
MANUAL DO DATAVERSE

LUCAS N. PAGANINE
VANDERLINO C. BARRETO NETO

LUCASPAGANINE@IBICT.BR
VANDERLINONETO@IBICT.BR
[HTTPS://WWW.OVERLEAF.COM/READ/VQYTYSGJBZND](https://www.overleaf.com/read/VQYTYSGJBZND)

REVISÃO POR:
LUCIANA DOS SANTOS NAHUZ

Lista de Figuras

1	Página principal	22
2	Citar dados	22
3	Arquivos	23
4	Diagrama esquemático	26
5	Facetas	27
6	Campos	29
7	Permissões	31
8	Página permissões	31
9	Tipos de permissão	32
10	Papéis	33
11	Designando papéis	34
12	Termos de uso	35
13	Livros de visita	35
14	Esquema de um Dataset	39
15	Download	42
16	Visualização	44
17	Proveniência	48
18	Arquivo de proveniência	48
19	Versões	52
20	Dashboard	56
21	Harvest server	57
22	Página de verificação do OAI server	57
23	Criar um critério de coleta	58

Lista de Tabelas

1	metadataBlock	63
2	datasetField	64
3	controlledVocabulary	67
4	FieldType	67
5	displayFormat	68
6	Metadados mandatórios	91
7	Metadados mandatórios se aplicáveis	92
8	Metadados Recomendados	95
9	Metadados Opcionais	98

Conteúdo

1	INSTALAÇÃO	8
1.1	Pré instalações	8
1.1.1	Preparação do Servidor	8
1.1.2	Java	9
1.1.3	Glassfish	9
1.1.4	PostgreSQL(banco de dados)	11
1.1.5	Solr	11
1.1.6	Jq e ImageMagick	12
1.1.7	R	13
1.2	Instalação do Dataverse	13
1.3	Configurações	14
1.3.1	Provedor Doi	14
2	Manual do usuário: Criação de contas + Gerenciamento	16
2.1	Informações da conta	16
2.1.1	Opções de log-in	16
2.1.2	Criação de contas	17
2.1.3	Edição de contas	17
2.1.4	Conversão de contas	17
2.1.5	Resetar senha	17
2.2	Meus dados	18
2.3	Notificações	18
3	Manual do usuário: Encontrando e usando dados	19
3.1	Encontrado dados	19
3.1.1	Busca básica	19
3.1.2	Busca avançada	20
3.1.3	Navegando no Dataverse	21
3.2	Usando Dados	21
3.2.1	Ver Dataverses + Datasets	21
3.2.2	Ver arquivos	21
3.2.3	Citar Dados, versões e metadados	22
3.2.4	Download de arquivos	22
3.2.5	Explore Dados	24
4	Manual do usuário: Gerenciamento Dataverse	25
4.1	Criação de um Dataverse (Dentro do Dataverse “Root”)	26

4.2	Edite um Dataverse	28
4.3	Informação geral	28
4.4	Tema	29
4.5	Widgets	29
4.5.1	Widget Caixa de Pesquisa	29
4.5.2	Widget Lista	30
4.6	Papéis e Permissões	30
4.6.1	Definindo Configurações de Acesso	32
4.6.2	Definindo papéis para Usuários e Grupos	33
4.7	Templates de Dataset	34
4.8	Livro de visitas de Datasets	36
4.9	Dataverses destaques	37
4.10	Vinculação de Datasets	37
4.11	Vinculação de Dataverses	38
4.12	Publique seu Dataverse	38
5	Manual do usuário: Dataset + Gerenciamento de arquivos	39
5.1	Metadata suportados	39
5.1.1	Formatos de exportação de metadados	40
5.2	Adicionando um novo Dataset	40
5.3	Upload de Arquivo	41
5.4	Manuseio de Arquivo	41
5.5	Arquivos de dados tabulares	42
5.6	Arquivos compactados	42
5.7	Outros tipos de arquivo	43
5.8	Editar Arquivos	43
5.8.1	Editar Arquivo de Metadados	43
5.8.2	Caminho do Arquivo	44
5.8.3	Tags do Arquivo	45
5.9	Substituir Arquivo	45
5.10	Termos	45
5.10.1	CC0 Domínio Público Dedicação	46
5.10.2	Termos de Uso Customizados para Datasets	46
5.10.3	Arquivos Restritos + Termos de Acesso	46
5.10.4	Livro de Visitas	46
5.10.5	Nível de Dataset	46
5.10.6	Nível de Arquivo	47
5.11	Proveniência dos Dados	47

5.12	Thumbnails e Widgets	49
5.12.1	Thumbnails	49
5.12.2	Widgets	50
5.12.3	Widget dataset	50
5.12.4	Widget Citação de dataset	51
5.13	Publique Datasets	51
5.14	Submeter à Revisão	51
5.15	URL Privada para Revisão de Datasets Não publicados	51
5.16	Versões dos Datasets	52
5.16.1	Detalhes das Versões	53
5.17	Desativar acessos de Dataset	53
6	Manual do administrador	55
6.1	Painel de controle	56
6.1.1	Harvest	56
6.1.2	Usuários	58
6.2	Gerenciando clientes de coleta	58
6.2.1	Seu Dataverse como um Harvester de Metadados	58
6.2.2	Gerenciando clientes de coleta	58
6.3	Gerenciando conjuntos e servidor de coleta	59
6.3.1	Seu Dataverse como servidor OAI	59
6.3.2	Como funciona?	59
6.3.3	Conjuntos OAI	59
6.3.4	Atualizações do conjunto OAI	61
6.4	Customização de metadados	61
6.4.1	Introdução	61
6.4.2	Sobre o bloco de metadados TSV	62
6.4.3	Explorando blocos de metadados	68
6.4.4	Criação de Metadados	68
6.4.5	Editando arquivos TSV	69
6.4.6	Carregando arquivos TSV no Dataverse	69
6.4.7	Propriedades #metadataBlock	70
6.4.8	Propriedades #datasetField (campo)	70
6.4.9	Propriedades #controlledVocabulary (enumeradas)	70
6.4.10	Habilitando um bloco de metadados	71
6.4.11	Atualizando o esquema do Solr	71
6.5	Recarregando um bloco de metadados	71
6.6	Exportação de Metadados	72

6.6.1	Exportações automáticas	72
6.6.2	Exportações em lote através da API	72
6.6.3	Falhas na Exportação	72
6.7	Aplicação de Temporizadores do Dataverse	73
6.7.1	Servidor de timer dedicado em um cluster de servidor do Dataverse . .	73
6.7.2	Temporizadores de coleta	74
6.7.3	Temporizador de Exportação de Metadados	74
6.8	Make Data Count	74
6.8.1	Introdução	74
6.8.2	Arquitetura	75
6.8.3	Limitações para instalações de dataverse usando handles em vez de DOIs	75
6.8.4	Habilitar o log para o Make data count	75
6.8.5	Configurar o Contador do Processador	75
6.8.6	Popular visualizações e downloads pela primeira vez	76
6.8.7	Popular visualizações e downloads noturnos	77
6.8.8	Enviando métricas de uso para o DataCite Hub	77
6.8.9	Configurando o Dataverse para fazer citações do Make Data Count . .	77
6.8.10	Recuperando métricas do Make Data Count do hub DataCite	77
6.8.11	Recuperando métricas do MAke Data Counts de dados do Dataverse .	78
6.9	Integrações	78
6.9.1	Obtendo dados	78
6.10	Administração do usuário	78
6.10.1	Gerenciar usuários	78
6.10.2	Listar usuários via API	79
6.10.3	Confirmar e-mail	79
6.11	Excluindo um token de API	80
6.12	Gerenciando conjuntos de dados e Dataverses	80
6.12.1	Dataverses	80
6.12.2	Conjuntos de dados	81
6.13	Índice de Pesquisa Solr	82
6.13.1	Reindexar completo	82
6.13.2	Reindexação manual	83
6.13.3	Consultando manualmente o Solr	84
6.14	Grupos de IP	84
6.14.1	Listando grupos de IP	84
6.14.2	Criando um grupo de IPs	84
6.14.3	Listando um grupo de IPs	85

6.14.4	Excluindo um grupo de IPs	85
6.15	Tráfego HTTP	85
6.15.1	Monitorando o tráfego HTTP do lado do cliente	86
6.15.2	Temporizadores EJB	86
6.16	Solução de problemas	86
6.16.1	Glass fish	86
6.16.2	Falha na implementação: Serviço de Timer EJB Não Disponível	87
6.16.3	Temporizador não funciona	88
7	Manual do usuário: Apêndice	89
7.1	Metadados	89
7.2	OpenAIRE	90
7.3	Planejando seu repositório de dados	99
7.4	Política de dados	99
7.5	Plano de Gestão de Dados (PGD)	100
7.6	Checklist EIFL	101
8	Referências	102

INSTALAÇÃO

1.1 Pré instalações

1.1.1 Preparação do Servidor

O sistema operacional utilizado no servidor será o Debian 9. Esse sistema operacional foi escolhido por ser um dos mais utilizados em Add comment Recompile 3 servidores de instituições acadêmicas. Os programas são instalados utilizando o programa gestor de pacotes Apt, contudo, por conta de problemas de compatibilidade, alguns programas serão instalados manualmente.

É necessário a autorização de administrador ou de superusuário(*root*) em algumas etapas da instalação, com isso, recomenda-se estar logado como *root* durante a instalação. Caso seja necessário a mudança de usuário, haverá uma instrução. Para acessar como superusuário, tendo conhecimento da senha, pode-se usar o comando **su**.

```
su -
```

Quando não é conhecido a senha de root, porém o usuário está no grupo de sudores do servidor, pode-se usar o comando a seguir.

```
sudo su
```

A primeira coisa a se fazer, depois de acessar como superusuário, é atualizar a base de dados do gestor de pacotes.

```
apt-get update
```

Em seguida, deve ser instalado alguns pacotes básicos do sistema operacional e programas básicos, que serão utilizados durante a instalação, como o editor de texto vim, o descompactador unzip, o programa para download web wget e o programa para download no github git,

```
apt-get install vim unzip wget git
```

A maioria das etapas são realizadas como superusuários, entretanto, é necessário atribuir a autorização de usuário comum aos pacotes do DataVerse, para evitar conflito na execução. Dessa forma, vamos criar um usuário de nome *dataverse*, para vinculá-lo aos pacotes do Dataverse.

```
useradd -m dataverse
```

Ao criar o usuário, também será criada uma pasta no diretório */home/*. Acesse a pasta */home/dataverse* e crie um diretório temporário. Essa pasta será usada para baixar os programas requisitados.

```
mkdir /home/dataverse/temp
```

1.1.2 Java

O pacote Java pode ser instalado utilizando o programa Apt, contudo, por questão de compatibilidade, recomendamos baixar o pacote diretamente da Oracle. Acesse a página de downloads do pacote java pelo endereço <https://www.oracle.com/java/technologies/jdk8-downloads.html>. Faça o download do pacote *jdk-8uxxx-linux-x64.tar.gz* e transfira o arquivo para a pasta */home/dataverse/temp*.

Em seguida, crie uma pasta de nome java no diretório */usr*.

```
mkdir /usr/java
```

Copie o arquivo *jdk-8uxxx-linux-x64.tar.gz* para a pasta */usr/java*.

```
cp /home/dataverse/temp/jdk-8uxxx-linux-x64.tar.gz /usr/java/
```

Agora, crie-se um link simbólico para os arquivos *java* e *javac* na pasta */usr/bin*

```
ln -s /usr/java/jdk1.8.0_221/bin/java /usr/bin/java
```

```
ln -s /usr/java/jdk1.8.0_221/bin/javac /usr/bin/javac
```

1.1.3 Glassfish

O programa Glassfish não precisa ser instalado, a aplicação é executada diretamente do diretório do pacote. Por questão de organização, vamos criar um usuário glassfish.

```
useradd glassfish
```

Baixe o programa glassfish 4.1 no diretório */home/dataverse/temp*,

```
cd /home/dataverse/temp/
```

```
wget http://download.oracle.com/glassfish/4.1/release/glassfish-4.1.zip
```

e descompacte-o.

```
unzip glassfish-4.1.zip
```

Em seguida, mova a pasta *glassfish4* para o diretório */usr/local*.

```
mv -R glassfish4 /usr/local/
```

Por conta da forma de execução do programa, deve-se modificar as permissões de algumas pastas e arquivos do pacote.

```
chown glassfish:glassfish /usr/local/glassfish4/glassfish/lib
```

```
chown -R glassfish:glassfish /usr/local/glassfish4/  
glassfish/domains/domain1
```

Dando continuidade, remova o arquivo *weld-osgi-bundle.jar*, localizado no diretório */usr/local/glassfish4/glassfish/modules/*

```
rm /usr/local/glassfish4/glassfish/modules/weld-osgi-bundle.jar
```

baixe o novo arquivo no diretório */home/dataverse/temp* e

```
wget http://central.maven.org/maven2/org/jboss/weld/weld-osgi-bundle/  
2.2.10.SP1/weld-osgi-bundle-2.2.10.SP1-glassfish4.jar
```

copie o novo arquivo .jar para o diretório */usr/local/glassfish4/glassfish/modules/*.

```
cp /home/dataverse/temp/weld-osgi-bundle-2.2.10.SP1-glassfish4.jar\  
/usr/local/glassfish4/glassfish/modules/
```

Depois de concluído todas as configurações básicas do glassfish, vamos inicia-lo,

```
/usr/local/glassfish4/bin/asadmin start-domain
```

e verificar a versão do Weld.

```
/usr/local/glassfish4/bin/asadmin osgi lb | grep 'Weld OSGi Bundle'
```

1.1.4 PostgreSQL(banco de dados)

O banco de dados PostgreSQL é instalado, utilizando o programa Apt,

```
apt-get install postgresql postgresql-contrib
```

Para o funcionamento do Dataverse, devem ser feitas algumas alterações na configuração do Postgres. Primeiro, acesse a pasta de arquivos principais do postgres

```
cd /etc/postgresql/(\textcolor{red}{versão_do_PSQL})/main/
```

e edite o arquivo *pg_hba.conf*,

```
vim /etc/postgresql/(versão do PSQL)/main/pg_hba.conf
```

Na sessão do script de configuração "*IPv4 local connection*", substitua o último termo do trecho(*md5*),

```
# IPv4 local connections:
host all all 127.0.0.1/32 md5
```

pelo termo *trust*.

```
# IPv4 local connections:
host all all 127.0.0.1/32 trust
```

Finalizado as modificações no Postgres, reinicie o banco de dados para subir as modificações.

```
/etc/init.d/postgres stop
/etc/init.d/postgres start
```

1.1.5 Solr

O programa Solr é executado diretamente da sua pasta de pacote, dessa forma, mantendo o padrão da instalação, baixe o pacote do programa Solr na pasta */home/dataverse/temp*,

```
cd /home/dataverse/temp
wget https://archive.apache.org/dist/lucene/solr/7.3.1/solr-7.3.1.tgz
```

e descompacte no mesmo diretório.

```
tar -xvzf solr-7.3.1.tgz
```

Crie um usuário solr, onde esse usuário será o proprietário dos arquivos do pacote.

```
useradd solr
```

Crie uma pasta solr no diretório */usr/local*, onde ficará os arquivos do pacote,

```
mkdir /usr/local/solr
```

em seguida, copie os arquivos baixados para a pasta `/usr/local/solr`.

```
cp -r /home/dataverse/temp/solr-7.3.1 /usr/local/solr/
```

O solr separa suas aplicações por *core*, assim, devemos preparar uma para o Dataverse. Copiamos os arquivos de configuração padrão para a pasta `collection1`.

```
cp -r /usr/local/solr/server/solr/configsets/_default  
/usr/local/solr/server/solr/collection1
```

Depois, copiamos os arquivos do Dataverse, que são para a configuração para o Solr, *schema.xml* e *solrconfig.xml*. Este arquivos estão na pasta de instalação do Dataverse, */home/dataverse/dvinstall*, e devem ser transferidos para a pasta de configuração da *collection1*, */usr/local/solr/solr-7.3.1/server/solr/collection1/conf*.

```
cp /home/dataverse/dvinstall/schema.xml  
/usr/local/solr/solr-7.3.1/server/solr/collection1/conf
```

```
cp /home/dataverse/dvinstall/solrconfig.xml  
/usr/local/solr/solr-7.3.1/server/solr/collection1/conf
```

Atribuímos como proprietário da pasta e dos arquivos o usuário `solr`.

```
chown -R solr:solr /usr/local/solr}
```

Iniciamos o Solr.

```
/usr/local/solr/solr-7.3.1/bin/solr start
```

1.1.6 Jq e ImageMagick

Primeiro, vamos adicionar o binário do Jq na lista de binários do sistema. Baixe o binário do Jq na pasta `temp` do Dataverse, */home/dataverse/temp/*.

```
cd /home/dataverse/temp/  
wget http://stedolan.github.io/jq/download/linux64/jq
```

Altere a permissão do arquivo para ser executável

```
chmod +x jq
```

e faça um link simbólico para o diretório `/usr/bin`

```
ln -s /home/dataverse/temp/jq /usr/bin/jq
```

Já o programa ImageMagick é instalado utilizando o programa Apt, pois não há nenhum registro de problema de compatibilidade.

```
apt install imagemagick
```

1.1.7 R

O programa R é instalado utilizando o pacote do fornecedor. Baixe o pacote na pasta */home/dataverse/temp*,

```
cd /home/dataverse/temp
wget https://cran.rstudio.com/src/base/R-3/R-3.6.1.tar.gz
```

descompacte o arquivo,

```
tar -zxvf R-3.6.1.tar.gz
```

e entre na pasta descompacta.

```
cd R-3.6.1
```

Para configurar a compilação e verificar se há alguma pendência, execute o comando,

```
./configure
```

depois, compile o código,

```
make
```

e instale a compilação.

```
make install
```

Finalizado esses passos, criamos um link simbólico do binário do R para o diretório */usr/bin*.

```
ln -s /home/dataverse/temp/R-3.6.1/bin/R /usr/bin/R
```

Para o funcionamento do Dataverse, é necessário a instalação de alguns pacotes do programa R. Acesse o programa R e instale os pacotes utilizando os comando a seguir.

```
install.packages("R2HTML", repos=https://cloud.r-project.org/)
install.packages("rjson", repos=https://cloud.r-project.org/)
install.packages("DescTool", repos=https://cloud.r-project.org/)
install.packages("Rserve", repos=https://cloud.r-project.org/)
install.packages("haven", repos=https://cloud.r-project.org/)
```

1.2 Instalação do Dataverse

A pré-instalação está concluída, agora, iniciamos a instalação do Dataverse. Vamos acessar o diretório de instalação do Dataverse,

```
cd /home/dataverse/dvinstall
```

certifique que está logado com usuário administrador ou superusuário e rode o script shell *setup.sh*.

```
./setup.sh
```

Ao executar o comando, vai aparece uma lista de perguntas para a determinar as configurações básicas para a instalação. Segue abaixo o formulário de perguntas, já com algumas respostas padrão, meramente.

```
Internet Address of your host: localhost
Glassfish Directory: /usr/local/glassfish4
Glassfish User: glassfish
Administrator e-mail address for this Dataverse:
SMTP (mail) server to relay notification messages: localhost
Postgres Server Address: [127.0.0.1]
Postgres Server Port: 5432
Postgres ADMIN password: secret
Name of the Postgres Database:
Name of the Postgres User:
Postgres user password: secret
Remote Solr indexing service: LOCAL
Rserve Server: localhost
Rserve Server Port: 6311
Rserve User Name: rserve
Rserve User Password: rserve
```

Depois de preenchido os campos, o script dará prosseguimento, finalizando a instalação. Durante todo o processo, é possível visualizar os *logs* da instalação. Caso aconteça algum erro, a instalação será interrompida.

Instalação concluída, o acesso ao seu Dataverse deve ser feito utilizando o endereço de acesso do seu *host*. O primeiro login é feito usando o usuário *dataverseAdmin* e senha *admin*, um resumo segue abaixo.

```
URL: http://localhost:8080
username: dataverseAdmin
password: admin
```

1.3 Configurações

1.3.1 Provedor Doi

O Dataverse vem com a opção de vinculação do dataset a um número de registro pertinente, DOI. Existe a opção de utilizar o DataCite e o EZID. A baixo, segue o comando para ativação do provedor do DOI, para exemplificar, o comando segue com o provedor DataCite seja escolhido o Datacite

```
curl -X PUT -d DataCite http://localhost:8080/api/admin/settings/:DoiProvider
```

Caso a instituição, de forma provisória ou nem queira contratar nenhum provedor, mas existe a necessidade de publicação dos dados, como por exemplo, para a coleta das informações dos dados. O Dataverse possibilita a ativação, no qual é utilizado um provedor DOI falso. Para isso, no comando, é necessário substituir o nome do provedor por "FAKE", da forma que é mostrado abaixo.

```
curl -X PUT -d FAKE http://localhost:8080/api/admin/settings/:DoiProvider
```


Manual do usuário: Criação de contas + Gerenciamento

Este manual é uma tradução, adaptação e sintetização das informações encontradas no guia disponível em: <http://guides.dataverse.org>.

2.1 Informações da conta

Como usuário registrado, você pode:

- Criar seu próprio Dataverse, e se permitido, personalize-o.
- Adicionar conjuntos de dados aos dados, se permitido.
- Contribuir com os conjuntos de dados existentes, se permitido.
- Solicitar acesso a arquivos restritos, se permitido.

2.1.1 Opções de log-in

O Dataverse foi configurado para uma ou mais das seguintes opções de login:

- Nome de usuário / e-mail e senha
- Login Institucional
- ORCID
- GitHub
- Google

Observe que depois de criar sua conta do Dataverse, ela será associada apenas a uma das opções de login acima.

2.1.2 Criação de contas

Para criar uma conta do Dataverse com a opção de login Nome de usuário / e-mail, use a página "Sign Up(Inscreva-se)". Preencha os campos e clique no botão "Create Account (Criar conta)". Observe que o campo sername (Nome de usuário) não suporta endereços de e-mail, mas permitirá os seguintes caracteres: a-Z, 0-9, _ (sublinhados), - (hífens) e. (períodos).

Para criar uma conta do Dataverse associada à opção de login da sua instituição, ORCID, GitHub ou Google, use a página "Login" e selecione um dos provedores de autenticação.

2.1.3 Edição de contas

1. Para editar sua conta após o login, clique no nome da conta no cabeçalho à direita e clique em Account Information (Informações da conta).
2. No canto superior direito da página da sua conta, clique no botão "Edit Account (Editar conta)" e, a partir daí, você pode optar por editar as informações ou a senha da conta.
3. Selecione "Save Changes (Salvar alterações)" quando terminar.

Observe que você não pode editar as informações da sua conta no Dataverse se usar a opção Login institucional. Em vez disso, você deve entrar em contato com sua instituição para alterar seu nome, e-mail, etc. Depois que a alteração for feita por sua instituição, ela será refletida no Dataverse na próxima vez que você fizer login. Os usuários da opção Log In Institutional não precisam verificar seu endereço de e-mail porque a instituição que fornece o endereço de e-mail é confiável.

2.1.4 Conversão de contas

Se mais de uma opção de login estiver disponível, você poderá converter de uma para outra e usar a nova opção a partir de agora.

Se você estiver convertendo a partir da opção de login por e-mail / nome de usuário, precisará ter sua senha pronta para concluir o processo de conversão. Clique em Logon, selecione a nova opção de logon e siga o processo de logon. Ao retornar ao Dataverse, procure uma opção que permita converter e digitar sua senha para concluir a conversão.

2.1.5 Resetar senha

Somente contas do Dataverse que usam a opção de login Nome de usuário / e-mail têm uma senha associada armazenada (com segurança) no Dataverse. Se você não conseguir se lembrar dessa senha, clique em Log In no canto superior direito de

qualquer página e clique no link "Forgot Password? (Esqueceu a senha?)" Abaixo, onde você digitaria seu nome de usuário / e-mail e senha. Digite seu endereço de e-mail e clique em "Submit Password Request (Enviar solicitação de senha)" para receber um e-mail com um link para redefinir sua senha.

Observe que, se você esqueceu seu nome de usuário, pode usar esse mesmo processo para receber seu nome de usuário em um e-mail.

2.2 Meus dados

A seção Meus dados da página da sua conta exibe uma lista de todos os dados, conjuntos de dados e arquivos que você criou, enviou ou que você tem acesso para editar. Você pode filtrar todos os dados, conjuntos de dados e arquivos listados lá usando a caixa de filtro. Você também pode usar as facetas no lado esquerdo para exibir apenas um Status ou função de publicação específico.

