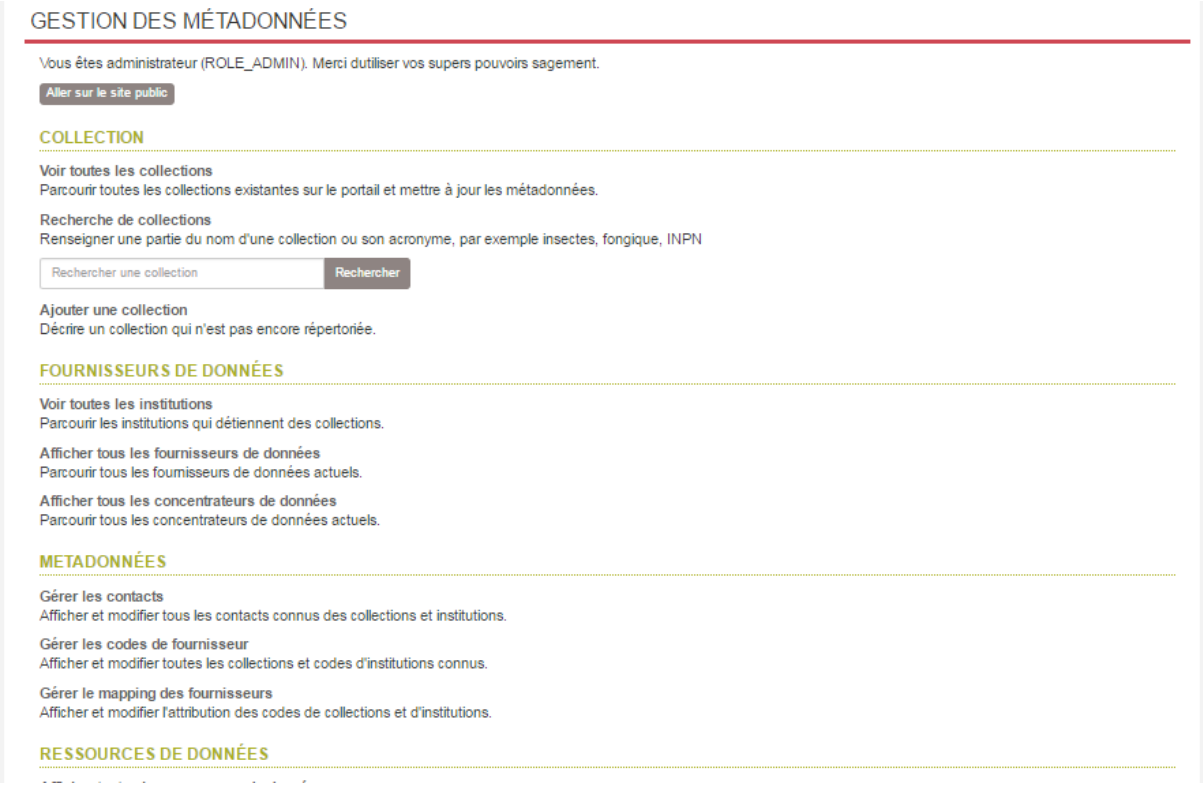


Ilustração 49: Página principal de administração

Os dados disponíveis são: coleções, provedores de dados, instituições, hubs de dados e recursos de dados. Os metadados são: contactos, códigos de provedor (provider codes), e mapas de provedor (provider maps).

Coleções

As coleções são apresentadas numa lista ordenada incluindo o nome, acrónimo (não obrigatório), tipo de instituição e coleção...

O nome da coleção e o nome da instituição incluem ligações para as páginas respetivas.

Ilustração 50: Página de informação de uma coleção

Uma sequência de letras do alfabeto no topo da página funciona como um índice e facilita o acesso à primeira coleção cujo nome começa pela letra selecionada.

Instituições

Ilustração 51: Página das Instituições

As instituições são um tipo particular de provedores de dados. A visualização numa lista contém o nome, acrónimo e tipo de cada instituição. O nome contém, para cada uma, a ligação à respetiva página que lhe é dedicada.

Provedores de dados

AFFICHER TOUTES LES FOURNISSEUR DE DONNÉESS

Accueil

Ajouter un nouveau fournisseur de données

Nom	UID	N ° ressources
Association Peau-Bleue	dp4	1
Institut National de Recherche Agronomique	dp2	1
IRD - Institute of Research for Development	dp3	0
Service du Patrimoine Naturel - Inventaire National du Patrimoine Naturel	dp0	5
Tela Botanica	dp1	0

Ilustração 52: Página de Provedores de Dados

A visualização dos provedores de dados mostra uma lista com o nome, o UID e o tipo de cada provedor de dados. O nome contém, para cada um, a ligação à sua página dedicada.

Hubs de Dados

O mesmo que na componente das "Instituições"; a diferença notável está em que os hubs de dados contém um UID em vez de um acrónimo.

Recursos de Dados

AFFICHER TOUTES LES RESSOURCE DE DONNÉESS

Accueil

Ajouter un nouveau ressource de données

Nom	UID	Type		Fournisseur de données
Actualisation de la cartographie des Ptéridophytes de France et d'Europe occidentale	dr154	records		
Aix-Marseille Université - MARS herbarium – Cytogenetic data-base	dr61	records		
Apterygota collection (EA) of the Muséum national d'Histoire naturelle (MNHN - Paris)	dr131	records		
Araneae collection (AR) of the Muséum national d'Histoire naturelle (MNHN - Paris)	dr132	records		
arim	dr14	records		
Atlas des plantes vasculaires de Lorraine	dr33	records		
Aves MZS	dr12	records		
Baillon Collection	dr19	records		
Barley French National Collection	dr22	records		

Ilustração 53: Página de Recursos de Dados

Os recursos de dados são mostrados numa lista com diferentes colunas: nome, UID, tipo de recurso de dados e provedor de dados associado. O nome de um recurso de dados leva à página que lhe é dedicada.

7.2 Funcionalidades de Importação

A importação de dados é tratada no capítulo Carregamento de dados.

7.3 Funcionalidades de Exportação

A exportação maciça de todos os dados é possível nos formatos JSON ou CSV.

7.4 Metadatos

7.4.1 Contactos

Os contactos são organizados numa lista de acordo com seis campos: ID, email, título, primeiro nome, nome e telefone.