2.3 Notificações

As notificações aparecem na guia de notificações da página da sua conta e também são exibidas como um número ao lado do nome da sua conta.

Você receberá uma notificação quando:

- Você criar sua conta;
- Você criar um dataverse ou adicionar um conjunto de dados;
- Outro usuário do Dataverse solicitar acesso a um arquivo restrito em um de seus conjuntos de dados.

As notificações serão enviadas por e-mail apenas uma vez, mesmo que você não tenha lido a notificação no site do Dataverse.

Manual do usuário: Encontrando e usando dados

3.1 Encontrado dados

Sem fazer login no Dataverse, os usuários podem navegar, procurar dados, conjuntos de dados e arquivos, visualizar descrições e arquivos de conjuntos de dados publicados e subconjunto, analisar e visualizar dados de arquivos de dados publicados (restritos e não restritos). Para visualizar um dataverse, conjunto de dados ou arquivo não publicado, será necessário que o usuário tenha permissão do administrador do administrador de dados para acessá-lo.

Um usuário pode pesquisar dentro de um Dataverse específico os conjuntos de dados e arquivos que ele contém, usando a barra de pesquisa e as facetas exibidas na página desse Dataverse.

3.1.1 Busca básica

Você pode pesquisar todo o conteúdo da instalação do Dataverse, incluindo os dataverses, conjuntos de dados e arquivos. Você pode acessar a pesquisa através da barra de pesquisa na página inicial ou clicando no ícone de lupa no cabeçalho de todas as páginas. A barra de pesquisa aceita termos, consultas ou frases exatas (entre aspas).

Organizando e visualizando resultados da busca

Facetas: à esquerda dos resultados da pesquisa, existem várias facetas nas quais um usuário pode clicar para restringir o número de resultados exibidos.

- Escolhendo uma faceta: para escolher uma faceta para restringir seus resultados, clique nessa faceta.
- Remoção de uma faceta: uma faceta escolhida pode ser removida clicando no "X" nela, no painel de facetas à esquerda ou acima dos resultados.
- Visualizando mais ou menos facetas: cada categoria no painel de facetas lista as 5 principais facetas mais comuns dessa categoria. Para ver mais, clique em "More...(Mais...)" no canto inferior direito dessa categoria. Depois de optar por ver mais, uma opção para ver menos será exibida na parte inferior esquerda da faceta.

Cartões de resultados: depois de inserir um termo ou consulta de pesquisa, os cartões de resultados que correspondem ao seu termo ou consulta aparecem embaixo da barra de pesquisa e à direita das facetas.

- Relevância dos resultados: cada cartão de resultados mostra quais campos de metadados correspondem à consulta de pesquisa ou termo digitado na barra de pesquisa, com o termo ou consulta correspondente em negrito. Se o termo ou consulta de pesquisa foi encontrado no título ou nome do Dataverse, conjunto de dados ou arquivo, o termo ou consulta de pesquisa será destacado em negrito.

Outros recursos básicos de pesquisa:

- Classificação dos resultados: os resultados da pesquisa podem ser classificados por nome (A-Z ou Z-A), por data (mais recente ou mais antiga) ou por relevância dos resultados. O botão de classificação pode ser encontrado acima dos resultados da pesquisa, no canto superior direito.
- URLs que podem ser marcados como favoritos: os URLs de pesquisa podem ser copiados e enviados para um colega pesquisador ou podem ser marcados para sessões futuras.

3.1.2 Busca avançada

Para realizar uma pesquisa avançada, clique no link "Advanced Search (Pesquisa avançada)" ao lado da barra de pesquisa. Lá, você poderá inserir termos de pesquisa para dados, metadados do conjunto de dados (citação e domínio específico) e metadados no nível do arquivo. Se você estiver procurando por arquivos de dados tabulares, também poderá procurar no nível da variável nome e rótulo. Para descobrir mais sobre o que cada campo pesquisa, passe o mouse sobre o nome do campo e aparecerá uma descrição detalhada dele.

3.1.3 Navegando no Dataverse

No Dataverse, a navegação é a visualização padrão quando um usuário não inicia uma pesquisa na página inicial ou na página de um dataverse específico. Ao navegar, apenas os conjuntos de dados e dados aparecem na lista de resultados e os resultados podem ser classificados por Nome (A-Z ou Z-A) e por Mais novo ou Mais antigo.

3.2 Usando Dados

3.2.1 Ver Dataverses + Datasets

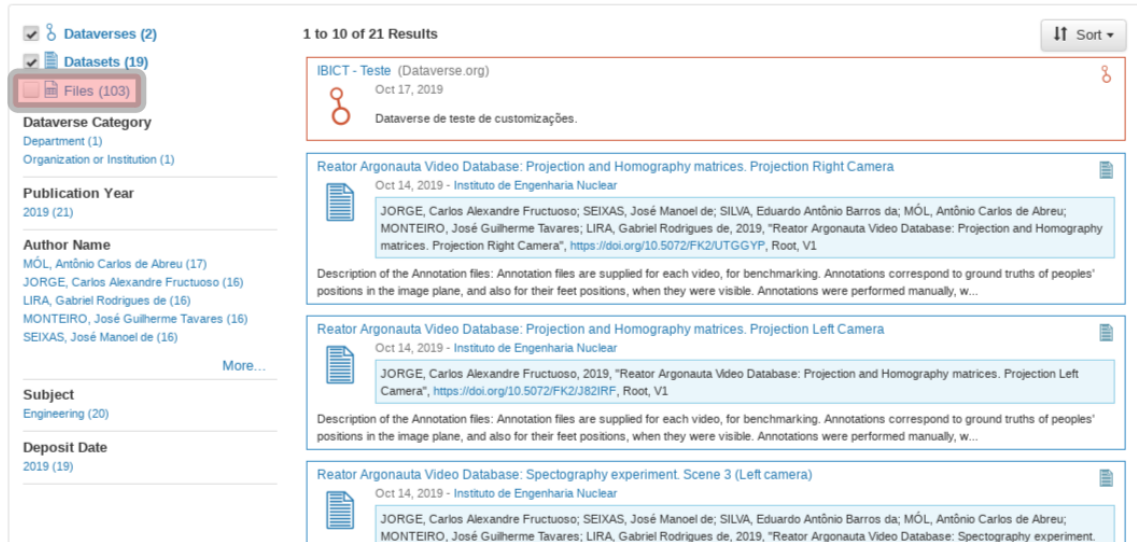
Após realizar uma pesquisa e localizar o dataverse ou conjunto de dados que você está procurando, clique no nome do dataverse ou conjunto de dados ou no thumbnail para ser levado para a página desse dataverse ou conjunto de dados. Uma vez em uma página do dataverse, é possível visualizar os dataverses, conjuntos de dados e arquivos nesse dataverse.

Quando estiver na página do conjunto de dados, você verá o título, a citação, a descrição e vários outros campos, além de um botão para enviar um e-mail ao contato do conjunto de dados e um botão para compartilhar o conjunto de dados nas mídias sociais. Abaixo dessas informações, estão disponíveis os arquivos, metadados, termos de uso e informações de versão do conjunto de dados.

3.2.2 Ver arquivos

Os arquivos no Dataverse têm sua própria página que pode ser acessada pelos resultados da pesquisa ou pela tabela Arquivos na página do conjunto de dados pai. A página do conjunto de dados e a página do arquivo oferecem a mesma funcionalidade em termos de exibição e edição de arquivos, com algumas exceções. A página do arquivo inclui o identificador persistente do arquivo (DOI ou identificador), que pode ser encontrado na guia Metadados [1]. Além disso, a guia Versões [2] da página do arquivo fornece um histórico de versões mais focado no arquivo individual e não no conjunto de dados como um todo.

Figura 1: Página principal

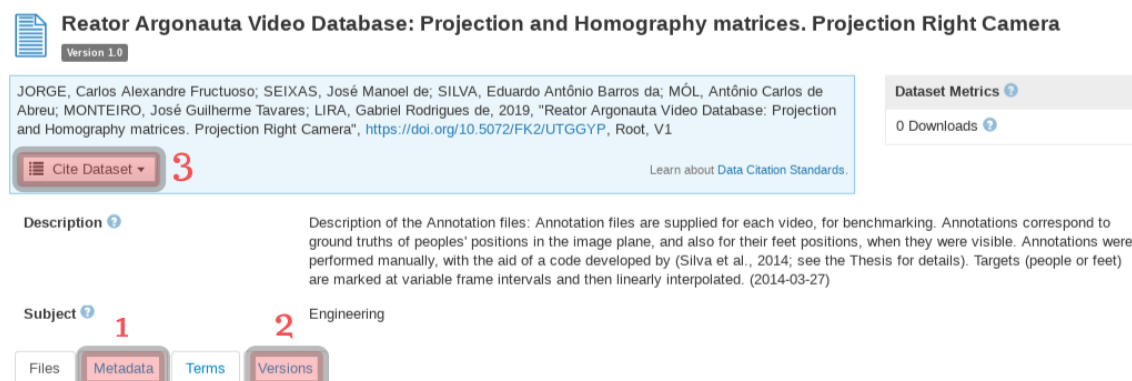


Fonte: Elaboração própria

3.2.3 Citar Dados, versões e metadados

Você pode encontrar a citação para o conjunto de dados na parte superior da página do conjunto de dados em uma caixa azul. Além disso, há um botão Citar Dados [3] que oferece a opção de baixar a citação como EndNote XML, RIS Format ou BibTeX Format.

Figura 2: Citar dados



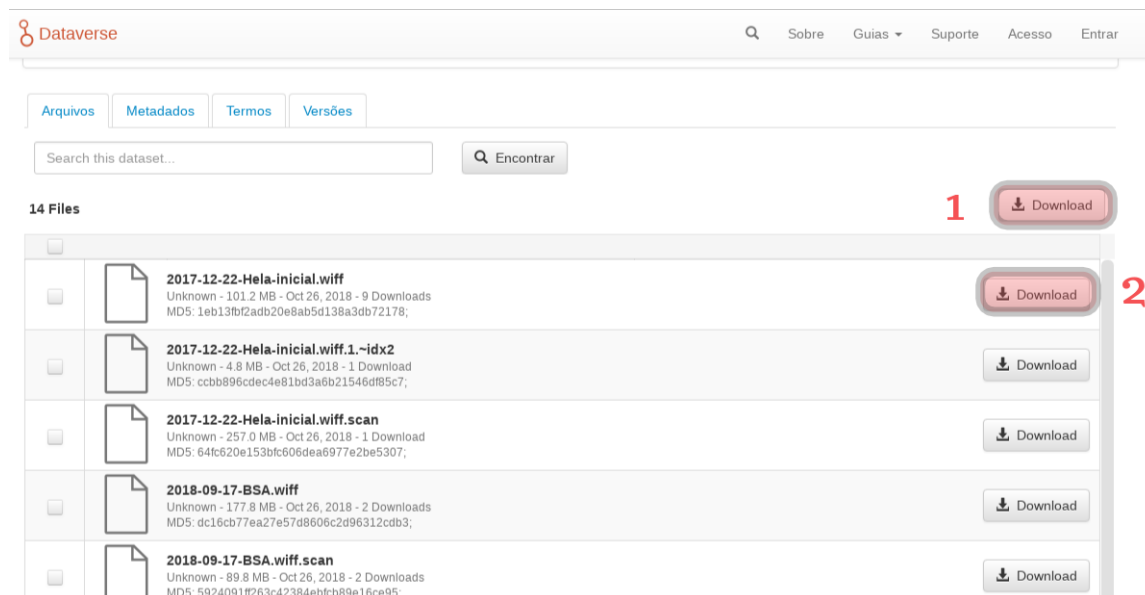
Fonte: Elaboração própria

3.2.4 Download de arquivos

Na guia Arquivos em uma página de conjunto de dados, você pode fazer o download dos arquivos nesse conjunto de dados. Para baixar mais de um arquivo por vez, selecione os arquivos que você deseja baixar e clique no botão Download [1] acima dos arquivos. Os arquivos selecionados serão baixados no formato .zip que preserva qualquer estrutura de pastas que o proprietário do conjunto de dados configurou.

Um arquivo da página do dataset, pode ser baixado, clicando no botão Download [2] no canto superior direito da página ou Fazendo o download via URL na guia Metadados na metade inferior da página.

Figura 3: Arquivos



Fonte: Elaboração própria

a Dados Tabulares

Os arquivos ingeridos podem ser baixados de várias maneiras diferentes:

- A opção padrão é baixar um arquivo de valores separados por tabulação, que é um padrão fácil e gratuito de usar;
- O arquivo original, que pode estar em um formato proprietário, requer software especial;
- Formato Rdata se a instalação configurou este;
- Os metadados da variável para o arquivo no formato DDI;
- Um subconjunto das colunas dos dados.

b Downloading via URL

O Dataverse exibe um URL de texto sem formatação para o local do arquivo na guia Metadados na página do conjunto de dados. Este URL de download pode ser usado para acessar diretamente o arquivo via API (ou em um navegador da Web, se necessário). Ao baixar arquivos maiores, para garantir um download confiável e retomável, recomendamos o uso do GNU Wget em um terminal de linha de comando ou o software gerenciador de download de sua escolha.

Certos arquivos não fornecem URLs de download por motivos técnicos: aqueles que são restritos, têm termos de uso associados ou fazem parte de um dataverse com um livro de visitas ativado.

c Downloading um pacote Dataverse via URL

Os pacotes de dataverse são normalmente usados para representar arquivos ou pacotes configuráveis extremamente grandes que contêm um grande número de arquivos. Os pacotes de dataverse geralmente são grandes demais para serem baixados com segurança usando um navegador da web. Quando você clica para baixar um pacote do Dataverse, em vez de iniciar automaticamente o download no navegador da Web, o Dataverse exibe um URL em texto sem formatação para o local do arquivo. Para garantir um download confiável e retomado, recomendamos o uso do GNU Wget em um terminal de linha de comando ou o software gerenciador de download de sua escolha. Se você tentar simplesmente colar o URL no seu navegador da Web, o download poderá sobrecarregá-lo, resultando em um download interrompido, excedido o tempo limite ou com falha.

3.2.5 Explore Dados

Observe que a instalação do Dataverse que você está usando pode ter ferramentas adicionais ou diferentes configuradas como o TwoRavens ou o WorldMap.

Manual do usuário: Gerenciamento Dataverse

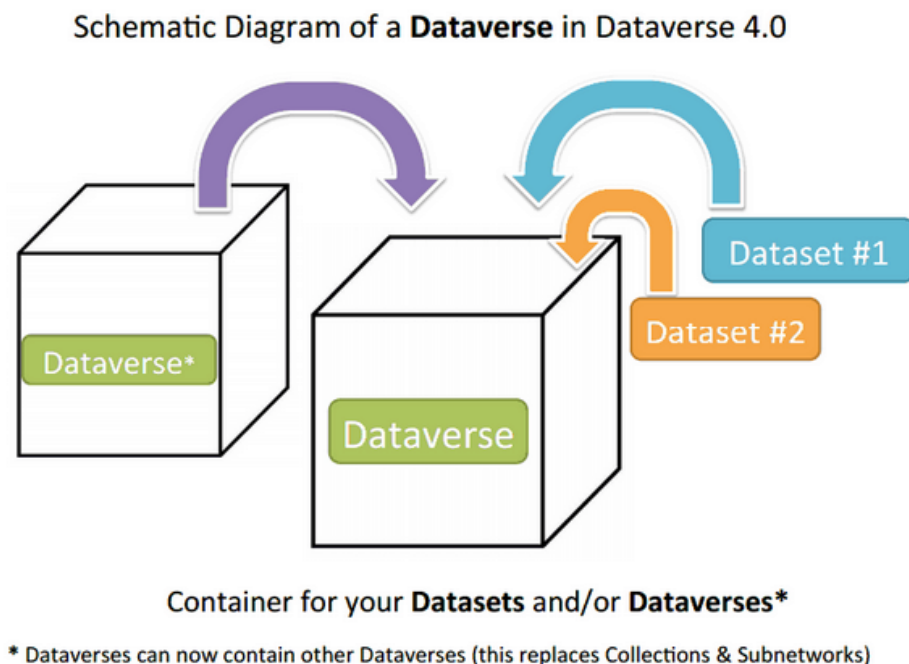
Um Dataverse é um contêiner para conjuntos de dados, ou seja:

- dados de pesquisa;
- código;
- documentação;
- metadados;
- outros dataverses;
- entre outros.

Que podem ser configurados para pesquisadores, departamentos, periódicos e organizações individuais.

Depois que um usuário cria um dataverse, por padrão, ele se torna o administrador desse dataverse. O administrador do dataverse tem acesso para gerenciar as configurações.

Figura 4: Diagrama esquemático



Fonte: DATAVERSE PROJECT, 2019

4.1 Criação de um Dataverse (Dentro do Dataverse “Root”)

1. Para criar um dataverse primeiro você deve ser um usuário registrado. Uma vez logado, clique no botão "Add Data (Adicionar Dados)" e, no menu suspenso, selecione "New Dataverse (Novo Dataverse)".

2. Na página "New Dataverse", preencha os seguintes campos:

- Nome: digite o nome do seu dataverse.
- Identificador: é uma abreviação, geralmente em letras minúsculas, que se torna parte da URL do novo dataverse. Caracteres especiais (; ', !, @, #, \$, %, ; & E *) e espaços não são permitidos. Se você alterar o campo URL do Dataverse, o URL do seu Dataverse será alterado (http //.../ 'url'), e afetará os links para esta página.
- E-mail: este é o endereço de e-mail que será usado como contato para esse dataverse específico. Você pode ter mais de um endereço de e-mail de contato para o seu dataverse.
- Afiliação: adicione qualquer Afiliação que possa ser associada a esse dataverse específico (por exemplo, nome do projeto, nome do instituto, nome do departamento, nome do diário, etc.). Isso é preenchido automaticamente se você adicionou uma afiliação à sua conta de usuário.
- Descrição: forneça uma descrição desse dataverse. Ela será exibida na página inicial do seu dataverse e na lista de resultados da pesquisa. O campo de descrição suporta determinadas tags HTML.

- **Categoria:** selecione uma categoria que melhor descreva o tipo de dataverse. Por exemplo, se esse é um dataverse para os conjuntos de dados de um pesquisador individual, selecione "Pesquisador". Se for um dataverse para uma instituição, selecione "Organização e instituição".
- Escolha os conjuntos de Elementos de metadados para conjuntos de dados neste dataverse: Por padrão, os elementos de metadados serão do dataverse do host em que esse novo dataverse é criado. O Dataverse oferece padrões de metadados para vários domínios.
- Selecionar facetas para esse dataverse: por padrão, as facetas que aparecerão na sua página de destino do dataverse serão do dataverse host no qual esse novo dataverse foi criado. As facetas são simplesmente campos de metadados que podem ser usados para ajudar outras pessoas a encontrar facilmente dataverses e conjuntos de dados dentro desse dataverse. Você pode selecionar quantas facetas desejar.

Figura 5: Facetas

The screenshot shows the 'New Dataverse' form. At the top, there's a breadcrumb 'Root > New Dataverse'. The form has several input fields: 'Dataverse' (with a red asterisk), 'Identifier' (with a red asterisk), 'Category' (a dropdown menu), 'Email' (with a red asterisk), 'Affiliation', 'Host Dataverse', and 'Description'. Below these are two sections: 'Metadata Fields' and 'Browse/Search Facets'. 'Metadata Fields' has a checkbox 'Use metadata fields from Root' (checked) and a list of metadata standards with checkboxes and 'View fields' links. 'Browse/Search Facets' has a checkbox 'Use browse/search facets from Root' (checked) and a dropdown menu for selecting facets.

Fonte: Elaboração própria

3. Os elementos de metadados selecionados também são usados para escolher quais campos de metadados você gostaria de usar para criar modelos para seus conjuntos de dados. Os campos de metadados podem ser ocultos ou selecionados conforme necessário ou opcional. Depois de selecionar todos os campos que gostaria de usar, você poderá criar seu (s) modelo (s) depois de terminar de criar seu dataverse.

4. Clique no botão "Create Dataverse (Criar Dataverse)"

Os campos obrigatórios são indicados por um asterisco vermelho.

4.2 Edite um Dataverse

Para editar seu dataverse, navegue até a página inicial do dataverse e selecione o botão "Edit Dataverse (Editar dataverse)", onde serão apresentadas as seguintes opções de edição:

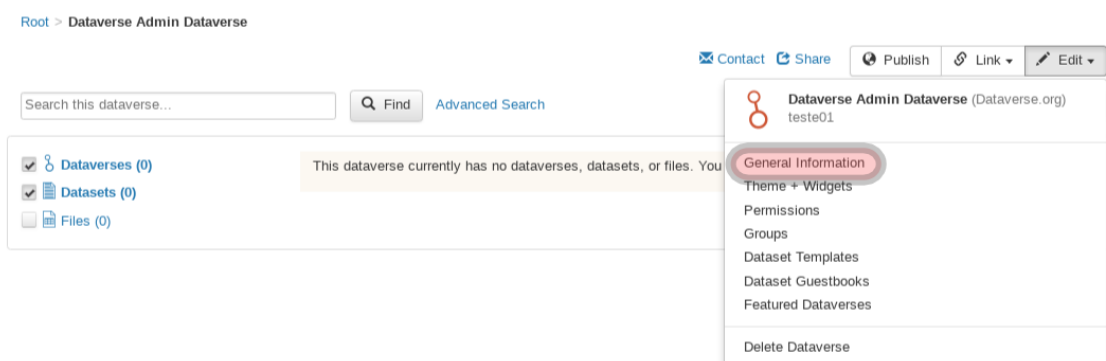
- Informações gerais: editar nome, identificador, categoria, e-mail de contato, afiliação, descrição, elementos de metadados e facetas para o seu dataverse;
- Tema: faça upload de um logotipo para o seu dataverse, adicione um link ao seu departamento ou site pessoal e selecione cores para o dataverse para marcá-lo;
- Widgets: obtenha um código para adicionar ao seu site e exibir seu dataverse nele;
- Permissões: conceda aos usuários do Dataverse permissões para o seu dataverse, ou seja, pode editar conjuntos de dados e ver quais usuários já possuem quais permissões para quais dataverse;
- Modelos de conjunto de dados: são úteis quando você tem vários conjuntos de dados com as mesmas informações em vários campos de metadados que você não precisará digitar manualmente toda vez;
- Livros de visitas do conjunto de dados: permite coletar dados sobre quem está baixando os arquivos dos seus conjuntos de dados;
- Dataverses em destaque: se você tiver um ou mais dataverses, poderá usar esta opção para mostrá-los na parte superior da página do dataverse para ajudar outras pessoas a encontrar facilmente dataverses interessantes ou importantes;
- Excluir o Dataverse: você pode excluir o seu dataverse desde que não seja publicado e não possua nenhum rascunho de conjuntos de dados.

4.3 Informação geral

A página "Informações gerais" é como você edita as informações preenchidas enquanto cria seu dataverse. Se você precisar alterar ou adicionar um endereço de e-mail de contato, este é o local para fazê-lo. Além disso, você pode atualizar os elementos de metadados usados para conjuntos de dados no dataverse, alterar quais campos de metadados estão ocultos, necessários ou opcionais e atualizar as facetas que você gostaria que fossem exibidas para navegar no dataverse. Se você planeja usar modelos, será necessário selecionar os campos de metadados na página Informações gerais.

Os campos de metadados que você selecionar conforme necessário aparecerão no formulário Criar conjunto de dados quando alguém adicionar um conjunto de dados ao dataverse.

Figura 6: Campos



Fonte: Elaboração própria

4.4 Tema

O recurso Tema fornece uma maneira de personalizar a aparência do seu dataverse. Você pode optar por usar o tema do dataverse que contém o dataverse (até o dataverse raiz, também conhecido como a página inicial) ou fazer upload de seu próprio arquivo de imagem. Os tipos de imagem suportados são JPEG, TIFF ou PNG e não devem ter mais que 500 KB. O tamanho máximo de exibição de um arquivo de imagem no tema de um dataverse é de 940 pixels de largura por 120 pixels de altura. As cores para o cabeçalho do seu dataverse e o texto que aparece no dataverse também podem ser selecionadas. Um link também pode ser adicionado ao seu site pessoal, o site da sua organização ou instituição, seu departamento, diário, etc.

4.5 Widgets

O recurso Widgets fornece código para você colocar em seu site pessoal e exibir seu dataverse lá. Existem dois tipos de Widgets para um dataverse, um widget Caixa de Pesquisa do Dataverse e um widget da Listagem do Dataverse. Na guia Widgets da página Tema + Widgets, você pode copiar e colar os trechos de código do widget que você deseja adicionar ao seu site. Se você precisar ajustar a altura do widget em seu site, poderá fazê-lo editando o parâmetro `heightPx = 500` no snippet de código.

4.5.1 Widget Caixa de Pesquisa

O Widget da Caixa de Pesquisa do Dataverse adicionará uma caixa de pesquisa ao seu site vinculada ao seu dataverse. Os usuários são direcionados ao seu dataverse em uma nova janela do navegador para exibir os resultados dos termos de pesquisa inseridos no campo de entrada.

4.5.2 Widget Lista

O widget de listagem de dados dataverse fornece uma lista de todos os seus dataverses e conjuntos de dados para os usuários procurarem, classificarem, filtrarem e pesquisarem. Quando alguém clica em um dataverse ou conjunto de dados no widget, ele exibe o conteúdo no widget no seu site. Eles podem baixar arquivos de dados diretamente dos conjuntos de dados no widget. Se um arquivo for restrito, eles serão direcionados ao seu dataverse para efetuar login, em vez de fazer o login através do widget em seu site.

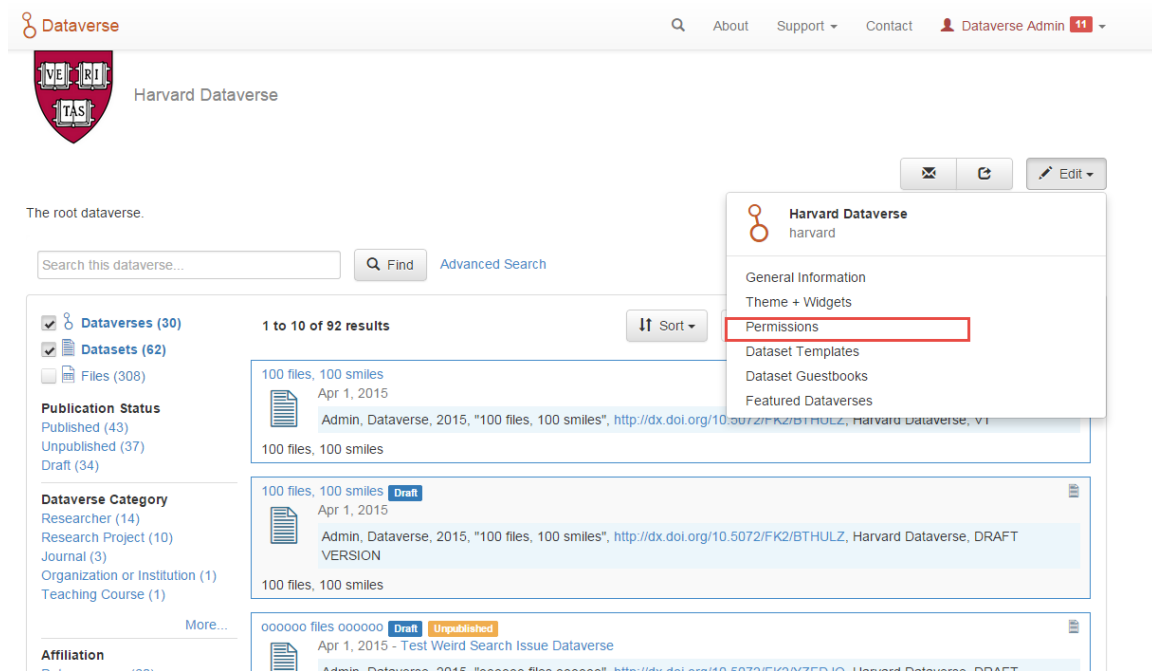
4.6 Papéis e Permissões

As contas de usuário do dataverse podem receber funções que definem quais ações têm permissão para executar em dataverses, conjuntos de dados e / ou arquivos específicos. Cada função vem com um conjunto de permissões, que definem as ações específicas que os usuários podem executar.

Funções e permissões também podem ser concedidas a grupos. Grupos podem ser definidos como uma coleção de contas de usuário do Dataverse, uma coleção de endereços IP (por exemplo, todos os usuários dos computadores de uma biblioteca) ou uma coleção de todos os usuários que efetuam login usando um login institucional específico (por exemplo, todos que efetuam login com um determinado credenciais de conta da universidade).

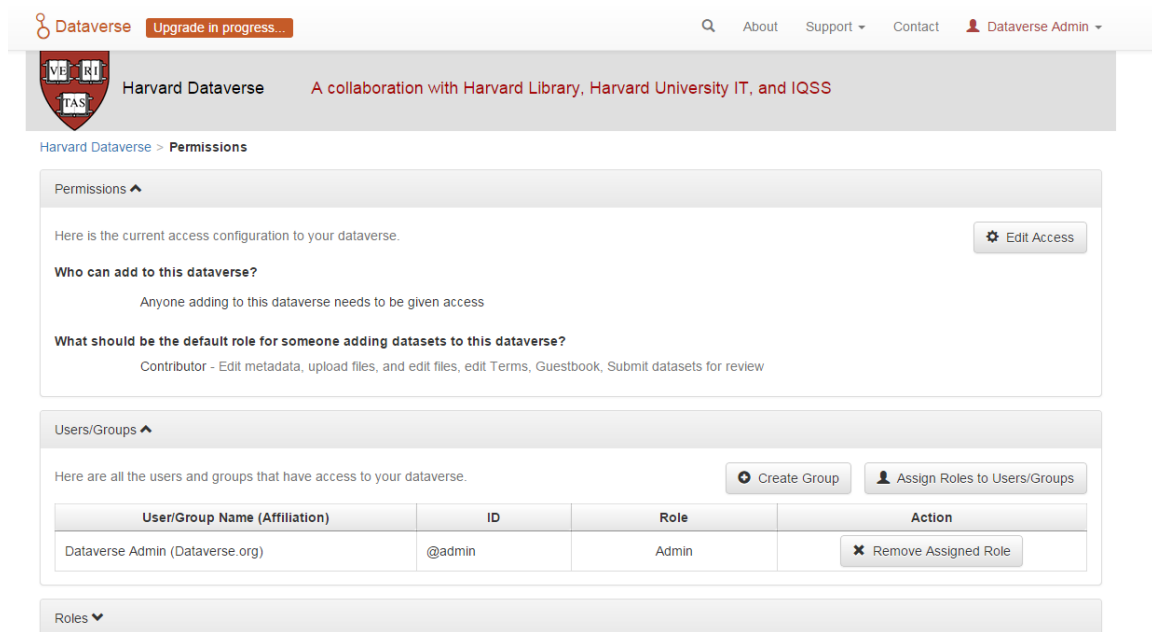
Os administradores de um dataverse podem atribuir funções e permissões aos usuários desse dataverse. Se você é um administrador em um dataverse, encontrará o link para a página "Permissões", no menu suspenso Editar, na página do dataverse.

Figura 7: Permissões



Fonte: DATAVERSE PROJECT, 2019

Figura 8: Página permissões



Fonte: DATAVERSE PROJECT, 2019

Ao acessar a página de permissões de um dataverse, você visualizará três seções:

Permissões: Define os requisitos que determinam quais tipos de usuários podem adicionar conjuntos de dados e subavaliadores de dados ao seu dataverse e quais per-

missões serão concedidas quando o fizerem.

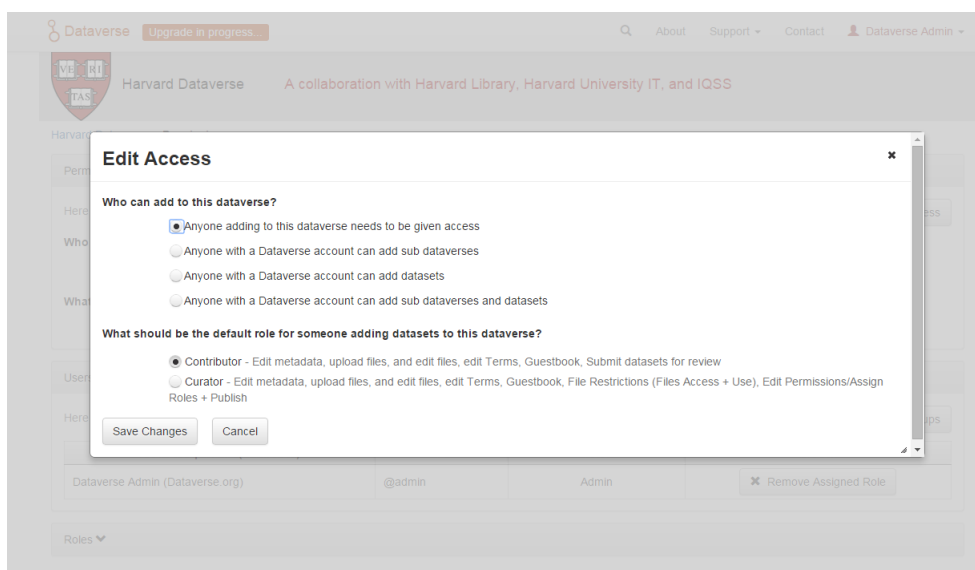
Usuários / Grupos: Atribui funções a usuários ou grupos específicos, determinando quais ações eles têm permissão para executar no seu dataverse. Poderá também pode fazer referência a uma lista de todos os usuários que têm funções atribuídas a eles para o seu dataverse e remover as funções, se desejar.

Funções: Elabora uma lista completa de funções que podem ser atribuídas aos usuários do seu dataverse. Cada função lista as permissões que oferece. Observe que, mesmo em um dataverse recém-criado, você poderá ver que usuários e grupos já receberam funções, se sua instalação tiver: Conjunto de InheritParentRoleAssignments.

4.6.1 Definindo Configurações de Acesso

Na guia Permissões, você pode clicar no botão "Edit Access (Editar acesso)" para abrir uma caixa onde você pode definir quem irá adicionar itens ao seu dataverse e quais permissões são concedidas àqueles que os adicionam.

Figura 9: Tipos de permissão



Fonte: DATAVERSE PROJECT, 2019

A primeira pergunta nesta página permite determinar como é aberto o seu dataverse para novas adições - você pode definir se a base de usuários inteira (todos os usuários logados) tem ou não a capacidade de adicionar conjuntos de dados ou sub-dataverses ao seu dataverse.

A segunda pergunta nesta página permite escolher a função (e, portanto, as permissões) concedidas aos usuários que adicionam um conjunto de dados ao seu dataverse. A função que você selecionar será concedida automaticamente a qualquer usuário que criar um conjunto de dados no seu dataverse, nesse conjunto de dados, no momento

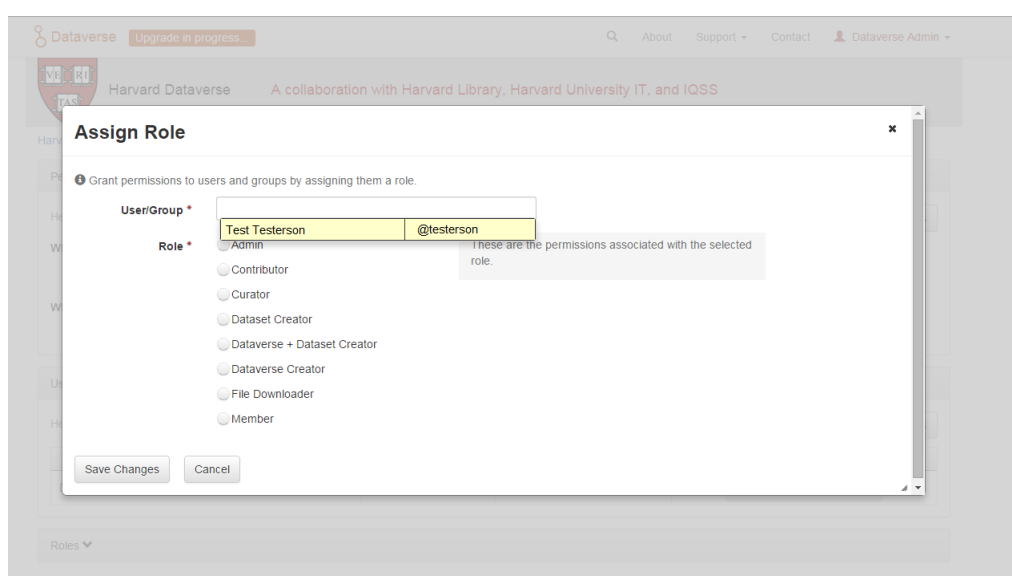
em que o criar. A função que o usuário recebe determina suas permissões para o conjunto de dados que ele criou. A principal diferença entre as duas funções é que os curadores podem publicar seus próprios conjuntos de dados, enquanto os colaboradores devem enviar o conjunto de dados para revisão antes da publicação. Além disso, os curadores podem gerenciar as permissões do conjunto de dados. Observe que essa configuração não aplica funções retroativamente aos usuários que adicionaram conjuntos de dados anteriormente ao seu dataverse; aplica-se apenas aos usuários que adicionam novos conjuntos de dados daqui para frente. Ambas as configurações podem ser alteradas a qualquer momento.

4.6.2 Definindo papéis para Usuários e Grupos

Na guia Usuários / Grupos, você pode adicionar, editar ou remover as funções concedidas a usuários e grupos no seu dataverse. Uma função é um conjunto de permissões concedidas a um usuário ou grupo que estiver usando seu dataverse. Por exemplo, atribuir ao seu assistente de pesquisa a função "Colaborador" lhe daria as seguintes permissões autoexplicativas no seu dataverse e em todos os conjuntos de dados dentro do dataverse: "ViewUnpublishedDataset", "DownloadFile", "EditDataset" e "DeleteDatasetDraft".

Essa função, no entanto, não teria a permissão "PublishDataset" e, portanto, seria incapaz de publicar conjuntos de dados no seu dataverse. Se você quisesse dar a ela essa permissão, você daria a ela um papel com essa permissão, como por exemplo o papel de Curador. Usuários e grupos podem ter várias funções ao mesmo tempo, se necessário. As funções podem ser removidas a qualquer momento. Todas as funções e suas permissões associadas são listadas na guia "Funções" da mesma página.

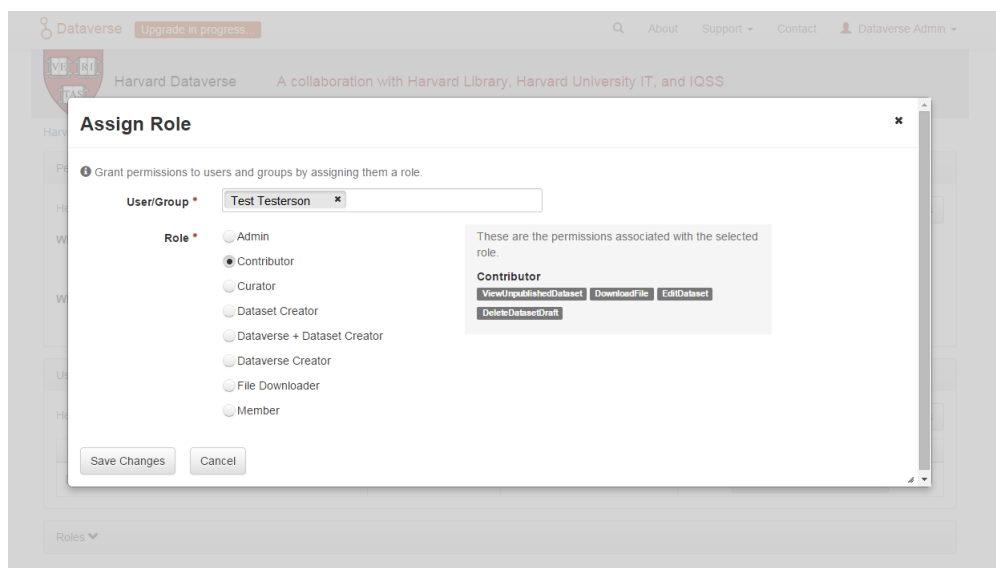
Figura 10: Papéis



Fonte: DATAVERSE PROJECT, 2019

Observe que a função Criador do conjunto de dados e a função Colaborador às vezes são confusas. A função Criador do conjunto de dados é atribuída no nível do dataverse e permite que um usuário crie novos conjuntos de dados nesse dataverse. A função Colaborador pode ser atribuída no nível do conjunto de dados, concedendo ao usuário a capacidade de editar esse conjunto de dados específico. Como alternativa, a função Colaborador pode ser atribuída no nível do dataverse, concedendo ao usuário a capacidade de editar todos os conjuntos de dados nesse dataverse.

Figura 11: Designando papéis



Fonte: DATAVERSE PROJECT, 2019

Se você precisar atribuir uma função a TODAS as contas de usuário do Dataverse, poderá atribuir a função ao grupo: "usuários autenticados".

4.7 Templates de Dataset

Os modelos são úteis quando você tem vários conjuntos de dados com as mesmas informações em vários campos de metadados que você prefere não continuar digitando manualmente ou se deseja usar um conjunto personalizado de "Termos de Uso e Acesso" para vários conjuntos de dados em um dataverse.

No Dataverse, os modelos são criados no nível do dataverse, podem ser excluídos (para que não sejam exibidos para futuros conjuntos de dados), definidos como padrão (não obrigatório) ou podem ser copiados para que você não precise recomendar ao criar um novo modelo com metadados semelhantes de outro modelo. Quando um modelo é excluído, ele não afeta os conjuntos de dados que já o usaram, ou seja, os conjuntos de dados que já utilizam um modelo não terão seu modelo trocado ou excluído.

Figura 12: Termos de uso

Terms of Use

Waiver

Datasets will default to a [CC0 public domain dedication](#). CC0 facilitates reuse and extensibility of research data. Our [Community Norms](#) as well as good scientific practices expect that proper credit is given via citation. If you are unable to give datasets a CC0 waiver you may enter custom Terms of Use for datasets.

☐ Yes, apply CC0 - "Public Domain Dedication" ☒ No, do not apply CC0 - "Public Domain Dedication"

If you are unable to use CC0 for datasets you are able to set custom terms of use. Here is an example of a [Data Usage Agreement](#) for datasets that have de-identified human subject data.

Terms of Use

Confidentiality Declaration

Special Permissions

Restrictions

Fonte: Elaboração própria

Para criar um modelo:

1. Navegue para o seu dataverse, clique no botão Editar dataverse e selecione Modelos de conjunto de dados.

Figura 13: Livros de visita

Root > teste

teste

Search this dataverse...

Find Advanced Search

☒ Dataverses (0)

☒ Datasets (0)

☐ Files (0)

This dataverse currently has no dataverses, datasets, or files. You

Contact Share Publish Link Edit

teste (Dataverse.org)

teste

General Information

Theme + Widgets

Permissions

Groups

Dataset Templates

Dataset Guestbooks

Featured Dataverses

Delete Dataverse

Fonte: Elaboração própria

2. Depois de clicar em Modelos de conjunto de dados, você será levado à página Modelos de conjunto de dados. Nesta página, você pode primeiro decidir usar os modelos de conjunto de dados do seu dataverse pai depois criar um novo modelo de conjunto de dados ou fazer as duas coisas.

3. Clique no Criar modelo de conjunto de dados para começar. Você verá que o modelo é igual à página de criação de conjunto de dados com um campo adicional na parte superior da página para adicionar um nome ao modelo.

4. Depois de adicionar informações aos campos de metadados para as quais você tem informações e clicar em Salvar e adicionar termos, você será levado à página onde poderá adicionar Termos de uso e acesso personalizados. Se você não precisar de Termos de uso e acesso personalizados, clique no modelo Salvar conjunto de dados e apenas os campos de metadados serão salvos.

5. Depois de clicar em Salvar modelo de conjunto de dados, você será levado de volta à página Gerenciar modelos de conjunto de dados e visualizará seu modelo listado agora com as opções de predefinição, edição, exibição ou exclusão.

6. Um dataverse não precisa ter um modelo padrão e os usuários podem selecionar qual modelo eles gostariam de usar na página Criar conjunto de dados.

7. No botão Exibir na página "Gerenciar modelos de conjunto de dados", será possível visualizar quais campos de metadados têm informações preenchidas. Observe que a capacidade de escolher quais campos de metadados estão ocultos, obrigatórios ou opcionais é feita na página Informações gerais do dataverse.

4.8 Livro de visitas de Datasets

Os livros de visitas permitem coletar dados sobre quem está baixando os arquivos dos seus conjuntos de dados. Você pode decidir coletar informações da conta (nome de usuário, nome e sobrenome, afiliação, etc.), além de criar perguntas personalizadas (por exemplo, para o que planeja usar esses dados?). Você também pode fazer o download dos dados coletados dos livros de visitas ativados como arquivos do Excel para armazenar e usar fora do Dataverse.

Para criar um livro de visitas:

1. Depois de criar um dataverse, clique no botão Editar dataverse e selecione Conjunto de dados do livro de visitas

2. Por padrão, os livros de visitas criados no dataverse em que o dataverse está, serão exibidos. Se você não quiser usar ou ver esses livros de visitas, desmarque a caixa de seleção: Incluir Livros de Visitas do Root Dataverse.

3. Para criar um novo livro de visitas, clique no botão Criar livro de visitas do conjunto de dados no lado direito da página.

4. Nomeie o livro de visitas, determine as informações da conta que você gostaria que fossem solicitadas (todos os campos de informações da conta aparecem quando alguém baixa um arquivo) e adicione Perguntas personalizadas (podem ser necessárias ou não).

5. Clique no botão Criar livro de visitas do conjunto de dados depois de terminar.

Depois de criar um livro de visitas, observará que existem várias opções para um livro de visitas e aparecerá na lista de livros de visitas.

Se você deseja usar um livro de visitas que criou, primeiro clique no botão no menu ativo que diz Ativar. Depois que um livro de visitas for ativado, você poderá acessar os Termos de Licença + para um conjunto de dados e selecionar um livro de visitas para ele. Também há opções para visualizar, copiar, editar ou excluir um livro de visitas. Depois que alguém fizer o download de um arquivo em um conjunto de dados em que um livro de visitas foi atribuído, uma opção para baixar os dados coletados será exibida.

4.9 Dataverses destaques

Os Dataverses em destaque são uma maneira de exibir sub-dataverses no seu dataverse que você deseja destacar para que as pessoas vejam facilmente quando o visitam.

Clique em Dataverses em destaque e um pop-up aparecerá. Selecione quais sub-dataverses você gostaria que fossem exibidos. Dataverses em destaque podem ser usados apenas com dataverses publicados.

4.10 Vinculação de Datasets

A vinculação de conjunto de dados permite que um proprietário do dataverse "vincule" seu dataverse a um conjunto de dados que exista fora desse dataverse, para que apareça na lista de conteúdos do dataverse sem realmente estar nesse dataverse. Você pode vincular os conjuntos de dados de outros usuários ao seu dataverse, mas isso não transfere a edição ou outras permissões especiais para você. O conjunto de dados vinculado ainda estará sob o controle do usuário original.

Por exemplo, os pesquisadores que trabalham em um estudo colaborativo entre instituições podem cada um vincular seus próprios dados institucionais individuais ao conjunto de dados colaborativo, facilitando a localização do estudo pelas partes interessadas de cada instituição.

Para vincular um conjunto de dados, você precisará da sua conta com a permissão "Adicionar conjunto de dados" no Dataverse que está fazendo o link. Se você criou o dataverse, já deve ter essa permissão, mas, caso contrário, precisará solicitar ao administrador desse dataverse que atribua essa permissão à sua conta. Você não precisa de permissões especiais no conjunto de dados que está sendo vinculado.

Para vincular um conjunto de dados ao seu dataverse, você deve navegar até esse

conjunto de dados e clicar no botão branco "Link" no canto superior direito da página do conjunto de dados. Isso abrirá uma janela na qual você pode digitar o nome do dataverse que deseja vincular o conjunto de dados. Selecione seu dataverse e clique no botão Salvar. Isso estabelecerá o link, e o conjunto de dados aparecerá agora no seu dataverse. Atualmente, não há como remover links estabelecidos na interface do usuário.

4.11 Vinculação de Dataverses

A vinculação de conjunto de dados permite que um proprietário do dataverse "vincule" seu dataverse a um conjunto de dados que exista fora desse dataverse, para que ele apareça na lista de conteúdos do dataverse sem realmente estar nesse dataverse. Você pode vincular os conjuntos de dados de outros usuários ao seu dataverse, mas isso não transfere a edição ou outras permissões especiais para você. O conjunto de dados vinculado ainda estará sob o controle do usuário original.