AFFICHER TOUTES LES CONTACTS

Accueil

Ajouter un Contact

ID	E-mail	Titre	Prénom	Nom de famille	Téléphone
263	mberthel@snv.jussieu.fr		Martine Berthelin		[redacted]
264	regine.vignes_lebbe@upmc.fr		Regine Vignes-Lebbe		[redacted]
312	gbifprovider@mnhn.fr		Simon Chagnoux		
313	gbifprovider@mnhn.fr				
361	marie.meister@zool-ulp.u-strasbg.fr		Marie Meister		[redacted]
374	philippe.lebaron@obs-banyuls.fr		Philippe LEBARON		[redacted]
375	curator.mola@obs-banyuls.fr		Laurent INTERTAGLIA		[redacted]

Ilustração 54: Página de contactos

Os seus IDs estão ligados aos seus perfis individuais que contêm com dados, em adição à entidade e coordenadas, a lista de coleções onde estão assinalados como contactos.

CONTACT -

Accueil

Afficher toutes les Contacts

Ajouter un Contact

ID	158
Titre	
Prénom	CRBIP technical contact
Nom de famille	
Téléphone	
Portable	
E-mail	crbip@pasteur.fr
Fax	
Commentaires	
Rendre public	Oui
Contact pour	• Centre de Ressources Biologiques de l Institut Pasteur • Centre de Ressources Biologiques de l'Institut Pasteur

Editer Supprimer

Ilustração 55: Página de detalhes do contacto

7.4.2 Códigos de provedor

AFFICHER TOUTES LES PROVIDERCODES

Accueil

Ajouter un ProviderCode

Code
38
agoralogie
AIX
Aix-Marseille Université
Ananas
ANDUZE
Apterygota collection (EA) of the Muséum national d'Histoire naturelle (MNHN - Paris)
AR
Araneae collection (AR) of the Muséum national d'Histoire naturelle (MNHN - Paris)
arim
Atlantique Nord Est__ovipa
AUT
AUT

Ilustração 56: Páginas de código de provedor

A vista de códigos de provedor (Provider code) é uma lista de códigos de diferentes provedores.

7.4.3 Mapa de provedor

Os mapas de provedor (Provider maps) incluem informação da coleção e do provedor que a fornece.

AFFICHER TOUTES LES PROVIDERMAPS

Accueil

Ajouter un ProviderMap

Id	Collection	Exact	Égal à tous les codes collection	Code institution	Code collection
7	Aix-Marseille Université - MARS herbarium – Cytogenetic data	Oui	Non	Aix-Marseille Université	MARS Herbarium Cytogenetic Database
9	Apterygota collection (EA) of the Muséum national d'Histoire	Oui	Non	EA	MNHN
10	Araneae collection (AR) of the Muséum national d'Histoire na	Oui	Non	MNHN	AR
11	arim	Oui	Non	Institute of Research for Development	arim

Ilustração 57: Página de mapas de provedor

Uma vista em lista mostra toda a sua informação: o id e nome da coleção, o código da instituição e o código da coleção.

Duas informações adicionais são também dadas, sobre se a coleção é igual ao código da coleção no código de mapeamento.

7.5 Deteção e relatórios de erros

Uma das melhores caraterísticas deste portal de dados é a sua capacidade de detetar erros nos registos de ocorrência. Desta forma podemos melhorar a qualidade de dados da nossa informação armazenada no portal de dados. Quando os utilizadores detetam um erro, podem enviar um comentário desse erro específico, e nós podemos pô-lo em contacto com o provedor de modo a tentar resolvê-lo.

Abaixo pode ver o processo para relatar um erro no portal de dados espanhol:

1. A primeira coisa que tem de fazer é encontrar um erro no portal de dados. A melhor opção para encontrá-lo é através da vista em mapa, mas antes disso, para facilitar a tarefa, pode filtrar os registos para um conjunto de dados particular.
2. Neste exemplo, filtrámos os registos de dr190 "**NEOTROPICMYXO. A database of Myxomycetes from the Neotropics**". Este é o URL para filtrar a informação: http://datos.gbif.es/generic-hub/occurrences/search?q=data_resource_uid:dr190&lang=en#tab_mapView

3. Pode ver o resultado na vista em mapa:

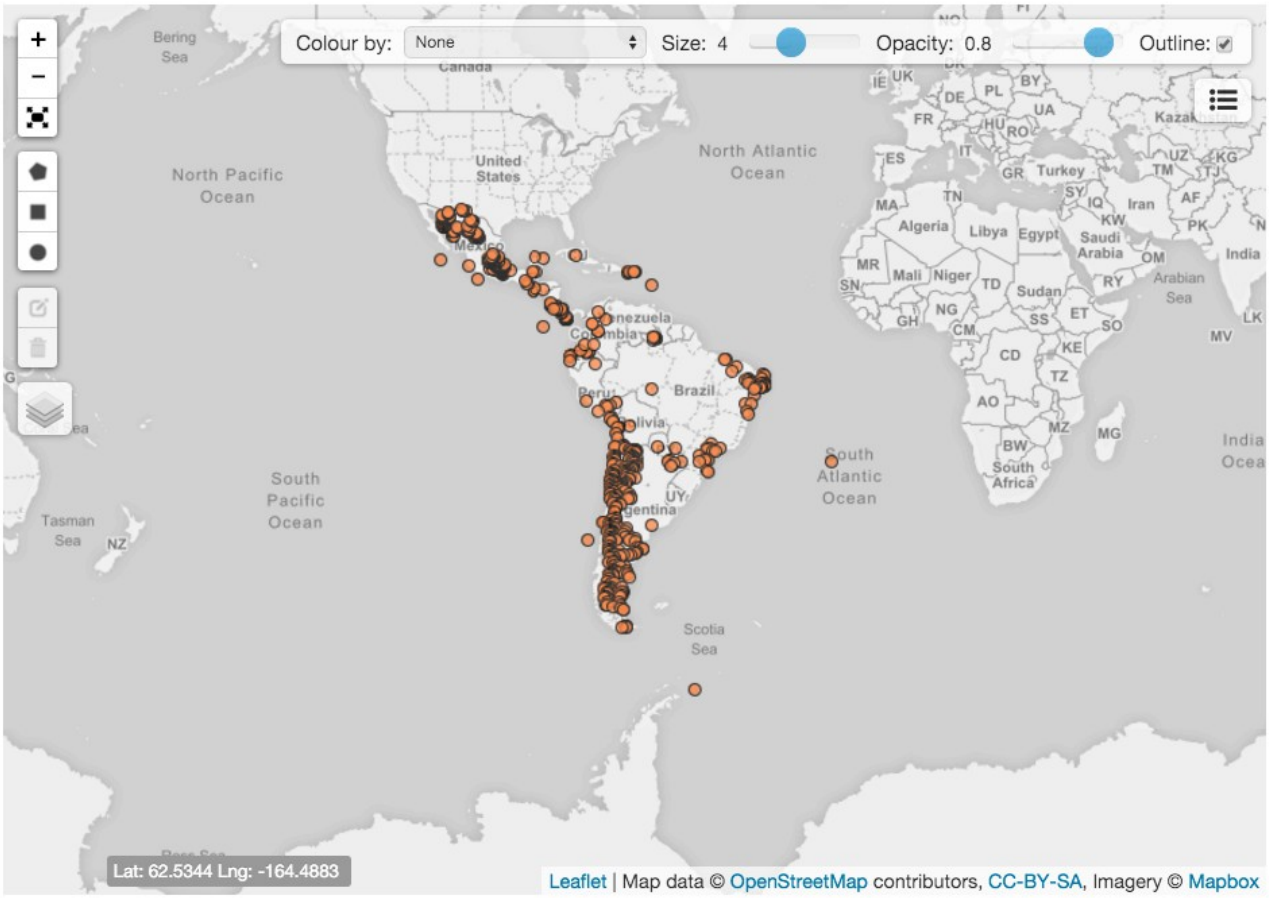


Ilustração 58: Vista em mapa

4. Os registos devem estar localizados na região dos Neotrópicos, mas nesta imagem podemos ver que pelo menos dois registos estão for a desta região. Ambos os registos são identificados pelo seu URL: http://datos.gbif.es/generic-hub/occurrences/934e4116-0de9-4f4c-bd3f-033785e940ab

http://datos.gbif.es/generic-hub/occurrences/9c60080f-69c0-442d-9478-959244954353

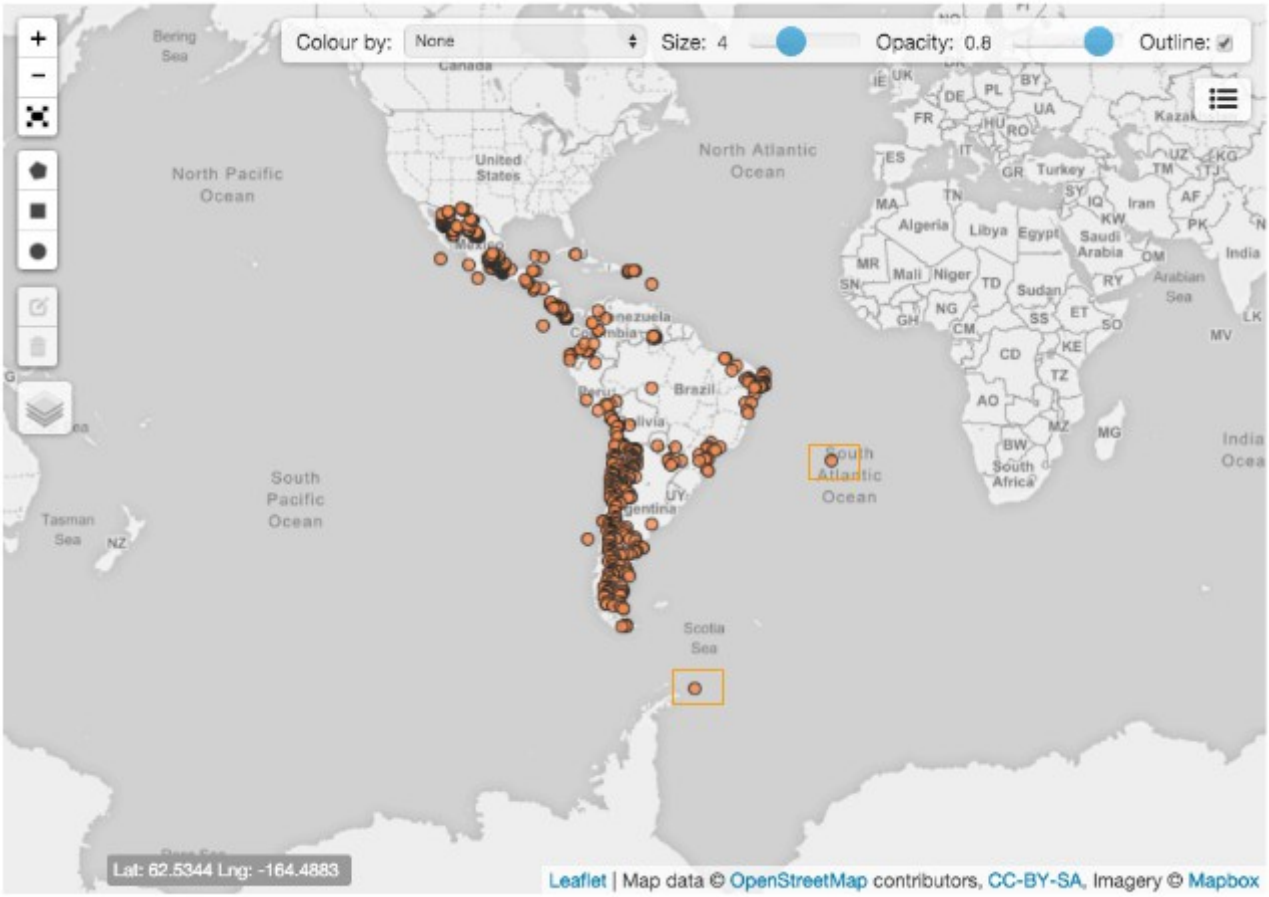


Ilustração 59: Detetar erros na vista em Mapa

5. Agora, pode enviar um email à equipa de suporte do GBIF, utilizando a ferramenta de "feedback". É muito fácil, a única coisa que tem de fazer é copiar o url onde detetou o erro e depois clicar no botão “Assinalar um problema”, que irá encontrar na vista dos registos, para o registo específico:

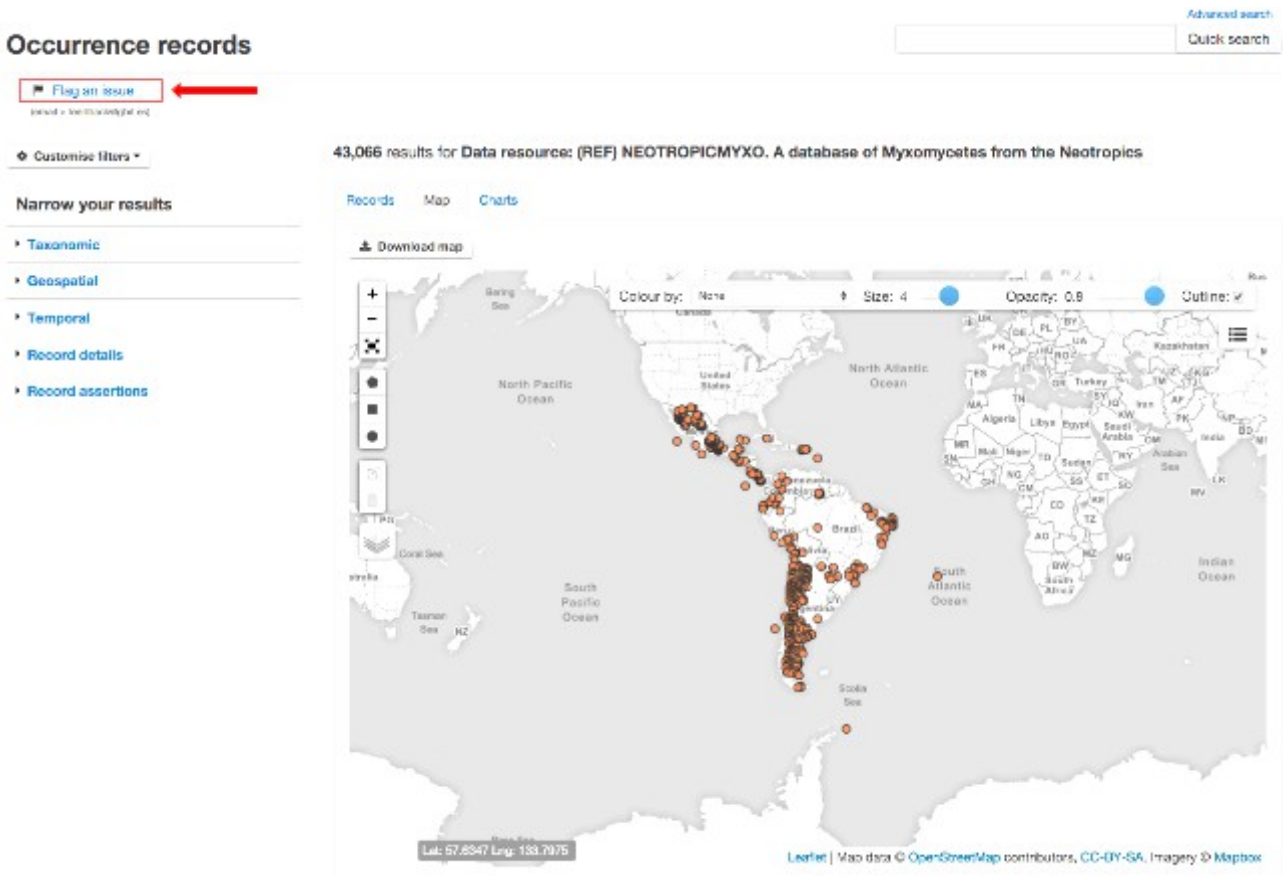


Ilustração 60: Assinalar um problema na vista por Mapa

Ilustração 61: Assinalar um problema na vista do registos

6. Quando clicar neste botão, o seu programa de email será aberto e pode enviar o relatório de erro que detetou.
7. A partir desse momento, a equipa de suporte do GBIF irá estudar o caso e pô-lo em contacto com o detentor destes registos para que sejam resolvidos os erros, nos casos em que seja necessário.

8 Leitura adicional

Para mais informação detalhada sobre a plataforma ALA, tecnologias e plataformas de desenvolvimento, padrões de dados, verifique os seguintes recursos online:

8.1 Sobre o software Atlas of Living Australia

Wiki de documentação Atlas of Living Australia

https://github.com/AtlasOfLivingAustralia/documentation/wiki

ALA Install (README file)

https://github.com/AtlasOfLivingAustralia/ala-install

Apresentações sobre o portal ALA em congressos internacionais e workshops técnicos do ALA

https://github.com/AtlasOfLivingAustralia/documentation/wiki/Presentation-ala-project

8.2 Tecnologias e plataformas de desenvolvimento

Ansible

Plataforma para a configuração e gestão de computadores

http://docs.ansible.com/

Criação de playbooks Ansible para o software ALA (ficheiro README)

https://github.com/AtlasOfLivingAustralia/ala-install/blob/master/ansible/README.md

Bootstrap

Coleção de ferramentas para criar sítios e aplicações web

http://getbootstrap.com/2.3.2/getting-started.html

GitHub

Gestão distribuída de código fonte e controlo de revisões

https://github.com/

Groovy

Linguagem de programação orientada a objetos para a plataforma Java

http://groovy-lang.org/documentation.html

Grails

Framework para desenvolvimento de aplicações Web, versão Groovy de Ruby on Rails, para utilizar no desenvolvimento de aplicações Java

https://grails.org/wiki/Documentation

http://grails.org/doc/2.3.x/guide/

Vagrant

Criar e configuram ambientes de desenvolvimento

https://www.vagrantup.com/docs/

Tomcat

Servidor Web em código aberto com implementação de tecnologias Java Servlet, JavaServer Pages, Java Expression Language e Java WebSocket

http://tomcat.apache.org/tomcat-7.0-doc/index.html

8.3 Gestão de dados

Apache Cassandra

Sistema de gestão de base de dados distribuída em código aberto

http://cassandra.apache.org/

MySQL

Sistema de gestão de bases de dados relacionais em código aberto

https://dev.mysql.com/doc/

https://www.mysql.com/

Apache Solr

Plataforma de pesquisa de nível Empresarial em código aberto

http://lucene.apache.org/solr/

http://wiki.apache.org/solr/

8.4 Formatos de dados de biodiversidade

Darwin Core

Padrão de referência estável para partilha de informação sobre diversidade biológica

http://rs.tdwg.org/dwc/terms/

GBIF (2010). Darwin Core Archives – How-to Guide, version 1, released on 1 March 2011, (contributed by Remsen D, Braak, K, Döring M, Robertson, T), Copenhagen: Global Biodiversity Information Facility, 21 pp, acessível online em:

http://links.gbif.org/gbif_dwca_how_to_guide_v1

GBIF - Darwin Core Archive Assistant

Ferramenta para apoiar na preparação de um ficheiro Darwin Core

http://tools.gbif.org/dwca-assistant/

GBIF Community Site

http://community.gbif.org/

9 Internacionalização

O Grails usa a forma padrão do Java para assegurar a internacionalização – utilizando `messages.properties` e o `g:message` taglib para serem utilizados nas páginas GSP. A deteção da língua pode ser lançada pelo browser fornecendo um valor local e fornecendo um pedido adicional de um parâmetro chamado lang. e.g. adicionar `?lang=es` ao pedido do URL.