Da mesma forma que a vinculação de conjuntos de dados, a vinculação de dataverses permite que um proprietário de um dataverse "vincule" seu dataverse a outro dataverse, para que o dataverse que está sendo vinculado apareça na lista de conteúdos do vinculador de dados sem realmente estar nesse dataverse. Atualmente, a capacidade de vincular um dataverse a outro dataverse é um recurso exclusivo do superusuário.

4.12 Publique seu Dataverse

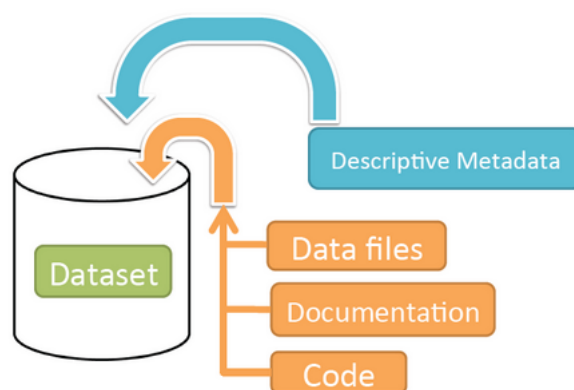
Quando o seu dataverse estiver pronto para ir a público, vá para a sua página, clique no botão "Publicar" no lado direito da página. Um pop-up aparecerá para confirmar que você está realmente pronto para publicar, uma vez que um dataverse é tornado público, esta ação não pode mais ser desfeita.

Manual do usuário: Dataset + Gerenciamento de arquivos

Um conjunto de dados no Dataverse é um contêiner para seus dados, documentação, código e os metadados que descrevem esse conjunto de dados.

Figura 14: Esquema de um Dataset

Schematic Diagram of a **Dataset** in Dataverse 4.0



Container for your data, documentation, and code.

Fonte: DATAVERSE PROJECT, 2019

5.1 Metadados suportados

Um conjunto de dados contém três níveis de metadados:

1. Metadados de citação: qualquer metadado necessário para gerar uma citação de dados e outros metadados gerais que possam ser aplicados a qualquer conjunto de

dados;

2. Metadados específicos de domínio: atualmente, com suporte específico para conjuntos de dados de Ciências Sociais, Ciências da Vida, Geoespaciais e Astronomia; e

3. Metadados no nível do arquivo: varia de acordo com o tipo de arquivo de dados.

5.1.1 Formatos de exportação de metadados

Depois que um conjunto de dados é publicado, seus metadados são exportados em vários formatos. Um botão na guia de metadados da página do conjunto de dados permitirá que um usuário exporte os metadados da versão publicada mais recente do conjunto de dados. Os formatos de exportação atualmente suportados são:

- Dublin Core
- DDI (Iniciativa de Documentação de Dados)
- DataCite 4
- JSON (formato nativo do Dataverse)
- OAI_ORE
- OpenAIRE

5.2 Adicionando um novo Dataset

1. Navegue para o dataverse no qual você deseja adicionar um conjunto de dados.
2. Clique no botão "Add Data (Adicionar dados)" e selecione "New Dataset (Novo conjunto de dados)" no menu suspenso.
3. Para começar rapidamente, insira no mínimo todos os campos obrigatórios com um asterisco (por exemplo, título do conjunto de dados, autor, descrição, e-mail de contato e assunto) para obter uma citação de dados com um DOI.
4. Role para baixo até a seção "Arquivos", e clique em "Select Files to Add (Selecionar arquivos a serem adicionados)" para adicionar todos os arquivos relevantes ao seu conjunto de dados. Você também pode enviar seus arquivos diretamente do seu Dropbox. Você pode arrastar e soltar ou selecionar vários arquivos de uma vez da área de trabalho diretamente no widget de upload. Seus arquivos aparecerão abaixo do botão "Select Files to Add (Selecionar arquivos a serem adicionados)", onde você pode adicionar uma descrição e tags (através do botão "Edit tag (Editar tag)") para cada arquivo. Além disso, uma soma de verificação MD5 será adicionada para cada arquivo. Se você

enviar um arquivo tabular, uma Impressão Digital Numérica Universal (UNF) será adicionada a esse arquivo.

5. Clique no botão "Save Dataset (Salvar conjunto de dados)" quando terminar. Seu conjunto de dados não publicado será criado.

Você pode adicionar metadados adicionais depois de concluir a criação do conjunto de dados inicial, indo em Editar conjunto de dados > Metadados.

5.3 Upload de Arquivo

O software Dataverse oferece vários métodos de upload de arquivos para um conjunto de dados. Esses métodos de upload são configuráveis pelo administrador de uma instalação do Dataverse, portanto, talvez você não veja algumas dessas opções no site do Dataverse que está usando.

Se houver várias opções de upload disponíveis, você deverá escolher qual delas usar para o seu conjunto de dados. Um conjunto de dados pode usar apenas um método de upload. Depois de fazer o upload de um arquivo usando um dos métodos de upload disponíveis, esse método é bloqueado para esse conjunto de dados. Se você precisar mudar os métodos de upload para um conjunto de dados que já contém arquivos, entre em contato com o Suporte clicando no link Suporte, na parte superior do aplicativo.

Você pode fazer upload de arquivos para um conjunto de dados enquanto cria esse conjunto de dados. Você também pode fazer upload de arquivos depois de criar um conjunto de dados, clicando no botão "Edit (Editar)" na parte superior da página do conjunto de dados e na lista suspensa selecionando "Arquivos (Upload)" ou clicando no botão "Upload files) Upload de arquivos" acima da tabela de arquivos em Arquivos aba. Em qualquer uma das opções, você será direcionado para a página Carregar arquivos para esse conjunto de dados.

Certos tipos de arquivo no Dataverse são suportados por funcionalidades adicionais, que podem incluir o download em diferentes formatos, subconjuntos, preservação de metadados no nível do arquivo, citação de dados no nível do arquivo, e exploração através da visualização e análise de dados.

5.4 Manuseio de Arquivo

Certos tipos de arquivo no Dataverse são suportados por funcionalidades adicionais, que podem incluir o download em diferentes formatos, subconjuntos, preservação de metadados no nível do arquivo, citação de dados no nível do arquivo; e exploração através da visualização e análise de dados. Consulte as seções abaixo para obter informações sobre funcionalidades especiais para tipos de arquivos específicos.

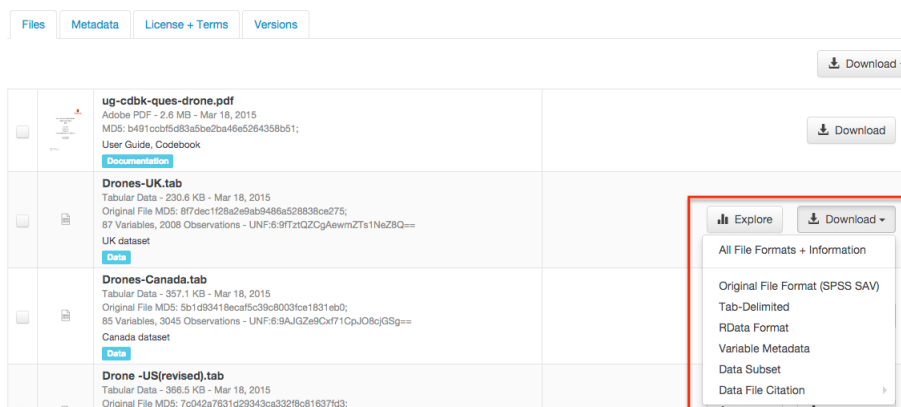
5.5 Arquivos de dados tabulares

Arquivos em determinados formatos - Stata, SPSS, R, Excel (xlsx) e CSV - podem ser ingeridos como dados tabulares. Os arquivos de dados tabulares podem ser mais explorados e manipulados.

Opções de download adicionais disponíveis para dados tabulares (encontradas no mesmo menu suspenso no botão "Download"):

- Como dados delimitados por tabulação (com os nomes das variáveis na primeira linha);
- O arquivo original carregado pelo usuário;
- Salvo como dados R (se o arquivo original não estava no formato R);
- Metadados variáveis (como um arquivo XML do DDI Codebook);
- Citação de arquivo de dados (atualmente no formato RIS, EndNote XML ou BibTeX);
- Tudo acima, como um pacote compactado.

Figura 15: Download



Fonte: DATAVERSE PROJECT, 2019

5.6 Arquivos compactados

Os arquivos compactados no formato .zip são descompactados automaticamente. Se um arquivo .zip falhar ao descompactar por qualquer motivo, ele será carregado como está. Se o número de arquivos internos for superior a um limite definido (1.000 por padrão, configurável pelo administrador), você receberá uma mensagem de erro e o arquivo .zip será carregado como está.

Se o arquivo .zip carregado contiver uma estrutura de pastas, o Dataverse acompanhará essa estrutura. A localização de um arquivo nessa estrutura de pastas é exibida nos metadados do arquivo como o Caminho do arquivo. Quando você baixa o conteúdo

do conjunto de dados, essa estrutura de pastas é preservada e os arquivos aparecem em seus locais originais.

Esses nomes de pasta estão sujeitos a regras estritas de validação. Somente os seguintes caracteres são permitidos: alfanuméricos, '_', '-', '.' E ' (espaço em branco). Quando um arquivo zip é carregado, os nomes das pastas são limpos automaticamente, com os caracteres inválidos substituídos pelo caractere '.' Quaisquer sequências de pontos são substituídas por um único ponto. Por exemplo, o nome da pasta data & info / code = 137 será convertido em data.info/code.137. Ao fazer o upload pela interface da Web, o usuário pode alterar ainda mais os valores no formulário de edição apresentado antes de clicar no botão "Save (Salvar)".

Se você carregar vários arquivos .zip em um conjunto de dados, todos os subdiretórios idênticos entre vários .zips serão mesclados quando o usuário fizer o download do conjunto de dados completo.

5.7 Outros tipos de arquivo

Existem várias opções avançadas disponíveis para determinados tipos de arquivo. Arquivos de imagem: arquivos .jpg, .png e .tif podem ser selecionados como miniatura padrão de um conjunto de dados. A miniatura selecionada aparecerá no cartão de resultados da pesquisa para esse conjunto de dados. Arquivos SPSS: os arquivos SPSS podem ser marcados com o idioma em que foram originalmente codificados. Isso pode ser encontrado clicando em Opções avançadas e selecionando o idioma na lista fornecida.

5.8 Editar Arquivos

5.8.1 Editar Arquivo de Metadados

Vá para o conjunto de dados que você deseja editar, onde verá a lista de arquivos. Selecione os arquivos que você deseja editar usando a caixa de seleção Selecionar tudo ou selecionando arquivos individualmente. Em seguida, clique no botão "Edit files (Editar arquivos)" acima da tabela de arquivos e, no menu suspenso, selecione se você deseja:

- Excluir os arquivos selecionados;
- Editar os metadados do arquivo (nome do arquivo, descrição) para os arquivos selecionados;
- Restringir os arquivos selecionados;
- Retirar restrições dos arquivos selecionados (somente se os arquivos selecionados forem restritos);
- Adicionar tags aos arquivos selecionados.

Você não precisará sair da página do conjunto de dados para concluir essas ações, exceto a edição de metadados do arquivo, que o levará à página Editar Arquivos. Lá, você terá que clicar no botão "Save Changes (Salvar alterações)" para aplicar suas edições e retornar à página do conjunto de dados.

Se você restringir os arquivos, também surgirá um pop-up solicitando que você preencha os Termos de acesso dos arquivos. Se os Termos de acesso já existirem, você será solicitado a confirmá-los. Observe que algumas instalações do Dataverse não permitem restrições de arquivo.

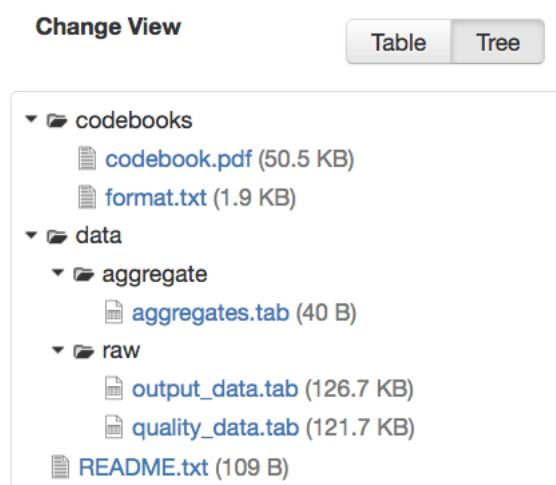
5.8.2 Caminho do Arquivo

O campo de metadados do caminho do arquivo é a maneira da Dataverse de representar a localização de um arquivo em uma estrutura de pastas. Quando um usuário carrega um arquivo .zip que contém uma estrutura de pastas, o Dataverse preenche automaticamente as informações do Caminho do arquivo para cada arquivo contido no arquivo .zip. Se um usuário fizer o download do conjunto de dados completo ou de uma seleção de arquivos, ele receberá uma estrutura de pastas com cada arquivo posicionado de acordo com o caminho do arquivo.

O caminho do arquivo de um arquivo pode ser adicionado ou editado manualmente na página Editar arquivos. Alterar o caminho do arquivo de um arquivo altera sua localização na estrutura de pastas criadas quando um usuário baixa o conjunto de dados completo ou uma seleção de arquivos dele.

Se houver mais de um arquivo no conjunto de dados e uma vez que pelo menos um deles tiver um caminho de diretório não vazio, a Página do Conjunto de Dados apresentará uma opção para alternar entre a exibição de tabela tradicional e a exibição em árvore dos arquivos mostrando a estrutura da pasta, como no exemplo abaixo:

Figura 16: Visualização



Fonte: DATAVERSE PROJECT, 2019

5.8.3 Tags do Arquivo

As tags de arquivo são compostas por tags personalizadas de categoria, (documentação, dados, código) e de tabela (evento, genômica, geoespacial, rede, painel, pesquisa, série temporal). Use os menus suspensos de seleção e a entrada de tag de arquivo personalizada para aplicar essas tags aos arquivos selecionados. Há também um recurso Excluir Tags que, se marcado, permitirá que você exclua as tags de arquivo não utilizadas dentro desse conjunto de dados.

5.9 Substituir Arquivo

Nos casos em que você deseja revisar um arquivo existente, em vez de adicionar um novo, é possível fazê-lo usando o recurso Substituir arquivo. Isso permitirá que você acompanhe o histórico desse arquivo nas versões do seu conjunto de dados, antes e depois de substituí-lo. Isso pode ser útil para atualizar seus dados ou corrigir erros nos dados. Como a substituição de um arquivo cria um link explícito entre a versão anterior do conjunto de dados e a versão atual, o recurso de substituição de arquivo não está disponível para rascunhos de conjuntos de dados não publicados. Observe também que a substituição de um arquivo não carrega automaticamente os metadados desse arquivo, mas, uma vez que o arquivo é substituído, seus metadados originais ainda podem ser encontrados fazendo referência à versão anterior do arquivo na guia "Versões" da página do arquivo.

Para substituir um arquivo, vá para a página do arquivo, clique no botão "Edit (Editar)" e, na lista suspensa, selecione "Replace (Substituir)". Isso o levará à página Substituir arquivo, onde você poderá ver os metadados da versão mais recente publicada do arquivo e fazer upload do seu arquivo de substituição. Depois de fazer o upload do arquivo de substituição, você poderá editar seu nome, descrição e tags. Quando terminar, clique no botão "Save Changes (Salvar alterações)".

Após a substituição bem-sucedida de um arquivo, será criada uma nova versão de rascunho do conjunto de dados. Um resumo de suas ações será gravado na guia "Versões", na página do conjunto de dados e na página do arquivo. A guia Versões permite acessar todas as versões anteriores do arquivo em todas as versões anteriores do seu conjunto de dados, incluindo a versão antiga do arquivo antes de substituí-lo.

5.10 Termos

Os termos do conjunto de dados podem ser visualizados e editados na guia Termos da página do conjunto de dados ou no botão suspenso Editar de um conjunto de dados. Lá, você pode configurar como os usuários podem usar seus dados depois de baixá-los (renúncia CC0 ou Termos de Uso personalizados), como eles podem acessar seus dados se você tiver arquivos restritos (termos de acesso) e ativar um Livro de Visitas para o seu conjunto de dados, para que você possa rastrear quem está usando seus dados e para quais finalidades.

5.10.1 CC0 Domínio Público Dedicação

Por padrão, todos os novos conjuntos de dados criados pela interface do usuário da web do Dataverse recebem uma Dedicação de Domínio Público Creative Commons CC0.

Se um proprietário de dados achar que CC0 não é adequado para seus dados, poderá inserir Termos de Uso personalizados.

5.10.2 Termos de Uso Customizados para Datasets

Se você não conseguir usar a Dedicação ao Domínio Público CC0 em seus conjuntos de dados, poderá especificar seus próprios Termos de Uso personalizados. Para fazer isso, selecione “Não, não aplique CC0 - “ Dedicação ao domínio público ”, e uma caixa de texto Termos de uso será exibida, permitindo que você insira seus próprios termos de uso personalizados para o seu conjunto de dados. Para adicionar mais informações sobre os Termos de Uso, fornecemos campos como Permissões Especiais, Restrições, Requisitos de Citação, etc.

5.10.3 Arquivos Restritos + Termos de Acesso

Se você restringir qualquer arquivo no seu conjunto de dados, será solicitado um pop-up para inserir os Termos de Acesso para os dados. Isso também pode ser editado na guia Termos ou selecionando Termos no botão suspenso "Edit (Editar)" no conjunto de dados. Você também pode permitir que os usuários solicitem acesso aos seus arquivos restritos, ativando "Solicitar acesso". Para adicionar mais informações sobre os Termos de acesso, fornecemos campos como Local de acesso a dados, Status de disponibilidade, Contato para acesso, etc.

5.10.4 Livro de Visitas

Aqui se habilita o Livro de Visitas específico para o seu conjunto de dados, que é configurado no nível do Dataverse.

5.10.5 Nível de Dataset

Os administradores ou curadores de um conjunto de dados podem atribuir funções e permissões aos usuários desse conjunto de dados. Se você é administrador ou curador de um conjunto de dados, pode acessar a página de permissões do conjunto de dados clicando no botão "Edit (Editar)", realçando "Permissões" na lista suspensa e clicando em "Dataset (Conjunto de dados)".

Ao acessar a página de permissões de um conjunto de dados, você verá duas seções:

Usuários / grupos: aqui você pode atribuir funções a usuários ou grupos específicos, determinando quais ações eles têm permissão para executar no seu conjunto de dados. Você também pode fazer referência a uma lista de todos os usuários que têm

funções atribuídas a eles para o seu conjunto de dados e remover as funções, se desejar. Alguns dos usuários listados podem ter funções atribuídas no nível do dataverse, nesse caso, essas funções podem ser removidas apenas da página de permissões do dataverse.

Funções: Aqui você pode fazer referência a uma lista completa de funções que podem ser atribuídas aos usuários do seu conjunto de dados. Cada função lista as permissões que oferece.

5.10.6 Nível de Arquivo

Se arquivos específicos no seu conjunto de dados tiverem acesso restrito, você poderá conceder acesso a usuários ou grupos específicos a esses arquivos, mantendo-os restritos ao público em geral. Se você é administrador ou curador de um conjunto de dados, pode acessar a página de permissões no nível do arquivo clicando no botão "Edit (Editar)", realçando "Permissões" na lista suspensa e clicando em "Arquivo".

Ao acessar a página de permissões no nível do arquivo de um conjunto de dados, você verá duas seções:

Usuários / Grupos: Aqui você pode ver quais usuários ou grupos tiveram acesso a quais arquivos. Você pode clicar no botão "Conceder acesso a usuários / grupos" para ver uma caixa onde você pode conceder acesso a arquivos específicos dentro do seu conjunto de dados para usuários ou grupos específicos. Se algum usuário tiver solicitado acesso a um arquivo no seu conjunto de dados, você poderá conceder ou rejeitar a solicitação de acesso aqui.

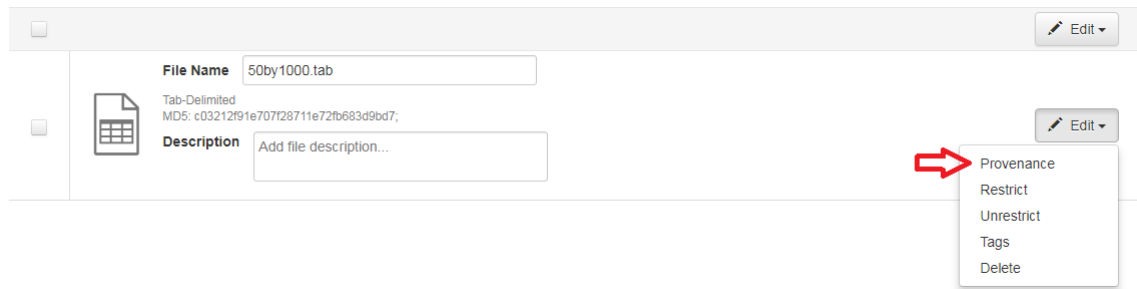
Arquivos restritos: Nesta seção, você pode ver as mesmas informações, mas discriminadas por cada arquivo individual no seu conjunto de dados. Para cada arquivo, você pode clicar no botão "Atribuir acesso" para ver uma caixa onde você pode conceder acesso a esse arquivo para usuários ou grupos específicos.

5.11 Proveniência dos Dados

A proveniência de dados é um registro de onde seus dados vieram e como atingiram sua forma atual. Ele descreve a origem de um arquivo de dados, quaisquer transformações feitas nesse arquivo e quaisquer pessoas ou organizações associadas a esse arquivo. A proveniência de um arquivo de dados pode ajudar na reprodutibilidade e conformidade com os regulamentos legais. O Dataverse disponibiliza apenas informações de proveniência para aqueles que têm permissões de edição no seu conjunto de dados.

O Dataverse aceita informações de proveniência de duas formas: um Arquivo de proveniência ou uma Descrição de proveniência de texto livre. Você pode anexar essas informações de proveniência aos seus arquivos de dados no Dataverse como parte do processo de upload do arquivo, clicando em Editar -> Proveniência:

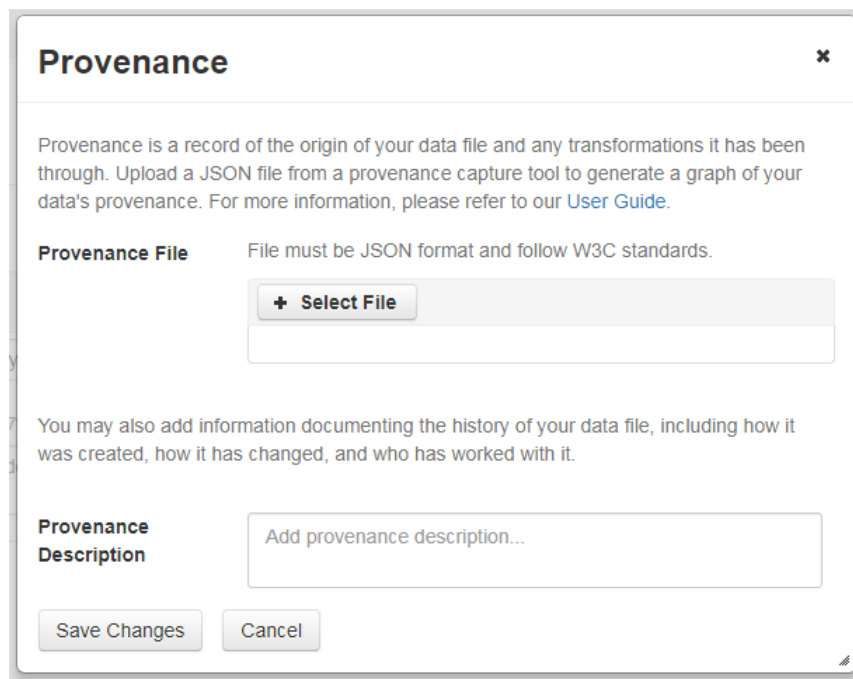
Figura 17: Proveniência



Fonte: DATAVERSE PROJECT, 2019

Isso abrirá uma janela onde você pode adicionar seu arquivo de proveniência e / ou descrição da proveniência:

Figura 18: Arquivo de proveniência



Fonte: DATAVERSE PROJECT, 2019

Depois de fazer o upload de um arquivo de proveniência, o Dataverse precisará de algumas informações adicionais para conectá-lo com precisão ao seu arquivo de dados. Quando o upload do arquivo de origem terminar, uma caixa de entrada chamada "Connect entity" aparecerá sob o arquivo. Os arquivos de proveniência contêm uma lista de "entidades", que incluem seu arquivo de dados e quaisquer objetos associados a ele (por exemplo, um gráfico, um corretor ortográfico, etc.). Você precisará informar ao Dataverse qual entidade no arquivo de proveniência representa seu arquivo de dados. Você pode digitar o nome da entidade na caixa ou clicar na seta ao lado da caixa e selecionar a entidade em uma lista de todas as entidades no arquivo de proveniência.