A aplicação web cliente (e.g. generic-hub) herda algumas propriedades de internacionalização (i18n) do plugin biocache-hubs, o qual, por sua vez, herda mais propriedades do biocache-service (via uma chamada webservice). As propriedades i18n específicas podem ser adicionadas/sobrepostas pela inclusão do ficheiro de mensagens apropriado. Uma lista completa das traduções dos códigos i18n (e em Inglês) podem ser vistos em:

```
~/generic-hub/messages/i18n/messages_en-US.properties
```

9.1 Sobre a Interface do Utilizador do portal ALA

Existem muitas formas de incluir a i18n no seu portal de dados. A forma mais usada é incluir o suporte ao seletor de diferentes línguas pelo portal de dados no cabeçalho . Aqui pode ver alguns exemplos:

Ilustração 62: Exemplos

9.2 Guia para a tradução do UI passo a passo

A primeira coisa que deve fazer para traduzir a interface é traduzir o seu `messages.properties` usando o portal Crowdin. O Crowdin é uma plataforma de gestão de projetos de localização e ferramenta de tradução para programadores e detentores de sítios web. O nosso trabalho torna-se fácil com o conteúdo sendo traduzido através desta plataforma de colaboração.

Vamos então supor que queremos suportar na nossa interface a língua Português do Brasil.

Primeiro deve registar-se em https://crowdin.com. Pode usar uma destas contas:

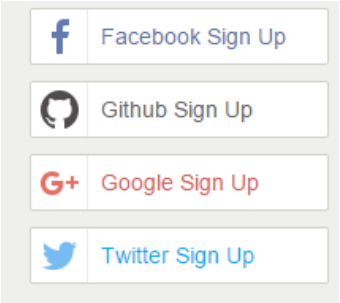


Ilustração 63: Login crowdin.com

ou criar uma nova conta crowdin em https://crowdin.com/join

O próximo passo que deve fazer é aceder à página do projeto: https://crowdin.com/project/ala-i18n e pedir o acesso ao projeto a um dos gestores do projeto que pode ver no canto superior direito da página:

Ilustração 64: Gestores

Quando tiver acesso como editor/tradutor, pode iniciar a tradução para uma língua específica. Para fazer isso, pode aceder à página do projeto que selecionou e tem de escolher uma língua que pretenda traduzir:

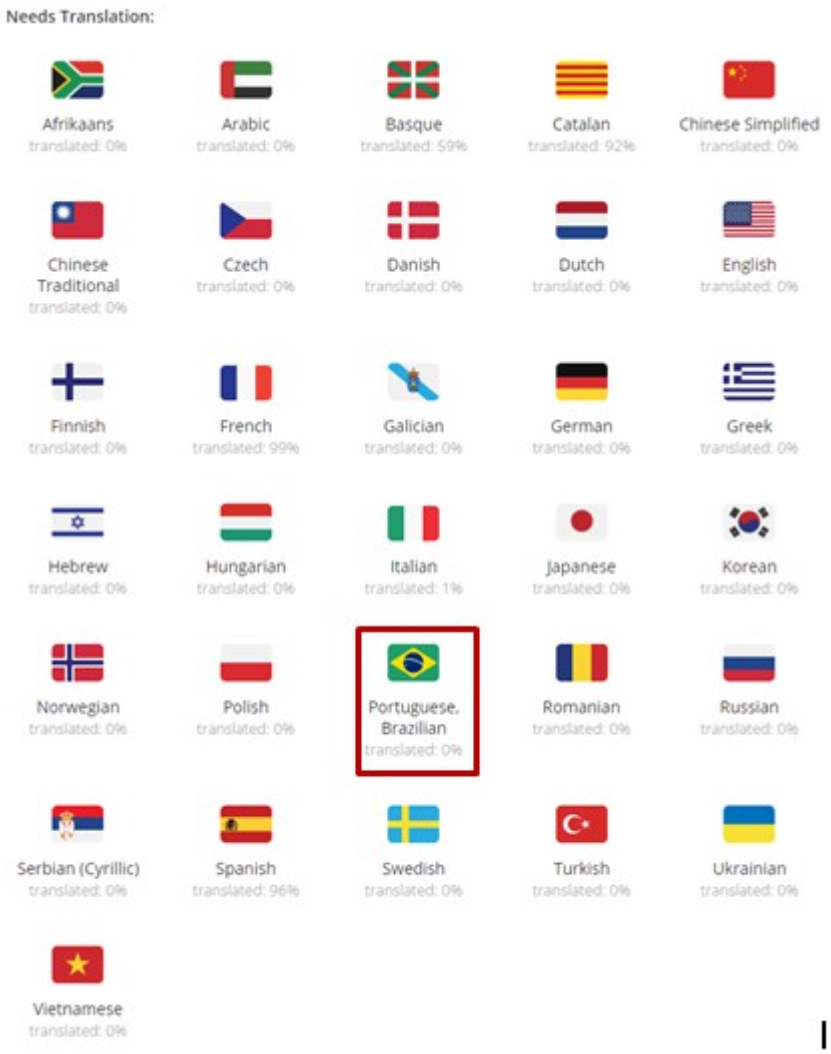


Ilustração 65: Línguas

Na imagem em baixo pode ver que temos 4 projetos e cada um tem a sua `messages.properties` que contém as propriedades que serão utilizadas nos módulos relacionados, fazendo desta forma a i18n:

Ilustração 66: Projetos

Neste momento, todas as `messages.properties` desta língua estão no nível 0 % do progresso, cujo valor irá aumentar à medida que fazemos a tradução.

Como podemos traduzir cada variável com esta plataforma? É muito fácil!!

Temos de selecionar um dos `messages.properties` prévios, e.g: biocache-hubs. A vista que temos abaixo é dividida nestas secções:

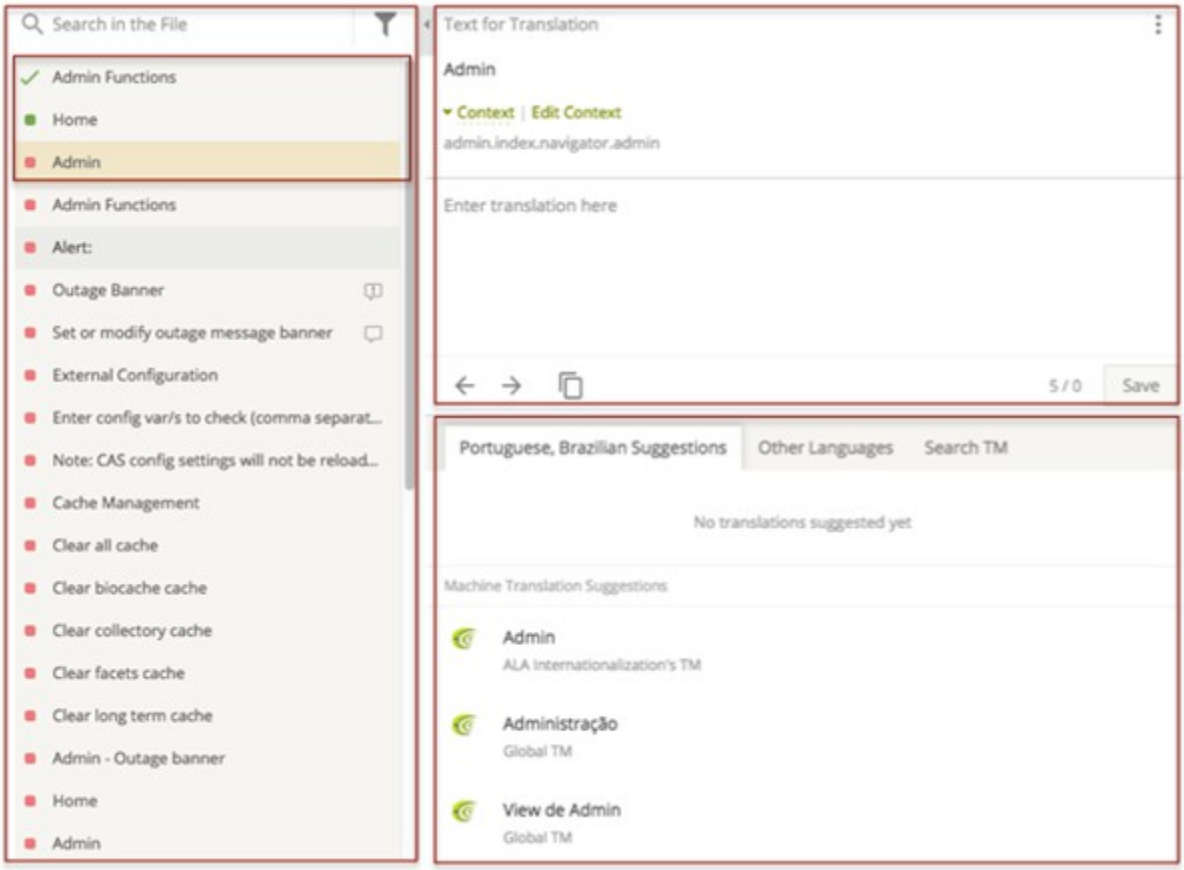


Ilustração 67: Tradução

A coluna à esquerda contém todo o texto para tradução, cada um destes relacionado com uma variável definida no ficheiro `messages.properties`. Cada variável tem três estados:

- Vermelho: não traduzido
- Verde: traduzido
- Verde claro: validado

A coluna da direita é dividida em duas secções:

- a secção de baixo sugere-nos várias possibilidades de tradução.
- e na secção superior devemos escrever a tradução final.

Uma vez escrita a tradução, devemos clicar no botão "Save", e depois a variável altera o seu estado, de não traduzido para traduzido, de vermelho a verde.

O mecanismo para fazer o resto da tradução é o mesmo para as restantes variáveis.

Outro aspeto relevante, e muito importante, é assegurar que o utilizador conhece o contexto de cada variável que está a traduzir. No caso de alguma dúvida, não hesite em contactar o gestor do projeto para obter mais informação.

Cada tradução precisa de ser validada por um gestor. Desta forma, quando tiver a certeza de que todas as traduções foram feitas de uma forma correta, deve pedir a validação por um gestor.

Quando tivermos concluído a tradução de um ficheiro `messages.properties` de um dos módulos, podemos instalar esse ficheiro no módulo respetivo. Se quiser descarregá-lo, deve enviar um pedido a um dos gestores, que irá compilar o projeto, e depois enviar o `messages.properties` final.

Podemos agora adicionar o novo `message.properties` na diretoria i18n do seu projeto. Iremos ver como fazer isso no próximo capítulo.

9.3 Configuração da língua

Os ficheiros de i18n do Grails estão em:

```
/grails-app/i18n
```

e seguem a convenção de nomeação do java de `messages_xx_XX.properties`. O ALA está interessado em obter contributos de tradutores de língua nativa não-inglesa, tanto através de pedidos de submissão de código ao projeto biocache-hubs como enviando-nos um ficheiro patch ou um ficheiro de propriedades individual.

> Nota: Se quiser contribuir no projeto i18n, por favor visite a página do projeto em: https://crowdin.com/project/ala-i18n/ e solicite a sua participação a um dos gestores.

9.4 Suporte Multi-língua

Como incluir um seletor de língua no seu projeto?

> Fonte: http://grails.org/plugin/lang-selector.

Nesta página tem as instruções para fazer corretamente a configuração do plugin lang-selector no seu módulo. Vamos, então, supor que quer instalar e configurar este plugin no generic-hub.

- passo 1: Tem de incluir na secção de plugins do BuildConfig.groovy o seguinte código:

```
runtime ":lang-selector:0.3"
```

- passo 2: Depois disto, deve adicionar o tag ao seu gsp. Por exemplo, pode incluir na secção do cabeçalho do /grails-app/views/layouts/generic.gsp para ver o seletor de língua em todas as vistas do módulo generic-hub.

```
<ul class="nav pull-right">
 <li class="dropdown">
       <a class="dropdown-toggle" data-toggle="dropdown" href="#">Language<b
class="caret"></b></a>
       <ul class="dropdown-menu">
           <li><langs:selector langs="ca"/></li>
           <li><langs:selector langs="en"/></li>
           <li><langs:selector langs="es"/></li>
       </ul>
   </li>
</ul>
```

- passo 3: Opcionalmente, pode adicionar esta propriedade ao ficheiro /grails-app/conf/Config.groovy para informar o plugin que bandeiras mostrar para que línguas:

```
com.mfelix.grails.plugins.langSelector.lang.flags = ["es":"es",
                                                     "en":"gb",
                                                     "fr":"fr",
                                                     "da":"dk",
                                                     "de":"de",
                                                     "it":"it",
                                                     "ja":"jp",
                                                     "nl":"nl",
                                                     "ru":"ru",
                                                     "th":"th",
                                                     "zh":"cn",
                                                     "pt":"pt",
                                                     "ca":"catalonia"]
```

Igualmente, pode fazer isto no módulo collectory.