Depois de fazer o upload do seu arquivo de proveniência e conectar a entidade

adequada, você pode clicar no botão "Visualizar" para visualizar o JSON bruto do arquivo de proveniência. Isso pode ajudar a confirmar que você enviou o arquivo certo. Verifique-o novamente, porque o arquivo de proveniência será permanente assim que for finalizado. Nesse ponto, você não poderá substituir, remover ou editar o arquivo de proveniência. Isso garante que o arquivo de proveniência mantenha um registro estável e imutável do histórico do arquivo de dados.

Essa finalização do arquivo de proveniência ocorre em diferentes pontos, dependendo do status do seu arquivo de dados. Se este for um novo arquivo de dados que nunca foi publicado antes, o arquivo de proveniência associado será tornado permanente assim que você publicar o conjunto de dados. Se esse arquivo de dados foi publicado em uma versão anterior do seu conjunto de dados, o arquivo de proveniência associado será tornado permanente assim que você fizer o upload do arquivo de proveniência e clique em "Save Changes (Salvar alterações)" no pop-up de aviso.

Uma Descrição de proveniência permite adicionar mais informações de proveniência além ou no lugar de um arquivo de proveniência. Este é um campo de texto livre que permite que você insira qualquer informação que considere relevante para os interessados em aprender sobre a procedência de seus dados. Esse pode ser um bom lugar para descrever fatores de proveniência, como qual sistema operacional você usou ao trabalhar com o arquivo de dados, quais funções ou bibliotecas você usou, como os dados foram mesclados no arquivo, qual versão do arquivo você usou etc. A descrição não é tão útil ou confiável quanto um arquivo de proveniência, mas ainda pode fornecer valor. Ao contrário do arquivo de proveniência, a descrição da proveniência nunca é permanente: você sempre pode editá-lo, removê-lo ou substituí-lo a qualquer momento.

Você pode retornar para anexar a proveniência ao seu arquivo de dados posteriormente, clicando no botão "Add + Edit Metadata (Adicionar + Editar Metadados)" na página do arquivo e, em seguida, clicando no botão "Edit -> Provenance (Editar -> Origem)".

5.12 Thumbnails e Widgets

5.12.1 Thumbnails

As imagens em miniatura podem ser atribuídas a um conjunto de dados manual ou automaticamente. A miniatura de um conjunto de dados aparece no cartão de resultados da pesquisa para esse conjunto de dados e na própria página do conjunto de dados. Se um conjunto de dados contiver um ou mais arquivos de dados que o Dataverse reconheça como uma imagem, uma dessas imagens será selecionada automaticamente como miniatura do conjunto de dados.

Se você deseja selecionar manualmente a miniatura do seu conjunto de dados, clique no botão "Edit (Editar)" no seu conjunto de dados e selecione "Miniaturas + Widgets" no menu suspenso.

Nesta página, na guia Miniatura, você verá três ações possíveis.

Selecionar arquivo disponível: clique no botão "Selecionar miniatura" para escolher uma imagem do seu conjunto de dados para usá-la como miniatura do conjunto de dados.

Carregar novo arquivo: carregue um arquivo de imagem do seu computador para usá-lo como miniatura do conjunto de dados. Embora, por padrão, sua imagem em miniatura seja desenhada a partir de um arquivo no seu conjunto de dados, isso permitirá que você envie um arquivo de imagem separado para usar como miniatura do conjunto de dados. Este arquivo de imagem enviado será usado apenas como miniatura do conjunto de dados, ele não será armazenado como um arquivo de dados no seu conjunto de dados.

Remover miniatura: se você clicar no botão "Remove (Remover)" abaixo da imagem em miniatura, removerá a miniatura atual do conjunto de dados. O conjunto de dados será revertido para exibir um ícone padrão básico como miniatura do conjunto de dados.

Quando terminar nesta página, clique em "Save Changes (Salvar alterações)" para salvar o que você fez. Se preferir, também é possível definir um arquivo de imagem em seu conjunto de dados como miniatura, selecionando o arquivo, indo em Edit Files -> Metadata e usando o botão "Set Thumbnail".

5.12.2 Widgets

O recurso Widgets fornece código para seu site pessoal, para que seu conjunto de dados possa ser exibido. Existem dois tipos de Widgets para um conjunto de dados: o Widget do Conjunto de Dados e o Widget de Citação do Conjunto de Dados. Os widgets podem ser encontrados indo para a página do conjunto de dados, clicando no botão "Edit (Editar)" (aquele com o ícone de lápis) e selecionando "Miniaturas + Widgets" no menu suspenso.

Na guia Widgets, você pode copiar e colar os trechos de código do widget que você deseja adicionar ao seu site. Se você precisar ajustar a altura do widget em seu site, poderá fazê-lo editando o parâmetro `heightPx = 500` no snippet de código.

5.12.3 Widget dataset

O Widget do conjunto de dados permite que a citação, metadados, arquivos e termos do seu conjunto de dados sejam exibidos no seu site. Quando alguém baixa um arquivo de dados no widget, ele é baixado diretamente dos conjuntos de dados no seu site. Se um arquivo for restrito, eles serão direcionados ao seu dataverse para efetuar login, em vez de fazer login através do widget em seu site.

Para editar seu conjunto de dados, você precisará retornar ao repositório do Dataverse onde o conjunto de dados está armazenado. Você pode fazer isso facilmente

clikando no link "Data Stored in (Name) Dataverse (Dados armazenados no (nome) dataverse)" localizado na parte inferior do widget.

5.12.4 Widget Citação de dataset

O Widget de Citação do Conjunto de Dados fornecerá uma citação para o seu conjunto de dados no site pessoal ou do projeto. Os usuários podem baixar a citação em vários formatos usando o botão Citar Dados. O URL persistente na citação direcionará os usuários para o conjunto de dados no seu dataverse.

5.13 Publique Datasets

Ao publicar um conjunto de dados (disponível para um administrador, curador ou qualquer função personalizada que tenha esse nível de permissão atribuído), você o disponibiliza ao público geral para que outros usuários possam procurá-lo. Quando seu conjunto de dados estiver pronto para ser publicado, vá para a página do seu conjunto de dados e clique no botão "Publish (Publicar)" no lado direito da página. Um pop-up aparecerá para confirmar que você está realmente pronto para publicar, pois uma vez que um conjunto de dados é tornado público, ele não pode mais ser publicado.

Sempre que você edita seu conjunto de dados, você pode publicar uma nova versão do conjunto de dados. O botão Publicar conjunto de dados reaparecerá sempre que você editar os metadados do conjunto de dados ou adicionar um arquivo.

Antes de publicar seu conjunto de dados, a Citação de dados indicará que este é um rascunho, mas o texto "DRAFT VERSION" será removido assim que você publicar.

5.14 Submeter à Revisão

Se você tiver uma função de Colaborador (pode editar metadados, fazer upload e editar arquivos, editar Termos, Livro de Visitas e Enviar conjuntos de dados para revisão) em um Dataverse, poderá enviar seu conjunto de dados para revisão quando terminar de enviar e preencher todos os dados dos campos de metadados relevantes. Para enviar para revisão, vá para o seu conjunto de dados e clique no botão "Submit for Review (Enviar para revisão)", localizado ao lado do botão "Edit (Editar)" no canto superior direito. Depois de enviado para revisão: o administrador ou o curador deste Dataverse será notificado para revisar esse conjunto de dados antes de decidir "Publicar" o conjunto de dados ou "Retornar ao autor". Se o conjunto de dados for publicado, o colaborador receberá uma notificação dizendo que o mesmo agora se encontra publicado. Se o conjunto de dados for retornado ao autor, o colaborador desse conjunto de dados será notificado que precisa fazer modificações antes de poder ser enviado para revisão novamente.

5.15 URL Privada para Revisão de Datasets Não publicados

A criação de uma URL privada para o seu conjunto de dados permite que você compartilhe seu conjunto de dados (para visualização e download de arquivos) antes de ser

publicado para um amplo grupo de indivíduos que podem não ter uma conta de usuário no Dataverse. Qualquer pessoa para quem você enviar o URL privado não precisará fazer login no Dataverse para visualizar o conjunto de dados:

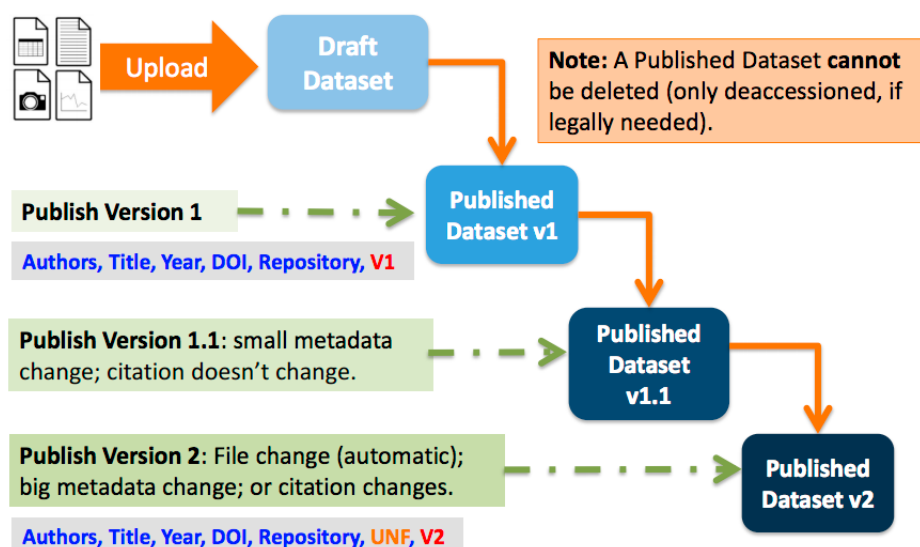
1. Vá para o seu conjunto de dados não publicado
2. Selecione o botão "Edit (Editar)"
3. Selecione "Private URL (URL privado)" no menu suspenso
4. No pop-up, selecione "Create Private URL (Criar URL privado)"
5. Copie o URL privado que foi criado para este conjunto de dados e agora pode ser compartilhado com qualquer pessoa que você deseja ter acesso para exibir ou baixar arquivos no seu conjunto de dados não publicado.

Para desativar um URL privado e revogar o acesso, siga as mesmas etapas acima até a etapa 3, quando você retornar ao pop-up, clique no botão "Disable Private URL (Desativar URL privada)".

5.16 Versões dos Datasets

O controle de versão é importante para o gerenciamento de dados de pesquisa de longo prazo, onde os metadados e / ou arquivos são atualizados ao longo do tempo. É usado para rastrear quaisquer metadados ou alterações de arquivo (por exemplo, fazendo upload de um novo arquivo, alterando metadados de arquivo, adicionando ou editando metadados) após a publicação do conjunto de dados.

Figura 19: Versões



Depois de editar seu conjunto de dados publicado, uma nova versão de rascunho desse conjunto de dados será criada. Para publicar esta nova versão do seu conjunto de dados, selecione o botão "Publish Dataset (Publicar conjunto de dados)" no canto superior direito da página. Se você estivesse na versão 1 do seu conjunto de dados, dependendo dos tipos de alterações feitas, seria solicitado publicar seu rascunho como versão 1.1 ou versão 2.0.

Se você adicionar um arquivo, seu conjunto de dados será automaticamente atualizado para uma versão principal (por exemplo, se você estivesse na versão 1.0, irá para a versão 2.0).

Na guia Versões de uma página de conjunto de dados, há uma tabela de versões que exibe o histórico de versões do conjunto de dados. Você pode usar os links do número da versão nesta tabela para navegar entre as diferentes versões do conjunto de dados, incluindo a versão de rascunho não publicada, se tiver permissão para acessá-lo.

Há também uma guia Versões na página do arquivo. A tabela de versões para um arquivo exibe as mesmas informações que o conjunto de dados, mas os resumos são filtrados para mostrar apenas as ações relacionadas a esse arquivo. Se uma nova versão do conjunto de dados fosse criada sem nenhuma alteração em um arquivo individual, o resumo da versão do arquivo para essa versão do conjunto de dados seria "Nenhuma alteração associada a esta versão".

5.16.1 Detalhes das Versões

Para visualizar exatamente o que mudou, iniciando da versão publicada originalmente para as versões publicadas subsequentes: clique na guia Versões na página do conjunto de dados para ver todas as versões e alterações feitas para esse conjunto de dados específico.

Depois de ter mais de uma versão (pode ser simplesmente a versão 1 e um rascunho), você pode clicar no link "View Details (Visualizar detalhes)" ao lado de cada resumo para saber mais sobre os campos e arquivos de metadados que foram adicionados ou editados. Você também pode clicar nas caixas de seleção para selecionar duas versões do conjunto de dados e clicar no botão "View Differences (Exibir diferenças)" para abrir o pop-up Detalhes das diferenças de versão e comparar as diferenças entre elas.

5.17 Desativar acessos de Dataset

Não é recomendável desativar um conjunto de dados ou uma versão de um conjunto de dados. Se você absolutamente precisar desaccessar, poderá desativar uma versão de um conjunto de dados ou de um conjunto de dados inteiro. Para desaccessar, vá para o conjunto de dados publicado (ou adicione um novo e publique-o), clique no botão "Edit (Editar)" e, no menu suspenso, selecione "Deaccession Dataset (Desaccessar dataset)". Se você tiver várias versões de um conjunto de dados, poderá selecionar aqui quais versões deseja desativar ou optar por desativar todo o conjunto de dados.

Você também deve incluir um motivo pelo qual esse conjunto de dados foi desativado. Selecione o motivo mais apropriado na lista suspensa de opções. Se você selecionar “Outro”, também deverá fornecer informações adicionais.

Adicione mais informações sobre o motivo por que foi desativado na caixa de texto livre. Se o conjunto de dados foi movido para um repositório ou site diferente, recomendamos que você inclua uma URL (de preferência persistente) para que os usuários possam continuar acessando esse conjunto de dados no futuro.

Se você desativar a versão publicada mais recentemente do conjunto de dados, mas nem todas as versões do conjunto de dados, poderá revisitar uma versão anterior e criar um novo rascunho não desativado para o conjunto de dados. Por exemplo, imagine que você tenha a versão 1 e a versão 2 de um conjunto de dados, ambas publicadas, e a versão 2 desaccessada. Você poderá editar a versão 1 do conjunto de dados e uma nova versão de rascunho será criada.

Uma página de lápide com os metadados básicos de citação sempre estará acessível ao público se eles usarem o URL persistente (Handle ou DOI) fornecido na citação para esse conjunto de dados. Os usuários não poderão ver nenhum dos arquivos ou metadados adicionais que estavam disponíveis anteriormente antes da desativação.



Manual do administrador

6.1 Painel de controle

O Dataverse oferece um painel de ferramentas administrativas apenas para superusuários. Se você é um superusuário logado, pode acessá-lo clicando no seu nome de usuário na barra de navegação e, em seguida, clicando em "Dashboard" no menu suspenso, fig. 20.

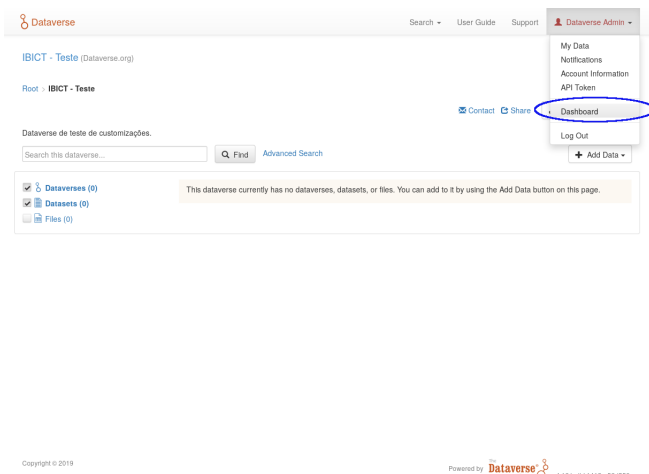


Figura 20: Dashboard

Você pode verificar se é um superusuário, verificando a cor do seu nome de usuário na barra de navegação. Se estiver vermelho, você tem as permissões corretas para usar o Painel. Os superusuários podem fornecer a outros usuários o status de superusuário via Administração do Usuário.

6.1.1 Harvest

Coletando Clientes

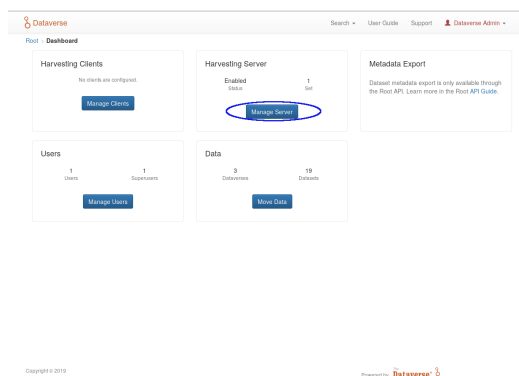
Essa ferramenta de painel permite configurar de quais outros repositórios sua instalação do Dataverse coleta registros de metadados. Você pode visualizar uma lista de clientes de coleta e adicioná-los, editar ou removê-los.

Coletando Servidores

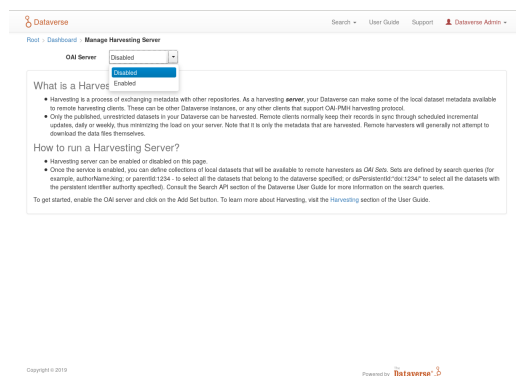
Essa ferramenta de painel permite definir conjuntos de conjuntos de dados locais para disponibilizar aos clientes de coleta remota. Você pode ver uma lista de conjuntos e adicioná-los, editar ou removê-los.

Depois de clicar na opção Dashboard, fig.20, o site abrirá uma janela com as opções, clique no *Manager Server*, do item *Harvest Server*, figura 21a. Na janela *Harvest Server*, figura 23a, estão as informações básicas sobre o colate.

Na parte logo acima da descrição tem o item *OAI Server*. Este item estará desabilitado, agora, modifique sua opção para *Enable*.



(a) Janela Dashboard



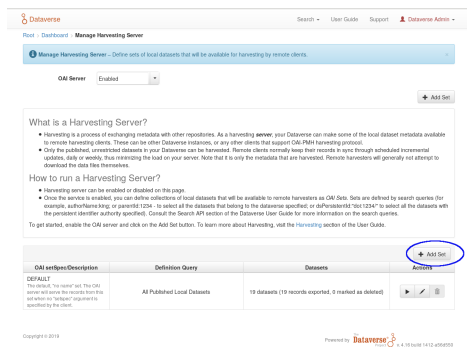
(b) Janela para habilitar o servidor OAI

Figura 21: Harvest server

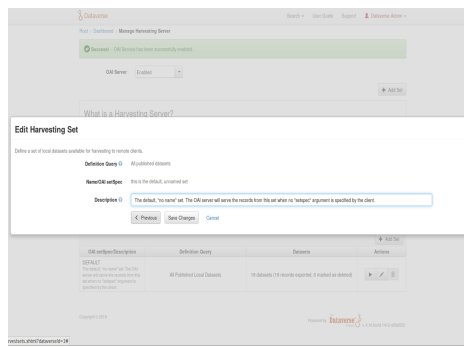
Para verificar se foi modificado, acesse o endereço *http:// URL DO DATAVERSE /oai*. Caso a página seja parecida com a figura 22, o OAI server está habilitado.



Figura 22: Página de verificação do OAI server



(a) Adicionar uma coleta



(b) Configurar coleta

Figura 23: Criar um critério de coleta

6.1.2 Usuários

Essa ferramenta de painel permite pesquisar uma lista de todos os usuários da sua instalação do Dataverse. Você pode remover funções das contas de usuário e atribuir ou remover o status de superusuário.

6.2 Gerenciando clientes de coleta

6.2.1 Seu Dataverse como um Harvester de Metadados

A coleta é um processo de troca de metadados com outros repositórios. Como cliente de coleta, seu Dataverse pode coletar registros de metadados de fontes remotas. Podem ser outras instâncias do Dataverse ou outros arquivos que suportam OAI-PMH, o protocolo de coleta padrão. Os registros de metadados coletados serão indexados e tornados pesquisáveis por seus usuários. Clicar em um conjunto de dados coletado nos resultados da pesquisa leva o usuário ao repositório original. Conjuntos de dados coletados não podem ser editados na instalação do Dataverse.

Os registros coletados podem ser mantidos em sincronia com o repositório original por meio de atualizações incrementais agendadas, diariamente ou semanalmente. Como alternativa, as coletas podem ser executadas sob demanda pelo administrador.

6.2.2 Gerenciando clientes de coleta

Para começar a coletar metadados de um repositório OAI remoto, primeiro crie e configure um Harvesting Client.

Os clientes são gerenciados na página "coleta de clientes", acessível através do Painel. Clique no botão Adicionar cliente para começar.

O processo de criação de um novo cliente ou edição de um cliente existente é amplamente autoexplicativo. Ele é dividido em etapas lógicas, de forma que permita ao usuário voltar e corrigir as entradas feitas anteriormente. O processo é interativo e o texto de orientação é fornecido. Por exemplo, é necessário que o usuário insira a URL do servidor OAI remoto. Quando ele clica em Avançar, o aplicativo tenta estabelecer uma conexão com o servidor para verificar se está funcionando e para obter informa-

ções sobre os conjuntos de registros de metadados e os formatos de metadados que ele suporta. As opções oferecidas ao usuário na próxima página serão baseadas nessas informações extras. Se o aplicativo falhar ao estabelecer uma conexão com o arquivo remoto no endereço especificado, ou se uma resposta inválida for recebida, o usuário terá a oportunidade de verificar e corrigir o URL digitado.

6.3 Gerenciando conjuntos e servidor de coleta

6.3.1 Seu Dataverse como servidor OAI

Como servidor de coleta, seu Dataverse pode disponibilizar alguns dos metadados do conjunto de dados local para clientes de coleta remota. Podem ser outras instâncias do Dataverse ou quaisquer outros clientes que suportam o protocolo de coleta OAI-PMH.

6.3.2 Como funciona?

Somente os conjuntos de dados publicados e irrestritos no seu Dataverse podem ser colhidos. Os clientes remotos normalmente mantêm seus registros sincronizados por meio de atualizações incrementais agendadas, diariamente ou semanalmente, minimizando a carga no servidor. Observe que são apenas os metadados que são coletados. Os harvesters remotos geralmente não tentam baixar os arquivos de dados associados aos conjuntos de dados coletados. O servidor de coleta pode ser ativado ou desativado na página "Servidor de coleta", acessível através do Painel. Por padrão, o servidor de coleta está desativado em um Dataverse novo e "pronto para uso".

6.3.3 Conjuntos OAI

Depois que o serviço é ativado, você define coleções de conjuntos de dados locais que estarão disponíveis para Harvester remotos como Conjuntos OAI. Os conjuntos são definidos por search queries. Qualquer consulta que encontre qualquer número de conjuntos de dados locais publicados (não coletados) pode ser usada para criar um conjunto OAI. Os conjuntos podem se sobrepor a locais de dados locais e podem incluir o número mínimo ou máximo de conjuntos de dados locais que você deseja.

Depois de inserir a search queries e clicar em Avançar, o número de resultados encontrados será mostrado na próxima tela. Dessa forma, se você perceber um número diferente do esperado, poderá voltar e tentar redefinir a consulta.

Alguns exemplos úteis de consultas de pesquisa para definir conjuntos OAI:

- Uma boa maneira de criar um conjunto que inclua todos os seus conjuntos de dados locais publicados é fazê-lo pela autoridade Identificador Único registrada no seu Dataverse, por exemplo:

```
dsPersistentId: "doi: 1234 /"
```

Observe que aspas duplas devem ser usadas, pois o valor do campo de pesquisa contém o símbolo de dois pontos.

Observe também que os termos de pesquisa que limitam os resultados aos conjuntos de dados publicados e locais são adicionados automaticamente à consulta. Portanto, você não precisa se preocupar com isso.

- Uma consulta para criar um conjunto para incluir os conjuntos de dados de uma coleta local específica:

```
parentId: NNN
```

Em que NNN é o ID do banco de dados do objeto DataVerse (consulte a tabela Dataverse do banco de dados SQL usado pelo aplicativo para verificar o ID do banco de dados).

- Uma consulta para encontrar todo o conjunto de dados de um determinado autor:

```
authorName: YYY
```

Onde AAA é o nome.

- Consultas complexas podem ser criadas com vários operadores lógicos AND e OR. Por exemplo,

```
(authorName: YYY OR authorName: ZZZ) E dsPublicationDate: NNNN
```

- Alguns exemplos adicionais de consulta: Para conjuntos de dados específicos usando um persistentID:

```
(dsPersistentId: 10.5000 /  
ZZYYXX / OU dsPersistentId: 10.5000 / XXYYZZ)
```

Para todos os conjuntos de dados em uma autoridade de ID específica:

```
dsPersistentId: 10.5000 / XXYYZZ
```

Para todos os dataverses com disciplinas de Astronomia e Astrofísica ou Ciências da Terra e do Ambiente:

```
(dvSubject: "Astronomia e astrofísica" OU  
dvSubject: "Terra e ciências ambientais")
```

Para todos os conjuntos de dados que contêm a palavra-chave "censura":

```
keywordValue: censura
```

6.3.4 Atualizações do conjunto OAI

Sempre que um novo conjunto de coleta é criado ou são feitas alterações em um conjunto existente, o conteúdo do conjunto é atualizado automaticamente - o aplicativo Dataverse encontrará os conjuntos de dados definidos pela consulta e tentará executar a exportação de metadados nos que ainda não foram exportados. Somente os conjuntos de dados para os quais a exportação foi concluída com êxito e os resultados armazenados em cache no sistema de arquivos são incluídos nos conjuntos OAI anunciados aos clientes da coleta.

Observe, no entanto, que as alterações feitas nos metadados reais do conjunto de dados não acionam automaticamente nenhum conjunto OAI correspondente para ser atualizado imediatamente. Por exemplo: digamos que você criou um conjunto OAI definido pela consulta de pesquisa

```
authorName: king
```

Que resultou em 43 registros de conjunto de dados. Se um novo conjunto de dados do mesmo autor for adicionado e publicado, isso não adicionará imediatamente o registro extra ao conjunto. Simplesmente seria muito custoso atualizar todos os conjuntos sempre que forem feitas alterações nos metadados.

No entanto, o conjunto OAI será atualizado automaticamente por um trabalho agendado de exportação de metadados que é executado todas as noites (às 2h da manhã, por padrão). Esse cronômetro de exportação é criado e ativado automaticamente sempre que o aplicativo é implantado ou reiniciado.

Ainda é possível, no entanto, fazer alterações como essa serem imediatamente refletidas no servidor OAI, acessando a página do Harvesting Server e clicando no ícone "Executar exportação" ao lado do conjunto OAI desejado.

6.4 Customização de metadados

O Dataverse possui um sistema flexível de metadados orientado a dados, alimentado por "blocos de metadados".

6.4.1 Introdução

Antes de iniciar a personalização de metadados no Dataverse, verifique a possibilidade de customização disponível na interface da web do Dataverse. É possível ocultar os campos e torná-los necessários, clicando em "Editar" no nível do dataverse, clicando em "Informações gerais" e fazendo ajustes em "Campos de metadados".

As possibilidades de personalização incluem:

- Editar e adicionar campos de metadados;
- Edição e adição de texto instrucional (dicas de ferramentas de rótulos de campo e marcas d'água de caixa de texto);

- Editar e adicionar vocabulários controlados;
- Alterar os campos que os depositantes devem usar para salvar os conjuntos de dados;
- Alterar como os valores salvos dos metadados são exibidos na interface do usuário.

De um modo geral, é mais seguro criar seu próprio bloco de metadados personalizado em vez de editar os blocos de metadados fornecidos com o Dataverse. Observe que os blocos de metadados enviados com o Dataverse são baseados em padrões (por exemplo, DDI para Ciências Sociais). O dataverse é enviado com um bloco de metadados de citação, que inclui campos obrigatórios necessários para atribuir IDs persistentes a conjuntos de dados e blocos de metadados específicos do domínio. As definições desses blocos são carregadas em uma instalação do Dataverse em valor separado por tabulação (TSV). Embora seja tecnicamente possível definir mais de um bloco de metadados em um arquivo TSV, é uma boa prática organizacional definir apenas um em cada arquivo.

6.4.2 Sobre o bloco de metadados TSV

Os propósitos de cada seção e propriedade do bloco de metadados TSV são:

1. metadataBlock

- Objetivo: Representa o bloco de metadados que está sendo definido.
- Cardinalidade:

0 ou mais por Dataverse

1 por definição de bloco de metadados

2. datasetField

- Objetivo: Cada entrada representa um campo de metadados a ser definido dentro de um bloco de metadados.
- Cardinalidade: 1 ou mais por metadadosBlock

3. controlVocabulary

- Objetivo: Cada entrada enumera um valor permitido para um determinado datasetField.
- Cardinalidade: zero ou mais por conjunto de dados

Cada uma das três seções principais possui conjuntos de propriedades:

Propriedades dos #metadataBlock

Tabela 1: metadataBlock

Propriedade	Próposito	Valores e restrições permitidos
name	Uma sequência definida pelo usuário usada para identificar um #metadataBlock	Sem espaços ou pontuação, exceto sublinhado. Por convenção, deve começar com uma letra e usar letras minúsculas em camelo. Não deve colidir com um campo com o mesmo nome na mesma ou em qualquer outra definição #datasetField, incluindo blocos de metadados definidos em outro local.
dataverseAlias	Se especificado, esse bloco de metadados estará disponível apenas para o dataverse designado aqui por seu alias e para filhos desse dataverse	Texto livre
displayName	Atua como um breve rótulo para exibição relacionada a este #metadataBlock	Deve ser relativamente breve. O limite é de 256 caracteres, mas nomes muito longos podem causar problemas de exibição.

Fonte: DATAVERSE PROJECT, 2019, adaptado

Propriedades #datasetField (field)

Tabela 2: datasetField

Propriedade	Próposito	Valores e restrições permitidos
name	Uma sequência definida pelo usuário usada para identificar um #datasetField. Mapeia diretamente para o nome do campo usado pelo Solr	(de DatasetFieldType.java) O nome interno do tipo DDI, sem espaços, etc. (do Solr) Os nomes dos campos devem consistir apenas em caracteres alfanuméricos ou sublinhados e não começar com um dígito. No momento, isso não é rigorosamente imposto, mas outros nomes de campo não terão suporte de primeira classe de todos os componentes e a compatibilidade com versões anteriores não é garantida. Os nomes com sublinhados iniciais e finais (por exemplo, _version_) são reservados. Não deve colidir com um campo com o mesmo nome em outra definição #metadataBlock ou com qualquer nome já incluído como campo no índice Solr
title	Atua como um breve rótulo para exibição relacionada a esse #datasetField	Deve ser relativamente breve
description	Usado para fornecer uma descrição do campo	Texto livre
watermark	Uma sequência a ser exibida inicialmente em um campo como um prompt para o que o usuário deve inserir	Texto livre
fieldType	Define o tipo de conteúdo que o campo, se não estiver vazio, deve conter	Nenhum, data, o email, texto, caixa de texto, URL, int, float
displayOrder	Controla a sequência na qual os campos são exibidos, para entrada e apresentação	Inteiro não negativo

displayFormat	Controla como o conteúdo é exibido para apresentação. O valor desse campo pode conter uma ou mais variáveis especiais. As tags HTML, provavelmente em conjunto com um ou mais desses valores, podem ser usadas para controlar a exibição do conteúdo na interface da web	
advancedSearchField	Especifique se este campo está disponível na pesquisa avançada	TRUE (disponível) or FALSE (não disponível)
allowControlledVocabulary	Especifique se os possíveis valores desse campo são determinados por valores na seção #controlledVocabulary	TRUE (controlado) or FALSE (não controlado)
allowmultiples	Especifique se este campo é repetitivo	TRUE (repetível) or FALSE (não repetível)
facetable	Especifique se o campo é facetável. Se um campo é facetável, ele aparece em "Procurar / pesquisar facetas" quando você edita "Informações gerais" para um dataverse. Definir esse valor como TRUE geralmente faz sentido para campos de vocabulário enumerados ou controlados, campos que representam identificadores e outros campos que provavelmente compartilham valores entre as entradas. É menos provável que faça sentido para campos que contenham descrições, números de ponto flutuante e outros valores que provavelmente sejam únicos	TRUE (controlado) or FALSE (não controlado)

displayoncreate	<p>Designar campos que devem ser exibidos durante a criação de um novo conjunto de dados, mesmo antes de o conjunto de dados ser salvo. Os campos não designados não serão exibidos até que o conjunto de dados seja salvo</p>	TRUE (exibir na criação) or FALSE (não exibir na criação)
required	<p>Especifique se o campo é obrigatório ou não. Isso significa que pelo menos uma instância do campo deve estar presente. Mais de um campo pode ser permitido, dependendo do valor de allowmultiples</p>	TRUE (requerido) or FALSE (não requerido)
parent	<p>Para subcampos, especifique o nome do pai ou campo contendo</p>	<p>Não deve resultar em uma referência cíclica. Deve fazer referência a um campo existente no mesmo #metadataBlock</p>
metadata_block_id	<p>Especifique o nome do #metadataBlock que contém esse campo</p>	<p>Deve fazer referência a um #metadataBlock existente. Como prática recomendada, o valor deve referenciar o #metadataBlock na definição atual</p>

Propriedades do #controlledVocabulary (enumeradas)

Tabela 3: controlledVocabulary

Propriedade	Propósito	Valores permitidos e restrições
DatasetField	Especifica o #datasetField ao qual esta entrada se aplica	Deve fazer referência a um #datasetField existente. Como prática recomendada, o valor deve referenciar um #datasetField na definição atual do bloco de metadados
Value	Uma string de exibição curta, representando um valor enumerado para este campo. Se a propriedade identificador estiver vazia, esse valor será usado como identificador	Texto livre
identifier	Uma sequência usada para codificar o valor enumerado selecionado de um campo. Se esta propriedade estiver vazia, o valor do campo "Valor" será usado como identificador	Texto livre
displayOrder	Controlar a ordem em que os valores enumerados são exibidos para seleção	Inteiro não negativo

Definições FieldType

Tabela 4: FieldType

Fieldtype	Definição
none	Usado para campos compostos, nesse caso o campo pai não teria valor e não exibiria controle de entrada de dados
date	Uma data, expressa em uma das três resoluções no formato AAAA-MM-DD, AAAA-MM ou AAAA
e-mail	Um endereço de e-mail válido. Não indexado por motivos de privacidade
text	Qualquer texto que não seja nova linha pode ser inserido neste campo
textbox	Qualquer texto pode ser inserido. Para entrada, o Dataverse apresenta uma área de multi-linhas que aceita novas linhas. Embora qualquer HTML seja permitido, apenas um subconjunto de tags HTML será renderizado na interface do usuário
url	Se não estiver vazio, o campo deve conter um URL válido
int	Um valor inteiro destinado a um campo numérico
float	Um número de ponto flutuante destinado a um campo numérico

Variáveis displayFormat

Essas são maneiras comuns de usar o displayFormat para controlar como os valores são exibidos na interface do usuário.

Tabela 5: displayFormat

Variável	Descrição
(blank)	O displayFormat é deixado em branco para campos primitivos e campos que não aceitam valores, pois displayFormats não funciona para esses campos
#VALUE	O valor do campo
#NAME	O valor do campo
#e-mail	Para exibir e-mails
#VALUE	Para exibir o valor como um link
#VALUE	Para exibir o valor como um link, com o valor incluído no URL
 	Para exibir a imagem de um URL de imagem inserido
#VALUE: - #VALUE: (#VALUE)	Anexa e / ou anexa caracteres ao valor do campo
; : ,	Exibe o caractere entre os valores dos campos nos campos compostos

6.4.3 Explorando blocos de metadados

Além de estudar os arquivos TSV, você pode encontrar os seguintes pontos da API altamente experimentais e sujeitos a alterações úteis para entender os blocos de metadados que já foram carregados em sua instalação do Dataverse:

Você pode obter um despejo de campos de metadados dessa forma:

```
curl http://localhost:8080/api/admin/datasetfield
```

Para ver detalhes sobre um campo individual como "title" siga o exemplo abaixo:

```
curl http://localhost:8080/api/admin/datasetfield/title
```

6.4.4 Criação de Metadados

Crie sua base de metadados, podendo ter como base os arquivos de metadados padrão. Os metadados estão localizados na pasta `/home/dataverse/dataverse-4.16/scripts/api/data/metablocks`. Depois de criado o arquivo de metadados, teste.tsv, transfira para a pasta

```
cp teste.csv /home/dataverse/dataverse-4.16/scripts/api/data/metablocks/
```

em seguida, recarregue os metadados.

```
curl http://localhost:8080/api/admin/datasetfield/load -H\  
  "Content-type: text/tab-separated-values\  
  -X POST --upload-file teste.tsv
```

6.4.5 Editando arquivos TSV

Observe que os campos de metadados compartilham um espaço para nome comum, portanto, eles devem ser exclusivos. O seguinte comando curl imprimirá a lista de campos de metadados já disponíveis no sistema.

```
curl http://localhost:8080/api/admin/index/solr/schema}
```

Usaremos este comando novamente abaixo para atualizar o esquema Solr para acomodar os campos de metadados que adicionamos.

6.4.6 Carregando arquivos TSV no Dataverse

Vários arquivos TSV são carregados no Dataverse a cada nova instalação, que se tornam os blocos de metadados que você vê na interface do usuário.

Junto com o arquivo TSV, há arquivos de propriedades ResourceBundle correspondentes com o par chave = valor.lang.directory e inclua um arquivo, por exemplo: "citation_lang.properties" no caminho especificado para a opção da JVM **`dataverse.lang.directory`** e, em seguida, reinicie o Glassfish.

Se você estiver alterando um bloco de metadados existente, o processo de instalação do Dataverse carregará o TSV para você, assumindo que você editou o arquivo TSV no local. O arquivo TSV para o bloco de metadados Citation, por exemplo, pode ser encontrado em **`scripts / api / data / metadatablocks / citation.tsv`**. Se algum dos valores de propriedade mencionados abaixo for alterado, o arquivo de propriedade ResourceBundle correspondente deverá ser editado e armazenado no local **`dataverse.lang.directory`**

- Nome, propriedade displayName em #metadataBlock
- Nome, título, descrição, propriedades da marca d'água em #datasetfield
- DatasetField, propriedade Value em #controlledVocabulary

Se você estiver criando um novo bloco de metadados customizados, o processo de instalação do Dataverse não saberá sobre o seu novo arquivo TSV, portanto, você deve carregar manualmente. O script que carrega os arquivos TSV no sistema é *scripts / api / setup-datasetfields.sh* e contém uma série de comandos curl. Aqui está um exemplo do comando curl necessário com o novo bloco de metadados personalizados no diretório `/ tmp`.

```
curl http://localhost:8080 / api / admin / datasetfield / load -H\  
  "Tipo de conteúdo: texto / valores separados por tabulação" -X\  
  POST --upload-file /tmp/new-metadata-block.tsv
```

Para criar um novo ResourceBundle, a seguir serão apresentadas as etapas para gerar o par chave = valor para as três seções principais:

6.4.7 Propriedades #metadataBlock

metadatablock.name = (o valor da propriedade **name** de #metadatablock)

metadatablock.displayName = (o valor da propriedade **displayName** de #metadatablock)

exemplo:

metadatablock.name = citação

metadatablock.displayName = Metadados da citação

6.4.8 Propriedades #datasetField (campo)

datasetfieldtype. (o valor da propriedade **name** de #datasetField) .title = (o valor da propriedade **title** de #datasetField)

datasetfieldtype. (o valor da propriedade **name** de #datasetField) .description = (o valor da propriedade **description** de #datasetField)

datasetfieldtype. (o valor da propriedade **name** de #datasetField) .watermark = (o valor da propriedade **watermark** de #datasetField)

exemplo:

datasetfieldtype.title.title = Título

datasetfieldtype.title.description = Título completo pelo qual o conjunto de dados é conhecido.

datasetfieldtype.title.watermark = Digite o título ...

6.4.9 Propriedades #controlledVocabulary (enumeradas)

Controladovocabulary. (o valor da propriedade **DatasetField** de #controlledVocabulary). (o valor da propriedade **Value** de #controlledVocabulary) = (o valor da proprie-

dade **Value** de #controlledVocabulary)

Como a propriedade **Value** de #controlledVocabulary é um texto livre, ao criar a chave, ela deve ser convertida para minúscula, substituir espaço por sublinhado e remover acentos.

exemplo:

```
controladovocabulary.subject.agricultural_sciences = Ciências Agrícolas
```

```
controlvocabulary.language.marathi_(marathi) = Marathi (Maru0101u1E6Dhu012B)
```

6.4.10 Habilitando um bloco de metadados

A execução de um comando curl como o exemplo "load" acima deve disponibilizar o novo bloco de metadados customizados no sistema, mas para começar a usar os campos, você deve habilitá-lo na GUI ou executando um comando curl como o abaixo, usando um token de API de superusuário. No exemplo abaixo, estamos habilitando os blocos de metadados “diário” e “geoespacial” para o dataverse raiz.

```
curl -H "X-Dataverse-key: \$ API\_TOKEN" -X POST -H\
"Tipo de conteúdo: application / json" -d\
["\" journal \", \" geospatial \"]"\
http: // localhost:8080/api/dataverses/:root/metadatablocks
```

6.4.11 Atualizando o esquema do Solr

Depois de ativar um novo bloco de metadados, você poderá ver os novos campos na GUI, mas antes de salvar o conjunto de dados, adicione campos adicionais ao seu esquema do Solr. Você deve executar o seguinte comando curl para que o Dataverse produza os elementos "nome do campo" e "copyField" para todos os campos de metadados que foram carregados no Dataverse:

```
curl http://localhost:8080/api/admin/index/solr/schema
```

Observe que, se você fizer uma solicitação pull atualizando */home/dataverse/dataverse-4.16/conf/solr/7.3.1/schema.xml* com os campos adicionados, primeiro carregue todos os blocos de metadados personalizados nos */home/dataverse/dataverse-4.16/scripts/api/data/metadatablocks* para criar uma lista completa de campos.

6.5 Recarregando um bloco de metadados

A sintaxe para recarregar é a mesma que para carregar. Aqui está um exemplo com o bloco de metadados "citation":


```
curl http:// localhost: 8080 / api / admin / datasetfield / load -H\  
"Tipo de conteúdo: texto / valores separados por tabulação" -X\  
POST --upload-file citation.tsv
```

Muito cuidado deve ser tomado ao recarregar um bloco de metadados. A correspondência é feita em nomes de campos (ou identificadores e, em seguida, em nomes de valores de vocabulário controlados), então é fácil criar acidentalmente campos duplicados.

A capacidade de recarregar blocos de metadados significa que os scripts de atualização SQL não precisam ser gravados para essas alterações.

6.6 Exportação de Metadados

6.6.1 Exportações automáticas

A publicação de um conjunto de dados inicia automaticamente um trabalho de exportação de metadados, que será executado em segundo plano, de forma assíncrona. Depois de concluído, ele fará com que os metadados do conjunto de dados sejam exportados e armazenados em cache em todos os formatos suportados listados em Formatos de exportação de metadados suportados na seção Conjunto de dados + Gerenciamento de arquivos do Guia do Usuário.

Um trabalho de timer agendado que será executado todas as noites tentará exportar quaisquer conjuntos de dados publicados que, por qualquer motivo, ainda não tenham sido exportados. Esse cronômetro é ativado automaticamente na implantação ou reinicialização do aplicativo. Portanto, não há necessidade de iniciar ou configurar manualmente.

6.6.2 Exportações em lote através da API

Além das exportações automatizadas, um administrador do Dataverse pode iniciar um trabalho em lotes por meio da API. As 2 chamadas de API a seguir são fornecidas:

```
/api/admin/metadados/exportAll
```

```
/api/admin/metadata/reExportAll
```

O primeiro tentará exportar todos os conjuntos de dados locais publicados (não coletados) que ainda não foram exportados. O último forçará a reexportação de todos os conjuntos de dados locais publicados, independentemente de já terem sido exportados ou não.

Observe que a criação, modificação ou reexportação de um conjunto OAI também tentará exportar todos os conjuntos de dados não exportados encontrados no conjunto.

6.6.3 Falhas na Exportação

Um trabalho em lote de exportação, iniciado pela API ou pelo cronômetro do aplicativo, deixará um log detalhado no diretório de logs configurado. Este é o mesmo local em que seu server.log principal do Glassfish é encontrado. O nome do arquivo de log é:

```
export_ [timestamp] .log
```

- por exemplo,

```
export_2016-08-23T03-35-23.log
```

O log conterá o número de conjuntos de dados processados com êxito e aqueles para os quais a exportação de metadados falhou, com algumas informações sobre as falhas detectadas.

6.7 Aplicação de Temporizadores do Dataverse

O Dataverse usa temporizadores para executar automaticamente tarefas de exportação agendadas de coleta e metadados.

6.7.1 Servidor de timer dedicado em um cluster de servidor do Dataverse

Ao executar um cluster do Dataverse, ou seja, vários servidores de aplicativos do Dataverse que conversam com o mesmo banco de dados, apenas um deles deve atuar como o servidor de timer dedicado. Isso evita iniciar tarefas em lote conflitantes em vários nós ao mesmo tempo. Isso não afeta uma instalação de servidor único.

A seguinte opção da JVM instrui o aplicativo a agir como o servidor de timer dedicado:

```
-Ddataverse.timerServer = true
```

Observe que esta opção é configurada automaticamente pelo script do instalador do Dataverse. Isso significa que, ao configurar um cluster com vários servidores, será responsabilidade do instalador remover a opção do `domain.xml` de cada nó, exceto aquele destinado ao servidor de timer. Também recomendamos que a seguinte entrada no `domain.xml`:

```
<ejb-timer-service timer-datasource = "jdbc / VDCNetDS">
```

Seja alterada novamente em todos os nós do servidor que não sejam do timer para:

```
<ejb-timer-service>
```

Da mesma forma, essa opção é configurada automaticamente pelo script do instalador. A alteração para a configuração padrão em um servidor que não precisa executar o cronômetro evitará uma condição potencial de concorrência, na qual vários servidores tentam bloquear o banco de dados do cronômetro.

Para que o timer funcione, a versão do driver JDBC do PostgreSQL que sua instância está usando deve corresponder à versão do banco de dados do PostgreSQL.

6.7.2 Temporizadores de coleta

Esses cronômetros são criados quando a coleta programada é ativada por um usuário administrador local (na página “Gerenciar clientes da coleta”).

Em um cluster com vários nós, todos esses cronômetros serão criados no node do cronômetro dedicado (e não necessariamente no nó em que os clientes de coleta foram criados e / ou salvos).

Um cronômetro será removido automaticamente quando um cliente de coleta com um agendamento ativo for excluído ou se o agendamento estiver desativado para um cliente existente.

6.7.3 Temporizador de Exportação de Metadados

Esse cronômetro é criado automaticamente sempre que o aplicativo é implantado ou reiniciado. Não há configuração acessível pelo usuário administrador para este timer.

Esse timer executa um trabalho diário que tenta exportar todos os conjuntos de dados publicados locais que ainda não foram exportados, em todos os formatos de metadados suportados, e armazena em cache os resultados no sistema de arquivos. Normalmente uma exportação ocorrerá automaticamente sempre que um conjunto de dados for publicado. Este trabalho agendado existe para capturar quaisquer conjuntos de dados para os quais essa exportação não teve êxito, por um motivo ou outro. Na primeira vez em que esse trabalho for executado, ele tentará exportar todos eles.

Esse trabalho diário também atualizará todos os conjuntos OAI coletáveis configuradas em seu servidor, adicionando conjuntos de dados novos e / ou recém-publicados ou marcando os conjuntos de dados desativados como "excluídos" nos conjuntos correspondentes, conforme necessário.