Agora, vamos supor que não pretende mostrar bandeiras, possivelmente por uma questão de acessibilidade, ou por outra razão. Neste caso, tem que modificar o código deste plugin. Para fazê-lo, tem de descarregar o plugin, e chamá-lo localmente no `/grails-app/BuildConfig.groovy`, e modificar `/lang-selector-0.3/grails-app/views/langSelector/_selector.gsp`:

BuildConfig.groovy:

```
grails.plugin.location.'lang-selector' = "plugins/lang-selector-0.3"
```

_selector.gsp:

```
<div id="lang_selector">
```

```
<span style="color:white"><g:message code="1" default="|"/></span>
<g:each in="${flags.keySet().sort() }" var="lang">
<%--
      uri + lang = with this expression, we are building the url that
redirects the application to the related view in that language “lang”.

--%>
    <a href="${ uri + lang }" title="${message(code:'title.lang_link')}">
        <span class="lang_flag ${ lang==selected.toString()? selected_class :
not_selected_class }" style="margin-left: 14px;">

<%--
In this section we are checking what language we are going to apply. We have
detected it, we can specify the message that we want to show.
--%>

        <g:if test="${flags[lang] == 'catalonia'}">
                <g:message code="1" default="Català"/>
        </g:if>


        <g:if test="${flags[lang] == 'gb'}">
                <g:message code="1" default="English"/>
        </g:if>

        <g:if test="${flags[lang] == 'es'}">
                <g:message code="1" default="Español"/>
        </g:if>

<%--

      We can also include the flag of language instead of a message

--%>

        <%--<img alt="" src="$
{resource(plugin:'langSelector',dir:'images/flags/png',file:flags[lang]
+'.png') }" border="0">--%>
        </span>
    </a>
    <span style="color:white"><g:message code="1" default="|"/></span>
</g:each>
```

10 Canais de colaboração

Pode ajudar a desenvolver o software ALA, mesmo que não seja um programador. Existem várias tarefas, e certamente uma destas é adequada para si.

Associar-se às mailing lists ALA

Associe-se à ALA mailing list gerida pelo GBIF (http://lists.gbif.org/mailman/listinfo/ala-portal) para ficar a conhecer o que está a ser desenvolvido, solicitar uma nova funcionalidade, relatar problemas e entrar em contacto com o resto da comunidade do software ALA.

Pode também associar-se ao Atlas of Living Australia HipChat, pedindo para juntar-se ao gestor do projeto Atlas of Living Australia (https://atlasoflivingaustralia.hipchat.com/home) para trocar mensagens chat com outros membros da comunidade ALA.

Contratar o seu nó GBIF

Será que o seu país/organização fez uma implementação do portal ALA? Se for esse o caso, é provável que o seu nó GBIF já esteja envolvido na comunidade ALA, e pode identificar os melhores contactos para que se associe a esta comunidade, e identificar que papel pode desempenhar nesta. Veja alista de participantes GBIF.

Como contribuir – papéis na comunidade

A sua contribuição para o software ALA pode ser de diferentes formas, dependendo das suas capacidades e talento. Junte-se à ALA mailing list e diga o como gostaria de contribuir. Existem diversos papeis que pode desempenhar:

Programador de software

Se for um programados, pode contribuir por:

Ajudar a fazer com que o suporte do ALA seja global

O Atlas of Living Australia foi desenvolvido no contexto australiano. Na sua fase inicial, não foi planeado para suportar uma comunidade global, pelo que muitas partes do código necessitam de ser generalizadas a partir do contexto australiano (e.g. regiões administrativas da Austrália) para um contexto global. Igualmente, alguns dos módulos não têm o suporte à internacionalização (i18n) inserido desde o inícios. Pode ajudar fazendo com que a plataforma ALA

tenha um suporte global, o que é também uma boa oportunidade para tornar o código melhor.

Desenvolver um novo módulo do Atlas

A plataforma Atlas of Living Australia é facilmente expansível pela adição de novos plugins ou módulos que adicionam funcionalidades ou características ao software. Igualmente, a arquitetura do sistema inclui uma camada de serviços web que pode facilmente permitir a integração com outras ferramentas ou programas. Se precisar e tiver a capacidade de expandir o software ALA, através do desenvolvimento de um novo módulo ou plugin, proponha essa nova funcionalidade para ser incluído no roteiro de desenvolvimento futuro do ALA. Visite https://github.com/AtlasOfLivingAustralia/documentation/wiki/RoadMap para verificar quais os planos de desenvolvimento futuro da plataforma ALA.

Adicionar mais características a um módulo existente

Do mesmo modo, adicionando caraterísticas a um módulo, para responder a implementações específicas do portal, pode ser um projeto de pequena dimensão onde a sua ajuda pode ser valiosa.

Relatar um problema

Faça um relatório de um problema, quer através de um envio de uma mensagem para a mailing list ALA do GBIF ou preenchendo um relatório de problema no repositório GitHub do módulo ALA correspondente.

Ajudar a implementar o ALA ao nível nacional para um Nó GBIF nacional

A implementação do portal ALA a nível nacional é uma tarefa intensa, onde qualquer ajuda é bem vinda. Neste documento identificámos muitas tarefas que necessitam de ser realizadas:

- alteração das interfaces de cada módulo para o layout nacional
- gestão de dados da lista nacional de espécies
- gestão de dados dos recursos nacionais de dados

Escrever documentação

Este manual pode ser expandido, quer pelo relatório sobre outros ambientes de implementação, quer pela adição da experiência na implementação de outros módulos do ALA. Se tiver este conhecimento, partilhe com a comunidade escrevendo um capítulo sobre o mesmo.

Utilizador Atlas

Se for um utilizador do Atlas, a sua contribuição é essencial para:

Dar Feedback

Encontrou um erro nos dados, ou na forma como os dados são manuseados? O seu feedback é essencial par detetar estes problemas e promover correções. Por favor use os contactos existentes na implementação do portal para relatar um problema, o que pode ajudar a beneficiar a qualidade do portal.

Propor uma nova funcionalidade / tópico

Tem ideias de como melhorar a usabilidade do portal, ou novas caraterísticas que beneficiariam, o seu uso? Proponha-as, ajudando a expandir o valor do portal.

Escrever documentação

Desde a pequenos como-fazer a manuais, pode contribuir para criar ou expandir a documentação para os utilizadores para muitas dos dos módulos existentes do ALA. Será ainda melhor se construir essa documentação com base me casos de estudo reais em que o ALA foi útil.

Traduzir documentação

Tal como o GBIF, devemos ajudar a fazer com que a plataforma ALA se torne global, pelo suporte a diferentes línguas. Pode contribuir para isto ajudando a traduzir a documentação ou manuais sobre o uso do ALA, ou o ficheiro de recursos que alimenta a interface web do portal ALA. A tradução de ficheiros de recursos ALA é feita utilizando a plataforma web amigável Crowdin. Verifique aqui quais as línguas em que é necessária a sua colaboração.