Este trabalho é agendado automaticamente para ser executado às 02:00, horário local, todas as noites. Caso seja necessário, é possível (para um usuário avançado) alterar esse horário editando diretamente a tabela do aplicativo de timer EJB no banco de dados.

6.8 Make Data Count

Make Data Count é um projeto para coletar e padronizar métricas sobre o uso de dados, especialmente visualizações, downloads e citações. O Dataverse pode integrar o Make Data Count para coletar e exibir métricas de uso, incluindo contagens de visualizações de conjuntos de dados, downloads de arquivos e citações de conjuntos de dados.

6.8.1 Introdução

A Make Data Count faz parte de um grupo de trabalho de métricas de uso de dados da Research Data Alliance (RDA), que ajudou a produzir uma especificação chamada Código de Prática para Dados de Pesquisa COUNTER (PDF, HTML), que o Dataverse

faz todos os esforços para cumprir. O Código de Prática (CoP) é construído sobre os padrões existentes, como COUNTER e SUSHI, que saem do mundo da publicação de artigos. O projeto Make Data Count enfatizou que eles gostariam de receber feedback sobre o código de prática.

6.8.2 Arquitetura

Instalações de dataverse que desejam dar suporte ao Make Data Count devem instalar o Counter Processor, um projeto Python criado pela California Digital Library (CDL) que faz parte do projeto Make Data Count e que executa o software em produção como parte de sua plataforma de compartilhamento de dados DASH .

6.8.3 Limitações para instalações de dataverse usando handles em vez de DOIs

Os repositórios de dados que usam Handles e outros identificadores não são suportados pelo Make Data Count. Dito isto, o log de uso do Dataverse ainda pode gerar logs e processá-los com o Counter Processor para criar json que detalha o uso no nível do conjunto de dados. O Dataverse pode ingerir esse json gerado localmente. Ao editar o arquivo mencionado abaixo

```
counter-processor-config.yaml
```

verifique se o booleano abaixo está definido como False.

```
upload_to_hub
```

6.8.4 Habilitar o log para o Make data count

Para fazer uso do conjunto de dados de log do Dataverse (visualizações e downloads) para Make Data Count, você deve definir a configuração do banco de dados

```
: MDCLogPath
```

Depois de ter seu primeiro dia de logs, você poderá processá-los no dia seguinte.

6.8.5 Configurar o Contador do Processador

- Primeiro, torne-se o usuário Unix “contador”.

```
sudo su - contador
```

- Mude para o diretório em que você instalou o Counter Processor.

```
cd /usr/local/counter-processor-0.0.1
```

- Faça o download do:

```
counter-processor-config.yaml
```

para

```
/usr/local/counter-processor-0.0.1.
```

- Edite o arquivo de configuração e preste atenção especial às linhas FIXME.

```
vim counter-processor-config.yaml
```

6.8.6 Popular visualizações e downloads pela primeira vez

Começaremos com uma única configuração bem-sucedida de execução do Counter Processor e chamadas para APIs do Dataverse.

- Mude para o diretório em que você instalou o Counter Processor.

```
cd /usr/local/counter-processor-0.0.1
```

- Se você estiver executando o Counter Processor pela primeira vez no meio de um mês, precisará criar arquivos de log em branco para os dias anteriores por exemplo:

```
cd /usr/local/glassfish4/glassfish/domínios/domain1/logs
```

```
touch counter_2019-02-01.log
```

```
...
```

```
touch counter_2019-02-20.log
```

- Execute o contador do processador.

```
CONFIG_FILE = contador-processor-config.yaml python36 main.py
```

Um arquivo JSON no formato SUSHI será criado no diretório que você especificou em “output_file” no arquivo de configuração.

- Popule visualizações e downloads para seus conjuntos de dados com base no arquivo SUSHI JSON. O diretório “/tmp” é usado no exemplo abaixo.

```
curl -X POST "http://localhost:8080/api/admin/makeDataCount/\naddUsageMetricsFromSushiReport? reportOnDisk = /tmp/\nmake-data-count-report.json"
```

- Verifique se as visualizações e downloads estão disponíveis via API. Agora que visualizações e downloads foram gravados no banco de dados Dataverse, você deve recuperá-los de um ou mais conjuntos de dados.

6.8.7 Popular visualizações e downloads noturnos

A execução de main.py para criar o arquivo SUSHI JSON e a chamada subsequente da API do Dataverse para processá-lo devem ser adicionadas como uma tarefa cron.

6.8.8 Enviando métricas de uso para o DataCite Hub

Altere "upload_to_hub" para "True" no arquivo de configuração. Na próxima vez que você executar o main.py, as seguintes métricas serão enviadas ao hub do DataCite para cada conjunto de dados publicado:

- Visualizações ("investigações" em COUNTER)
- Downloads ("solicitações" em COUNTER)

6.8.9 Configurando o Dataverse para fazer citações do Make Data Count

Observe que no exemplo curl, as variáveis de ambiente Bash são usadas com a ideia de que você pode definir algumas variáveis de ambiente e copiar e colar os exemplos como estão. Por exemplo, "\$ DOI" pode se tornar "doi: 10.5072 / FK2 / BL2IBM" emitindo o seguinte comando de exportação do Bash:

```
export DOI= "doi:10.5072/FK2/BL2IBM"
```

Para confirmar se a variável de ambiente foi definida corretamente, você pode usar o echo assim:

```
echo $ DOI
```

Em alguma base periódica (talvez semanalmente), você deve chamar o seguinte comando curl para cada conjunto de dados publicado para atualizar a lista de citações que foram feitas para esse conjunto de dados.

```
curl -X POST "http://localhost:8080/api/admin/makeDataCount/\n:persistentId/updateCitationsForDataset? persistentId = $ DOI"
```

As citações serão recuperadas para cada conjunto de dados publicado e registradas no banco de dados Dataverse. Observe que, embora o Dataverse tenha um campo de metadados para "Conjunto de dados relacionados", essas informações não são enviadas atualmente como uma citação para Crossref.

6.8.10 Recuperando métricas do Make Data Count do hub DataCite

As seguintes métricas podem ser baixadas diretamente do hub DataCite para conjuntos de dados hospedados por instalações do Dataverse que foram configuradas para enviar essas métricas para o hub:

- Total de visualizações para um conjunto de dados;

- Visualizações exclusivas para um conjunto de dados;
- Total de downloads para um conjunto de dados;
- Downloads para um conjunto de dados;
- Citações de um conjunto de dados (via Crossref).

6.8.11 Recuperando métricas do MAke Data Counts de dados do Dataverse

Os terminais da API do Dataverse para recuperar as métricas Make Data Count estão descritos abaixo em Métricas do conjunto de dados na seção Native APIs do API Guide. Observe que também é possível recuperar métricas do próprio hub DataCite via <https://api.datacite.org>

6.9 Integrações

6.9.1 Obtendo dados

Uma variedade de integrações é existente para tornar mais fácil para seus pesquisadores depositar dados em sua instalação do Dataverse, incluindo:Dropbox, Estrutura de Ciência Aberta (OSF), RSpace e Sistemas de periódicos abertos (OJS). Acerca da Preservação de Dados de Pesquisa encontramos o Archivematica e DuraCloud / Chronopolis.

6.10 Administração do usuário

6.10.1 Gerenciar usuários

A tabela Gerenciar usuários fornece ao administrador da rede uma lista de todas as contas de usuário no formato de tabela. Você pode acessá-lo clicando no botão "Gerenciar usuários" no painel, que está vinculado ao cabeçalho de todas as páginas do Dataverse (se você estiver conectado como administrador). Ele lista o nome de usuário, o nome completo, o endereço de e-mail, a afiliação, o método de autenticação que eles usam, as funções que sua conta recebeu e se eles têm ou não status de superusuário.

Os usuários são listados em ordem alfabética por nome de usuário. A barra de pesquisa acima da tabela permite pesquisar um usuário específico. Ele executa uma pesquisa coringa à direita truncada das colunas Nome de usuário, Nome e e-mail. Isso significa que, se você pesquisar "futeba", ele procurará nessas três colunas qualquer sequência de texto que comece com "futeba", por exemplo "Futebol" ou "Futebolbola".

Se você deseja atribuir ou remover o status de superusuário de um usuário, marque ou desmarque a caixa de seleção na coluna "Superusuário".

Se você deseja remover todas as funções / permissões da conta de um usuário (no caso de deixarem a organização, por exemplo), clique no botão "Remover tudo", na coluna Funções. Isso manterá a conta do usuário ativa, mas a reverterá para colocar a conta no nível de um usuário padrão com permissões padrão.

6.10.2 Listar usuários via API

Existem duas maneiras de listar usuários via API. Se você tiver relativamente poucos usuários, poderá obtê-los todos como um dump com este comando com um token de API de superusuário:

```
curl -H "Chave X do Dataverse: $ API_TOKEN" http://localhost:8080/api/admin/authenticatedUsers
```

Se você possui muitos usuários e deseja pesquisar e paginar pelos resultados, use o comando abaixo com um token de API de superusuário:

```
curl -H "Chave X do Dataverse: $ API_TOKEN" http://localhost:8080/api/admin/list-users
```

Com o formulário list-users, você pode incluir os seguintes parâmetros de consulta opcionais:

- searchTerm: Uma sequência que corresponde ao início de um identificador de usuário, nome, sobrenome ou endereço de e-mail;
- itemsPerPage: O número de resultados detalhados a serem retornados. O padrão é 25. Este número não tem limite. por exemplo. Você pode configurá-lo para 1000 para retornar 1.000 resultados;
- selectedPage: A página de resultados a retornar. O padrão é 1.

6.10.3 Confirmar e-mail

O Dataverse incentiva os usuários internos / locais a verificar seu endereço de e-mail ao se inscrever ou alterar o e-mail, para que os administradores de sistemas possam ter certeza de que os usuários podem ser contatados.

O aplicativo envia um e-mail de boas-vindas padrão com um URL no qual o usuário pode clicar e, quando ativado, armazena um lastconfirmed carimbo de hora confirmado na tabela authenticateduser do banco de dados. Sempre que isso é "nulo" para um usuário (imediatamente após a inscrição e / ou alteração do endereço de e-mail do Dataverse), o e-mail atual em arquivo é considerado não verificado. O link enviado expira após um tempo (o padrão é 24 horas), mas isso é configurável por um superusuário por meio da opção:

MinutesUntilConfirme-mailTokenExpiresconfig.

Se o token do URL dos usuários expirar, eles verão o botão "Verificar e-mail" na página de informações da conta para enviar outro URL.

Os administradores de sistemas podem determinar quais usuários verificaram seus endereços de e-mail procurando a presença do valor e-mailLastConfirmed na saída JSON ao listar usuários. Os endereços de e-mail dos usuários do Shibboleth são confirmados a cada login.

6.11 Excluindo um token de API

Se um token de API estiver comprometido, ele deverá ser excluído. Os usuários podem gerar um novo para si mesmos, mas você pode excluir preventivamente os tokens do banco de dados. Usando o token da API 7ae33670-be21-491d-a244-008149856437 como um exemplo:

```
delete from apitoken where tokenstring =  
'7ae33670-be21-491d-a244-008149856437';
```

Você deve esperar a saída DELETE 1 após emitir o comando acima.

6.12 Gerenciando conjuntos de dados e Dataverses

6.12.1 Dataverses

Excluir um Dataverse

Dataverses precisam estar vazios para excluí-los. Navegue até o coleta e clique em "Editar" e, em seguida, em "Excluir dataverse" para excluí-lo.

Mover um Dataverse

Move um dataverse cujo ID é passado para um novo dataverse cujo ID também é movido conjuntamente. O alias do dataverse também pode ser usado em vez do id. Se o coleta movido tiver um livro de visitas, modelo, bloco de metadados, link ou coleta em destaque que não seja compatível com o dataverse de destino, você será informado e terá a opção de forçar a movimentação e remover a associação. Acessível apenas para superusuários.

```
curl -H "X-Dataverse-key: $ API_TOKEN" -X POST http: // $ SERVER /  
api / dataverses / $ id / move / $ destination-id
```

Vincular um Dataverse

Cria um link entre um dataverse e outro dataverse. Acessível apenas para superusuários.

```
curl -H "X-Dataverse-key: $ API_TOKEN" -X PUT http: // $ SERVER /  
api / dataverses / $ alias-  
dataverse vinculado / link / $ alias-dataverse-alias
```

Desvincular um Dataverse

Remove um link entre um dataverse e outro dataverse. Acessível apenas para superusuários.

```
curl -H "X-Dataverse-key: $ API_TOKEN" -X DELETE http: // $ SERVER /  
api / dataverses / $ linked-dataverse-alias / deleteLink / $ binding-  
dataverse-alias
```

Adicionar atribuições de função de dataverse a dados de dataverses filho

Atribui recursivamente os usuários e grupos que possuem funções definidas no conjunto para serem herdadas por meio da configuração: InheritParentRoleAssignments, em um coleta especificado para ter as mesmas atribuições de função em todos os dataverses criados nele. A resposta indica sucesso ou falha e lista os indivíduos / grupos e os dados envolvidos na atualização. Acessível apenas para superusuários.

```
curl -H "X-Dataverse-key: $ API_TOKEN" http: // $ SERVER /  
api / admin / dataverse / $ dataverse-alias // addRoleAssignmentsTo  
Children
```

6.12.2 Conjuntos de dados

Mover um conjunto de dados

Os superusuários podem mover conjuntos de dados usando o painel. Se o conjunto de dados movido tiver um livro de visitas ou um link dataverse que não seja compatível com o dataverse de destino, você será informado e terá a opção de forçar a movimentação (com forceMove = true como parâmetro de consulta) e remover o livro de visitas ou o link (ou ambos). Acessível apenas aos usuários com permissão para publicar o conjunto de dados no coleta original e de destino.

```
curl -H "X-Dataverse-key: $ API_TOKEN" -X POST http: // $ SERVER  
/ api / datasets / $ id / move / $ alias
```

Vincular um conjunto de dados

Cria um link entre um conjunto de dados e um dataverse.

```
curl -H "X-Dataverse-key: $ API_TOKEN" -X PUT http: // $ SERVER / api /  
datasets / $ ID do conjunto de dados vinculado / link / $ alias-  
dataverse-  
alias
```

Desvincular um conjunto de dados

Remove um link entre um conjunto de dados e um dataverse. Acessível apenas para superusuários.

```
curl -H "X-Dataverse-key: $ API_TOKEN" -X DELETE http: // $ SERVER / api  
/ datasets / $ linked-dataset-id / deleteLink / $ binding-dataverse-alias
```

Criar novo PID para um conjunto de dados

Minta um novo identificador para um conjunto de dados registrado anteriormente com um identificador. Acessível apenas para superusuários.

```
curl -H "X-Dataverse-key: $ API_TOKEN" -X POST http: // $ SERVER / api /  
admin / $ dataset-id / reregisterHDLToPID
```

Enviar metadados do conjunto de dados ao provedor PID

Força a atualização dos metadados fornecidos ao provedor PID de um conjunto de dados publicado. Acessível apenas para superusuários.

```
curl -H "X-Dataverse-key: $ API_TOKEN" -X POST http: // $ SERVER / api /  
datasets / $ dataset-id / modifyRegistrationMetadata
```

Faça atualizações de metadados sem alterar a versão do conjunto de dados

Como superusuário, clique em "Atualizar versão atual" ao publicar.

6.13 Índice de Pesquisa Solr

O Dataverse exige que o Solr esteja operacional o tempo todo. Se você parar o Solr, verá um erro sobre isso na página inicial, que é alimentada pelo índice de pesquisa que o Solr fornece. O PostgreSQL é a fonte de informações e o aplicativo Dataverse copiará os dados do PostgreSQL para o Solr. Por esse motivo, o índice de pesquisa pode ser reconstruído a qualquer momento. Dependendo da quantidade de dados que você possui, esse pode ser um processo lento.

6.13.1 Reindexar completo

Há duas maneiras de executar uma reindexação completa do índice de pesquisa do Dataverse. O primeiro garante um índice completamente limpo, mas envolve tempo de inatividade. A reindexação no local não envolve tempo de inatividade, mas não garante um índice completamente limpo.

Limpar e Reindexar

Limpando dados do Solr

Observe que, no momento em que você emitir esse comando, os usuários finais que olharem para a página inicial parecerão que todos os dados se foram! Isso ocorre porque a página inicial é alimentada pelo índice de pesquisa.

```
curl http: // localhost: 8080 / api / admin / index / clear
```

Iniciar reinicialização assíncrona

Observe que esta operação pode levar horas, dependendo da quantidade de dados em seu sistema.

```
curl http: // localhost: 8080 / api / admin / index
```

Reindexar no local

Uma alternativa para limpar completamente o índice de pesquisa é reindexar no local.

Limpar timestamps de índice

```
curl -X DELETE http: // localhost: 8080 / api / admin / index /  
timestamps
```

Iniciar ou continuar reinicialização assíncrona

Se a indexação parar, esse comando deve continuar de onde parou, com base em quais registros de data e hora do índice foram definidos, e é por isso que começamos limpando esses registros de data e hora acima. Esses carimbos de hora são armazenados na tabela de banco de dados dvobject.

```
curl http: // localhost: 8080 / api / admin / index / continue
```

6.13.2 Reindexação manual

Se você fez alterações manuais em um conjunto de dados no banco de dados ou deseja reindexar um conjunto de dados que o solr não desejava indexar corretamente, é possível reindexar manualmente conjuntos de dados e dados.

Reindexando Dataverses

Dataverses devem ser referenciados pelo ID do objeto de banco de dados. Se você tiver acesso direto ao banco de dados, uma consulta SQL como

```
selecione id do dataverse em que alias = 'datavarsealias';
```

ou você pode clicar no menu "Editar" do Dataverse e procurar dataverseId = nos URLs produzidos pelo menu suspenso. Em seguida, para indexar novamente:

```
curl http: // localhost: 8080 / api / admin / index / dataverses / 135  
que deve retornar: _ {"status": "OK", "data": {"message": "reindexando  
o coleta 135"}} _
```

Reindexando conjuntos de dados

Os conjuntos de dados podem ser referenciados pelo ID persistente ou pelo ID do objeto de banco de dados. Para indexar novamente por ID persistente:

```
curl http: // localhost: 8080 / api / admin / index / dataset?  
persistentId = doi: 10.5072 / FK2 / AAA000
```

Para indexar novamente um conjunto de dados por seu ID do banco de dados:

```
curl http: // localhost: 8080 / api / admin / index / datasets / 7504557
```

6.13.3 Consultando manualmente o Solr

Se você suspeitar que algo não está indexado corretamente no solr, poderá ignorar a interface da Web do Dataverse e consultar a linha de comando abaixo diretamente para verificar o que o solr retorna vendo o JSON esperado ser passado para o Dataverse.

```
curl "http: // localhost: 8983 / solr / collection1 / select?  
q = dsPersistentId:doi: 10.15139 / S3 / HFV0A0"
```

6.14 Grupos de IP

Grupos de IP podem ser usados para permitir o download de arquivos restritos por endereços IP, e não por pessoas. Por exemplo, convém permitir que arquivos restritos sejam baixados por pesquisadores que entram fisicamente em uma biblioteca e fazem uso da rede da biblioteca.

6.14.1 Listando grupos de IP

Os grupos de IPs podem ser listados com o seguinte comando:

```
curl: http: // localhost: 8080 / api / admin / groups / ip
```

6.14.2 Criando um grupo de IPs

Os grupos de IPs devem ser expressos como intervalos no formato IPv4 ou IPv6. Veja um exemplo de todo o intervalo IPv4 e IPv6 que você pode baixar e editar para ter um intervalo mais restrito para atender às suas necessidades. Se você precisar que seu grupo de IP abranja apenas um único endereço IP, insira esse endereço IP para o "início" e o "fim" do intervalo. Se você não usar endereços IPv6, poderá excluir essa seção do JSON. Observe que o "alias" deve ser exclusivo se você definir vários grupos de IPs. Você deve atribuir um "nome" significativo, pois "alias" e "nome" aparecerão e poderão ser pesquisados na GUI quando seus usuários estiverem atribuindo funções.

```
{
  "alias": "ipGroupAll",
  "name": "Grupo de IPs para corresponder a todos os endereços IPv4 e IPv6",
  "gamas": [
    [
      "0.0.0.0",
      "255.255.255.255"
    ],
    [
      "::",
      "ffff:ffff:ffff:ffff:ffff:ffff:ffff:ffff"
    ]
  ]
}
```

Digamos que você faça o download do exemplo acima e edite-o para obter um intervalo usado por sua biblioteca, fornecendo um nome de arquivo ipGroup1.json e colocando-o no diretório / tmp. Em seguida, carregue-o no Dataverse usando o seguinte comando curl:

```
curl -X POST -H 'Tipo de conteúdo: application / json' http: //
localhost: 8080 / api / admin / groups / ip --upload-file /
tmp / ipGroup1.json
```

Observe que você pode atualizar um grupo da mesma maneira, desde que use o mesmo alias.

6.14.3 Listando um grupo de IPs

Digamos que você usou "ipGroup1" como o alias do grupo de IPs criado acima. Para listar apenas esse grupo de IPs, você pode incluir o alias no comando curl assim:

```
curl http: // localhost: 8080 / api / admin / groups / ip / ipGroup1
```

6.14.4 Excluindo um grupo de IPs

Não é recomendável excluir um grupo de IPs aos quais foram atribuídas funções. Se você deseja excluir um grupo de IPs, remova primeiro suas permissões. Para excluir um grupo de IPs com um alias de "ipGroup1", use o comando curl abaixo:

```
curl -X DELETE http: // localhost: 8080 / api / admin / groups / ip
/ ipGroup1
```

6.15 Tráfego HTTP

O tráfego HTTP pode ser monitorado do lado do cliente, do servidor ou de ambos.

6.15.1 Monitorando o tráfego HTTP do lado do cliente

O tráfego HTTP para clientes da Web que possuem cookies ativados (a maioria dos navegadores) pode ser rastreado pelo Google Analytics (<https://www.google.com/analytics/>). Para rastrear análises além das visualizações de página, foram adicionadas classes de estilo aos botões de ação do usuário final, que incluem:

```
btn-compute, btn-contact, btn-download, btn-explore, btn-export,  
btn-preview, btn-request, btn-share
```

6.15.2 Temporizadores EJB

Se você estiver interessado em monitorar os cronômetros EJB, este script pode ser usado como um exemplo:

```
#!/usr/bin/env bash

# exemplo script de monitoramento para timers EJB.
# atualmente assume que existem dois temporizadores
# comandos de monitoramento reais devem substituir as instruções echo
para uso em produção

r0 = `curl -s http://localhost:8080/ejb-timer-service-app/timer`

se [$? -ne 0]; então
echo "alert - no timer service" # coloque aqui o comando de alerta real
fi

r1 = `echo $ r0 | grep -c "Existem 2 temporizadores persistentes ativos
neste contêiner" `

if ["1" -ne "$ r1"]; então
echo "alert - no timers active" # coloque o comando de alerta real aqui
fi
```

6.16 Solução de problemas

6.16.1 Glass fish

server.log é o principal local para procurar quando você encontrar problemas. Provavelmente uma mensagem de erro tenha sido registrada. Se houver um rastreamento de stack, pode ser do interesse dos desenvolvedores, especialmente eles podem rastrear os números de linha de volta para uma versão marcada. Para fins de debugging, você pode achar útil aumentar os níveis de log.

6.16.2 Falha na implementação: Serviço de Timer EJB Não Disponível

Às vezes, o aplicativo Dataverse falha ao implantar ou o Glassfish falha ao reiniciar quando o aplicativo é implantado, com a seguinte mensagem de erro:

“falha remota: ocorreu um erro durante a implantação: exceção ao carregar o aplicativo: o Serviço de Timer EJB não está disponível. ”

Aqui está uma solução alternativa conhecida:

- Pare o Glassfish;
- Remova os diretórios gerados e osgi-cache;
- Exclua todas as linhas da tabela EJB__TIMER__TBL no banco de dados;
- Iniciar Glassfish.

O script de shell abaixo executa as etapas acima. Além dos valores de configuração que precisam ser alterados para refletir seu ambiente (o diretório Glassfish, nome do banco de dados, etc.), o script depende do banco de dados ser configurado de uma certa maneira para acesso.

```
#!/ bin / sh
```

```
# Os temporizadores EBJ às vezes causam problemas; utilitário para limpar  
diretórios gerados e linhas de banco de dados
```

```
# assume que este script está sendo executado como root e que o usuário do  
postgres não possui senhas
```

```
# acesso ao banco de dados (soquetes locais ou variáveis ambientais apropriadas).
```

```
# reiniciará o glassfish se estiver parado; comente o comando `start-domain`  
no final
```

```
# se você quiser evitar isso.
```

```
# diretório em que o glassfish está instalado
```

```
GLASSFISH_DIR = / usr / local / glassfish4
```

```
# diretório dentro do glassfish (padrões)
```

```
DV_DIR = $ {GLASSFISH_DIR} / glassfish / domínios / domínio1
```

```
# nome do banco de dados dataverse
```

```
DV_DB = dvndb
```

```
# Usuário do SO para o banco de dados
```



```
DB_USER = postgres
```

```
# interrompe o domínio glassfish (gera um aviso se o glassfish for parado)  
$ {GLASSFISH_DIR} / bin / asadmin stop-domain
```

```
rm -rf $ {GLASSFISH_DIR} / $ {DV_DIR} / generated /  
rm -rf $ {GLASSFISH_DIR} / $ {DV_DIR} / osgi-cache / felix
```

```
sudo -u $ {DB_USER} psql $ {DV_DB} -c 'delete de "EJB__TIMER__TBL";
```

```
# reinicie o domínio (também gera um aviso se o glassfish for parado)  
$ {GLASSFISH_DIR} / bin / asadmin start-domain
```

6.16.3 Temporizador não funciona

O Dataverse depende dos cronômetros EJB para executar tarefas agendadas: coleta de servidores remotos, atualização dos conjuntos OAI locais e execução de exportações de metadados. Se esses trabalhos agendados não estiverem em execução no servidor, isso pode ser o resultado da incompatibilidade entre a versão do banco de dados PostgreSQL que você está usando e o driver JDBC PostgreSQL em uso pela sua instância do Glassfish. Os sintomas:

Se você estiver vendo o seguinte em seu server.log ...

Handling timeout on ...

seguido por um rastreamento de pilha de exceção com estas linhas:

Internal Exception: java.io.StreamCorruptedException: invalid stream header ...

Exception Description: Could not deserialize object from byte array ...

Provavelmente significa que existe incompatibilidade do driver JDBC que impede o funcionamento correto do timer. Certifique-se de instalar a versão correta do driver. Por exemplo, se você estiver executando a versão 9.3 do PostgreSQL, verifique se possui o driver postgresql-9.3-1104.jdbc4.jar no diretório <GLASSFISH FOLDER> / glassfish / lib. Se você possui um driver mais antigo no glassfish / lib, remova-o, substitua-o pela nova versão e reinicie o Glassfish. (Pode ser necessário remover todo o conteúdo de <GLASSFISH FOLDER> / glassfish / domínios / domínio1 / gerado antes de iniciar o Glassfish).

Manual do usuário:

Apêndice

7.1 Metadados

Dataverse está comprometido com o uso de metadados compatíveis com o padrão para garantir que os metadados do Dataverse possam ser mapeados facilmente para esquemas de metadados padrão e serem exportados para o formato JSON (XML para metadados de arquivo tabular) para preservação e interoperabilidade.

A seguir, são detalhados os esquemas de metadados que oferecemos suporte para citação e metadados específicos de domínio no Dataverse:

- Metadados de citação: compatível com DDI Lite, DDI 2.5 Codebook, DataCite 3.1 e Termos de metadados DCMI do Dublin Core (consulte a versão .tsv). O campo de idioma usa o vocabulário controlado pela ISO 639-1.
- Geoespaciais: compatível com DDI Lite, DDI 2.5 Codebook, DataCite e Dublin Core (consulte a versão .tsv). O campo País / Nação usa vocabulário controlado pela ISO 3166-1.
- Ciências Sociais e Humanas: compatível com DDI Lite, DDI 2.5 Codebook e Dublin Core (consulte a versão .tsv).
- Astronomia e astrofísica: esses elementos de metadados podem ser mapeados / exportados para o formato de esquema VOResource da International Virtual Observatory Alliance (IVOA) e são baseados nos metadados de descoberta e proveniência do Virtual Observatory (VO) (consulte versão .tsv).

- Ciências da vida: com base na especificação ISA-Tab, juntamente com o vocabulário controlado de subconjuntos da ontologia do OBI e da taxonomia para organismos do NCBI.

7.2 OpenAIRE

O projeto Digital Repository Infrastructure Vision for European Research (DRIVER) em 2007 publicou as Guidelines for content providers: Exposing textual resources with OAI-PMH, que contém recomendações que permitem interoperabilidade entre repositórios. Desde então essas recomendações foram evoluindo e estão atualmente em sua versão 4.0 sendo chamadas de OpenAIRE Guidelines for Literature Repository Managers v4. A partir de 2013 a Comissão Europeia com a publicação do Guidelines on Open Access to Scientific Publications and Research Data está de acordo com os requerimentos do Horizon 2020 abrangendo desde repositórios de literatura à repositórios de dados e sistemas CRIS.

As diferentes versões das diretrizes estão apresentadas em [link](#) e versão mais recente em [link](#). A diretriz é dividida em três partes: Uma introdução, onde ela é apresentada; Uso do OAI-PMH, na qual são apresentadas algumas explicações sobre o uso do protocolo OAI-PMH e; por fim uma visão de um perfil de aplicação, que pode ser encontrado em [link](#). A aplicação é baseada em quatro padrões de metadados: Dublin Core; Dublin Core qualificado; Datacite e; Oaire (um padrão de metadados próprio desenvolvido pela OpenAIRE), além de serem utilizados diversos vocabulários controlados.

Aplicação

Foram definidos quatro níveis de obrigatoriedade para metadados: Mandatório (M); Mandatório se aplicável (MA); Recomendado (R) e; Opcional (O). Seguem as explicações de cada um deles.

- Mandatório (M): Qualquer campo com esse nível de obrigatoriedade deve ser preenchido independentemente da situação, caso isso não ocorra estará em não conformidade com as diretrizes. (ex: Título)
- Mandatório se aplicável (MA): Quando o campo apresenta esse nível de obrigatoriedade ele deve ser preenchido caso a informação em questão esteja disponível. (ex: Contribuidor)
- Recomendado (R): O campo que apresentar esse nível de obrigatoriedade não necessita ser preenchido porém é benéfico que o faça. (ex: Identificador Alternativo)
- Opcional (O): Este nível de obrigatoriedade significa que seu preenchimento é de caráter completamente discricionário, caso se julgue que seu preenchimento enriquecerá o registro pode-se preenchê-lo ou não. (ex: Tamanho)

É apresentado no perfil de aplicação uma tabela com os campos, elementos e vocabulários utilizados. São 32 campos no total, com seis mandatórios, oito mandatórios se aplicáveis, quinze recomendados e três opcionais. Para melhor visualização essa tabela foi reformulada e dividida de acordo com os níveis de obrigatoriedade a seguir. Os primeiros campos apresentados serão os mandatórios (M):

Tabela 6: Metadados mandatórios

Campo	Metadado	Definição	Exemplo	Exemplo de aplicação	Vocabulário controlado
Title (Título)	datacite: title	Nome ou título pelo qual o recurso é conhecido	National Institute	<datacite:title xml:lang="en-US"> National Institute </datacite:title>	title type
Creator (Criador)	datacite: creator	Autor do recurso, poder ser uma instituição ou nome pessoal	Evans, R.J.	<datacite: creator>< datacite:creato rName>Evans, R.J.</datacite :creatorName> </datacite: creator>	name type
Publication Date (Data de Publicação)	datacite: date	Data com alguma associação com o ciclo de vida do recurso, deve ser utilizado o formato W3C	2000-12-25	<datacite:date dateType="Issued">2000-12-25</datacite:date>	date type
Resource Type (Tipo de Recurso)	aire: resource Type	O gênero do recurso, tipo de produto científico	Journal Article (Artigo de periódico)	<aire:resourc eType resourc eTypeGeneral ="literature" uri="http://pu rl.org/coar/res ource_type/c_6501">journa l article</oi re:resourceT ype>	COAR Resource Type Vocabulary

Resource Identifier (Identificador do Recurso)	datacite:identifier	Identificador único do recurso	http://hdl.handle.net/1234/5628	<datacite:identifier identifier Type="Handle">http://hdl.handle.net/1234/5628</datacite:identifier>	identifier type
Access Rights (Direitos de Acesso)	datacite:rights	Informações sobre como o recurso pode ser acessado, utilizado e disseminado	Open Access (Acesso Aberto)	<datacite:rights rightsURI="http://purl.org/coar/access_right/c_abf2">open access</datacite:rights>	COAR Access Right Vocabulary

A seguir na próxima tabela aborda-se o nível de obrigatoriedade Mandatório se aplicável (MA):

Tabela 7: Metadados mandatórios se aplicáveis

Campo	Metadado	Definição	Exemplo	Exemplo de aplicação	Vocabulário controlado
Contributor (Colaborador)	datacite:contributor	Pessoa ou instituição que tenha tido algum tipo de colaboração na elaboração do recurso, faz ser obrigatório o uso do atributo contributorType como demonstrado no exemplo (contributorType possui vocabulário controlado)	Evans, R. J.	<datacite:contributor> <datacite:contributor Name>Evans, R. J.</datacite:contributor Name> <datacite:contributor>	name type e contributor type

Funding Reference (Referência de financiamento)	oaire: funding Reference	Informações sobre apoio financeiro recebido. Subpropriedade funderName obrigatória	European Commission (Comissão Euripéia)	<oaire: funding Reference> < oaire:funderName> European Commission </datacite: funderName > </oaire: funding Reference>	funder Identifier type
Embargo Period Date (Período de embargo)	datacite: date	Datas relevantes para descrever o período de embargo. O campo direitos de acesso deve estar como “embargoed access” e devem ser utilizados as datas de início e fim do período de embargo além dos atributos dateType	2011-12-01 2012-12-01	<datacite: dates> <datacite: date dateType="Accepted" 2011-12-01</datacite: date ><datacite: date date Type="Available" >2012-12-01 </datacite: date></datacite: dates>	date type
Language (Língua)	dc: language	Idioma do conteúdo do recurso	eng (Inglês)	<dc:language>eng</dc: language>	IETF BCP 47
Publisher (Publicador)	dc: publisher	O publicador do recurso, usar apenas nome completo.	John Wiley & Sons, Inc. (US)	<dc: publisher> John Wiley & Sons , Inc. (US)</dc:publisher >	

Description (Descrição)	dc: descriptio n	Um texto descrevendo o conteúdo do recurso	Um número de problemas no estado quântico são abordados	<dc: description xml:lang= "en- US»Um número de problemas no estado quântico são abordados </dc: description>	
Subject (Assunto)	datacite: subject	Assunto geral, palavra-chave, ou qualquer outro código ou termo que representa o recurso, se utilizar algum código utilizar o atributo subjectScheme	Earth sciences and geology (Ciências da terra e geologia)	<datacite: subject> Earth sciences and geology</ datacite: subject>	
File Location (Localização do arquivo)	oaire:file	referência para desambiguação do arquivo, utilizar um HTTP URI	http://purl. org/coar/ access_ right/c_ abf2" mimeType ="" application/ pdf	<oaire:file accessRights URI="http:// purl.org/coar /access_right /c_abf2"mi meType="" application/ pdf"object Type="fullte xt»http://lin k-to-the-fullt ext.org</oair e:file>	

A próxima tabela se trata dos campos Recomendados (R) :

Tabela 8: Metadados Recomendados

Campo	Metadado	Definição	Exemplo	Exemplo de aplicação	Vocabulário controlado
Alternate Identifier (Identificador alternativo)	datacite:alternateIdentifier	Outro identificador que não seja o principal. Caso esse campo seja utilizado é obrigatório o uso do atributo alternateIdentifierType	http://someUrl	<datacite:alternateIdentifiers><datacite:alternateIdentifier alternateIdentifierType="http://someUrl">datacite:alternateIdentifier</datacite:alternateIdentifiers>	alternateIdentifier type
Related Identifier (Identificador relacionado)	oaire:fundingReference	Identificador de um outro recurso associado ao recurso principal descrito. Obrigatório o uso do atributo relatedIdentifierType (vocabulário controlado disponível) e o uso do atributo relationType (também utiliza vocabulário)	http://someUrl	<datacite:relatedIdentifiers><datacite:relatedIdentifier relatedIdentifierType="URL" relationType="HasPart">http://someUrl</datacite:relatedIdentifier></datacite:relatedIdentifiers>	relatedIdentifier type, relation type e resource type general
Format (Formato)	dc:format	Manifestação física ou digital do recurso	application/xml	<dc:format>application/xml</dc:format>	

Source (Fonte)	dc:source	Referência à outro recurso do qual o original descrito é derivado	Ecology Letters (1461023X) vol.4 (2001)	<dc:source> Ecology Letters (1461023X) vol.4 (2001) </dc:source>	
License Condition (Condição da licença)	oaire: license Condition	Informações sobre direitos e licença como uso e acesso ao recurso	Creative Commons Attribution- Non Commercial	<oa:license Condition s tartDate="2019-02-01" uri="http:// creative commons.org/ licenses/by-nc/ 4.0/»Creative Commons Attribution- Non Commercial</ oa:license Condition>	
Coverage (Cobertura)	dc: coverage	O escopo ou abrangência do conteúdo recurso	2000-2010	<dc:coverage> 2000-2010</dc: coverage>	
Resource Version (Versão do recurso)	oaire: version	Dependendo do tipo de recurso pode indicar, o número da versão ou status da publicação	1.0.3	<oa:version> 1.0.3</oa: version>	COAR Version Vocabulary
Citation Title (Título para citação)	oaire: citation Title	Título de onde o recurso está inserido conforme constará na citação bibliográfica	Jornal Title	<oa:citation Title>some Journal Title</ oa:citation Title>	
Citation Volume (Volume para citação)	oaire: citation Volume	Volume conforme constará na citação bibliográfica	10	<oa:citation Volume>10</ oa:citation Volume>	

Citation Issue (Exemplar de citação)	oaire: citation Issue	Exemplar conforme constará na citação bibliográfica	1	<oaire:citation Issue>1</oaire:citation Issue>	
Citation Start Page (Página inicial para citação)	oaire: citation StartPage	Página inicial conforme constará na citação bibliográfica	100	<oaire:citation StartPage>100</oaire:citation StartPage>	
Citation End Page (Página final para citação)	oaire: citation EndPage	Página final conforme constará na citação bibliográfica	105	<oaire:citation EndPage>105</oaire:citation EndPage>	
Citation Edition (Edição para citação)	oaire: citation Edition	Edição conforme constará na citação bibliográfica	2	<oaire:citation Edition>2</oaire:citation Edition>	
Citation Conference Place (Local da Conferência para citação)	oaire: citation ConferencePlace	Local da conferência conforme constará na citação bibliográfica	Berlin	<oaire:citation Conference Place>Berlin</oaire:citation Conference Place>	
Citation Conference Date (Data da Conferência para citação)	oaire: citation ConferenceDate	Data da conferência conforme constará na citação bibliográfica, deve ser conforme as regras W3C	2013-10-22		

Por fim a tabela com os campos opcionais (O):

Tabela 9: Metadados Opcionais

Campo	Metadado	Definição	Exemplo	Exemplo de aplicação	Vocabulário controlado
Size (Tamanho)	datacite: size	Tamanho do recurso, pode ser digital ou físico	6 MB	<datacite:size>6 MB</datacite: size>	
Geo Location (Geo localização)	datacite: geo Location	Localização geográfica ou espacial à qual o recurso se refere ou onde foi desenvolvido. Utilizar latitude e longitude como apresentado no exemplo. Caso deseje utilizar áreas ao invés de ponto de localização utilizar as sub-propriedades disponíveis	31.233 -67.302	<datacite:geo Location><data cite:geoLocation Place>Atlantic Ocean</datacite: geoLocation Place><datacite: geoLocation Point><datacite :pointLongitude> 31.233</datacite: pointLongitude> <datacite:point Latitude>-67.302 </datacite:point Latitude></data cite:geoLocation Point>	
Audience (Público)	dcterms: audience	Classe o entidade à qual o recurso pode ser útil ou foi pensado para	Researchers (Pesquisado res)	<dcterms: audience> Researchers</dc terms:audience>	

7.3 Planejando seu repositório de dados

Assim como qualquer repositório, o repositório de dados se beneficia de um planejamento bem estruturado e completo. Um planejamento de um repositório de dados deve então abarcar aspectos como:

- Objetivos e Missão
- Definição dos serviços prestados e seu funcionamento
- Papéis necessários para implementação e manutenção
- Competências necessárias aos diferentes papéis
- Custos envolvidos com implementação e manutenção
- Métodos de avaliação e indicadores

7.4 Política de dados

Ao desenvolver uma política de dados para seu repositório deve inicialmente conhecer os atores envolvidos no processo como um todo, para isso deve-se perguntar:

- Quem produzirá os dados e com qual motivo?
- Quem será responsável por realizar o depósito e curadoria dos dados?
- Quem poderá se interessar e utilizar os dados depositados?

A partir destas respostas se terá um conhecimento inicial para orientar as decisões que formarão sua política de dados. Outras informações importantes a se conhecer antes de se iniciar o desenvolvimento de uma política de dados é o conteúdo de outras políticas institucionais garantindo assim a concordância, algumas políticas que podem ser importantes: as políticas da biblioteca, a política de preservação, política do repositório institucional, entre outras.

Alguns tópicos são então essenciais à uma política de dados:

- Definição clara dos atores envolvidos, e que papéis irão desempenhar (workflows e tutoriais detalhados dos processos serão de grande valia aqui, como por exemplo manual de depósito, fluxo de curadoria, entre outros);
- Definição de terminologia e objetos importantes (como por exemplo: Conjunto de dados, dado de pesquisa, dado para pesquisa, etc);
- Descrição do ciclo de vida do dado desde sua ingestão até descarte (ou outro) incluindo em especial a questão sobre embargo;
- Descrição dos metadados utilizados e diretrizes adotadas.

Além dos tópicos acima mencionados vale destacar que pode ser importante para uma política de dados dependendo da instituição a orientação de elaboração dos chamados Plano de Gestão de Dados (PGD).

7.5 Plano de Gestão de Dados (PGD)

O plano de gestão de dados é um documento que deverá esclarecer o modo como os dados foram criados, quem poderá os acessar, quem será responsável por sua preservação e curadoria, o modo como serão documentados, onde e como serão armazenados e condições de partilha e reuso bem como quaisquer outra informação que o pesquisador julgar importante como por exemplo software para acesso, objetivo, etc. Se percebe uma considerável semelhança entre o PGD e a política de dados de uma instituição, porém a grande diferença entre uma política de dados e um plano de gestão de dados é sua autoria, na primeira os dados são tratados em um âmbito institucional enquanto na última os dados são tratados enquanto no encargo do pesquisador ou coletor.

O Data Curation Center (DCC) desenvolveu o DMPonline, que é uma ferramenta gratuita de desenvolvimento de planos de gestão de dados. Ele fornece orientação e exemplos personalizados para ajudar os pesquisadores a elaborar planos de gerenciamento de dados. Em seu template básico podemos encontrar as seguintes questões orientadoras:

- Quais dados você coletará ou criará?
- Como os dados serão coletados ou criados?
- Que documentação e metadados acompanharão os dados?
- Como você gerenciará problemas éticos?
- Como você gerenciará questões de direitos autorais e direitos de propriedade intelectual (DPI)?
- Como os dados serão armazenados e haverá backup durante a pesquisa?
- Como você gerenciará o acesso e a segurança?
- Quais dados são de valor a longo prazo e devem ser retidos, compartilhados e / ou preservados?
- Qual é o plano de preservação de longo prazo para o conjunto de dados?
- Como você compartilhará os dados?
- São necessárias restrições ao compartilhamento de dados?
- Quem será responsável pelo gerenciamento de dados?
- Quais recursos você precisará para entregar seu plano?

7.6 Checklist EIFL

A organização não lucrativa EIFL (Electronic Information for Libraries/ (Informações eletrônicas para bibliotecas) desenvolveu um checklist voltado para a melhoria de repositórios DSpace de acesso aberto. Porém as orientações podem ser adaptadas para a realidade dos repositórios de dados, mais especificamente em Dataverse. As orientações se resumem então em:

- O software adotado está atualizado? (No caso do Dataverse a última versão disponível é a 4.14)
- Utiliza algum identificador persistente? (O Dataverse já vem com o DOI pré-configurado porém oferece a possibilidade de utilização H do HANDLE)
- Medidas de recuperação/backup?
- Compatibilidade com OAI-PMH? (O Dataverse possibilita essa funcionalidade)
- Utiliza Padrão de metadados estabelecido? (O Dataverse já vem com padrões de metadados internacionais e bem estabelecidos além de possibilitar a customização ou adoção de outros padrões)
- Possibilita a ação de Motores de busca? (O Dataverse é indexável e encontrável por motores de busca)
- Registro no OpenDOAR e no BASE- Bielefeld Academic Search Engine?
- Estatísticas de tráfego? (O Dataverse conta com plugins para essa finalidade, além do guestbook)
- Responsivo? (O Dataverse conta com o Bootstrap que fornece um sistema de grid de 12 colunas, fluido e responsivo usado para organizar os layouts de página)
- Política do repositório?
- Licenciamento adequado? (O Dataverse vem com a Creative Commons CC0 Public Domain Dedication por padrão mas ela pode ser trocada)
- Adota alguma Identificação de autores como ORCID E IDLatter? (O Dataverse possibilita a integração com o ORCID porém)



SECTION

Referências

DATA CITE, DataCite Metadata Working Group. **DataCite Metadata Schema Documentation for the Publication and Citation of Research Data**. v. 4.1. 2017. DataCite e.V. 10.5438/0014. 72 p. Disponível em: [https://schema.datacite.org/meta/kernel-4.0/doc/DataCite-Meta dataKernel_v4.0.pdf](https://schema.datacite.org/meta/kernel-4.0/doc/DataCite-Meta%20dataKernel_v4.0.pdf). Acesso em: 31 out. 2019.

DATAVERSE PROJECT, Institute for Quantitative Social Science. **Dataverse Documentation**. v. 4.17. 2019. Disponível em: <http://guides.dataverse.org/en/latest/>. Acesso em: 31 out. 2019.

DCC, DMPOnline. **Digital Curation Centre: DCC Template**. 2018. Disponível em: https://dmponline.dcc.ac.uk/template_export/1638514350.pdf 2010. Acesso em: 31 out. 2019.

FOSTER, Facilitate Open Science Training for European Research. **Planos de Gestão de Dados no Horizonte 2020**. Disponível em: <https://www.fosteropenscience.eu/node/26432019>. Acesso em: 31 out. 2019.

KUCHMA, Iryna. **EIFL Checklist: how to make your DSpace Open Access repository work really well**. 2019. 14 p. Disponível em: https://www.eifl.net/system/files/resources/201907/repositories_checklist_july_2019__0.pdf. Acesso em: 31 out. 2019.

LEITE, Fernando César Lima. **Como gerenciar e ampliar a visibilidade da informação científica brasileira: repositórios institucionais de acesso aberto**. Brasília: Ibict, 2009.

120 p., il. Disponível em: <http://livroaberto.ibict.br/bitstream/1/775/4/Como%20gerenciar%20e%20ampliar%20a%20visibilidade%20da%20informa%C3%A7%C3%A3o%20cient%C3%ADfica%20brasileira.pdf>. Acesso em: 31 out. 2019.

OPENAIRE, Open Access Infrastructure for Research in Europe. **OpenAIRE Guidelines for Literature Repository Managers**. v. 4. 2018. Disponível em: <https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/introduction.html>. Acesso em: 31 out. 2019.

RDP Brasil. **Acesso aberto a dados de pesquisa no Brasil**: planejamento para implantação de Comunidade Produtora de Dados para o Repositório Rede de Dados de Pesquisa – RDP Brasil de Dados. 2019. 40 p. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/185195/001082283.pdf?sequence=1&isAllowed=y>. Acesso em: 31 out. 2019.

ROSA, Juliane Alves, et al. **Manual para o gerenciamento de repositórios digitais criados em DSpace**. Brasília: Ibict, 2016. 132 p., il. Disponível em: https://wiki.duraspace.org/download/attachments/112528241/Manual_Gerenciamento_de_reposit%C3%B3rios_digita%20is_DSpace_final_2016.pdf?version=1&modificationDate=1556132200816&api=v2. Acesso em: 31 out. 2019.