Publicite o portal ALA

Tem conhecimento de colegas, organizações ou instituições que podem encontrar valor em utilizar ou desenvolver o portal ALA? Por exemplo, trabalha com escolas, na qual os projetos dos alunos podem ser ajudados com informação sobre biodiversidade?? Ou tem contactos com projetos de ciência cidadã que produzem informação primária de biodiversidade, mas necessitam de uma plataforma web ou aplicação móvel para gerir as observações? É membro de um projeto ou rede que necessita de uma plataforma web a nível nacional ou internacional com informação sobre biodiversidade para um tópico específico? A plataforma do ALA pode dar suporte a todos estes casos, pelo que é importante que todos os utilizadores potenciais tenham conhecimento da da disponibilidade do portal ALA para ver, aceder e desenvolver soluções sobre

dados de biodiversidade. E lembre-se, a melhor publicidade é feita da demonstração através de exemplos.

Anexo I: Glossário funcional

DataResource (Recurso de dados)	Este é a ligação do conjunto de dados ao Darwin Core Archive ou CSV carregado. Na página de administração dos recursos de dados, irá encontrar um link para o DwC Archive. Todos os dados estão ligados a um recurso de dados.
DataProvider (Provedor de dados)	DataProvider representa a organização global que pode ser composta por várias instituições. Pode ser utilizado, por exemplo, por um agregador de dados.
Contacto (Contacto)	Ligação ao utilizador do recurso de dados, coleção, instituição, ou provedor de dados. Pode atribuir permissões de editor a qualquer contacto.
Institution (Instituição)	Este é o detentor dos dados.
Collection (Coleção)	Diz respeito ao conjunto de dados de uma coleção(e.g. : espécimes de herbário). Pode estar ligado a um ou mais recursos de dados.
ProviderCode (Código de provedor)	Corresponde ao código de provedor utilizando o código da coleção e o código da instituição que encontra no seu recurso de dados.
Provider map (Mapa de provedor)	Este é o mecanismo que liga o recurso de dados à instituição e coleção, ou só à instituição.
Administrator (Administrador)	Este é uma das permissões que pode ser atribuída aos utilizadores do sistema ALA. Com esta permissão, pode gerir todo o software. O Administrador pode criar, modificar e apagar informação relacionada com os Provedor de Dados, Instituições, Coleções, Recursos de dados e contactos.
Editor	Esta é uma das permissões que pode ser atribuída aos utilizadores do sistema ALA. Com privilégios de editor, pode gerir e modificar metadados do seu próprio recurso de dados/coleção/instituição.
Metadata (Metadados)	Informação sobre o recurso de dados, coleção, instituição ou contacto.

Anexo II: Ilustrações

Índice de Ilustrações

Ilustração 1: Arquitetura ALA..... 8
Ilustração 2: Visão geral dos módulos.....9
Ilustração 3: Módulos de pesquisa de ocorrências..... 10
Ilustração 4: Módulos Collectory.....11
Ilustração 5: Módulo Espécies..... 12
Ilustração 6: Módulos Datahub.....13
Ilustração 7: Hub de Dados Myxotropic..... 14
Ilustração 8: Ocorrências Myxotropic.....15
Ilustração 9: Hub de Dados..... 16
Ilustração 10: Módulo Regiões..... 17
Ilustração 11: Módulos do Portal Espacial.....18
Ilustração 12: Portal Espacial.....19
Ilustração 13: Infraestrutura do portal de dados espanhol.....22
Ilustração 14: Infraestrutura Australiana..... 25
Ilustração 15: Infraestrutura Francesa.....26
Ilustração 16: Ubuntu.....27
Ilustração 17: Login Ubuntu..... 28
Ilustração 18: Ansible.....29
Ilustração 19: Demo ALA..... 30
Ilustração 20: Fluxo de informação.....31
Ilustração 21: Vista de árvore do Plugin..... 35
Ilustração 22: Vista de árvore GBIFrance-hub.....35
Ilustração 23: Criar o seu layout específico.....37
Ilustração 24: Ficheiro interno..... 38
Ilustração 25: Adicione um novo módulo..... 39
Ilustração 26: Página institucional (Austrália).....40
Ilustração 27: Página institucional (França).....40
Ilustração 28: Página de resultados (Austrália).....41
Ilustração 29: Página de resultados (França).....41
Ilustração 30: Diagrama de relações das entidades.....42
Ilustração 31: GBIF França.....43
Ilustração 32: Formulário de contacto.....45
Ilustração 33: Contactos do recurso.....45
Ilustração 34: Mapa de Provedor 1..... 46
Ilustração 35: Mapa de Provedor 2..... 46
Ilustração 36: Solr..... 51
Ilustração 37: Cópia de segurança do índice Solr..... 52
Ilustração 38: Páginas de boas-vindas.....55
Ilustração 39: Pesquisa básica..... 55
Ilustração 40: Pesquisa avançada.....56
Ilustração 41: Página de resultados de uma pesquisa (Mapa).....57
Ilustração 42: Página de resultados da pesquisa (vista de lista de registos)..57
Ilustração 43: Página de resultados da pesquisa (Gráficos).....58

Ilustração 44: Página das Coleções (Mapas).....59
Ilustração 45: Página de filtro das coleções de fauna.....59
Ilustração 46: Página de Conjuntos de Dados.....60
Ilustração 47: Página de detalhe dos conjuntos de dados.....61
Ilustração 48: Dados georreferenciados.....62
Ilustração 49: Página principal de administração.....63
Ilustração 50: Página de informação de uma coleção.....64
Ilustração 51: Página das Instituições.....64
Ilustração 52: Página de Provedores de Dados.....65
Ilustração 53: Página de Recursos de Dados.....65
Ilustração 54: Página de contactos.....66
Ilustração 55: Página de detalhes do contacto.....66
Ilustração 56: Páginas de código de provedor.....67
Ilustração 57: Página de mapas de provedor.....67
Ilustração 58: Vista em mapa.....69
Ilustração 59: Detetar erros na vista em Mapa.....70
Ilustração 60: Assinalar um problema na vista por Mapa.....71
Ilustração 61: Assinalar um problema na vista do registos.....71
Ilustração 62: Exemplos.....75
Ilustração 63: Login crowdin.com.....75
Ilustração 64: Gestores.....76
Ilustração 65: Línguas.....77
Ilustração 66: Projetos.....78
Ilustração 67: Tradução.....79