

ACESSO ABERTO A DADOS DE PESQUISA NO BRASIL

Dataverse: documentação de instalação e configurações –
Debian9



Controle de Versionamento:

Versão	Autores	Data
1.0	Universidade Federal do Rio Grande do Sul <ul style="list-style-type: none">• Rafael Port da Rocha• Sônia Elisa Caregnato• Caterina Groposo Pavão• Rene Faustino Gabriel Junior• Samile Andréa de Souza Vanz• Paula Caroline Schifino Jardim Passos• Victor Andrews Garcia Lima• Iván Andrés Fornos Angues Universidade Federal de Rio Grande <ul style="list-style-type: none">• Eduardo Nunes Borges• Luís Alberto Barbosa Azambuja	12/08/2019

Coordenador

Rafael Port da Rocha (Universidade Federal do Rio Grande do Sul)

Coordenadora adjunta

Sônia Elisa Caregnato (Universidade Federal do Rio Grande do Sul)

Pesquisadores Universidade Federal do Rio Grande do Sul

Caterina Groposo Pavão

Paula Caroline Schifino Jardim Passos

Rene Faustino Gabriel Junior

Samile Andréa de Souza Vanz

Universidade Federal do Rio Grande

Eduardo Nunes Borges

Luís Alberto Barbosa Azambuja

Bolsistas

Iván Andrés Fornos Angues

Victor Andrews Garcia Lima

Contato

e-mail: dadosdepesquisa@rnp.br - site: <https://dadosdepesquisa.rnp.br>

Rua Ramiro Barcelos, 2777 - Campus Saúde Sala: 106 - Anexo 1 - Brasil -

Porto Alegre - RS - CEP 90.035-007 - Telefone: +55(51)3308.5942



Como citar: GABRIEL JUNIOR, Rene Faustino; LIMA, Victor Andrews Garcia; AZAMBUJA, Luís Alberto Barbosa; CAREGNATO, Sônia Elisa; PASSOS, Paula Caroline Schifino Jardim; PAVÃO, Caterina Groposo; BORGES, Eduardo Nunes; ROCHA, Rafael Port da; VANZ, Samile Andrea de Souza. Acesso aberto a dados de pesquisa no Brasil: dataverse: documentação de instalação e configuração – Debian9. Disponível em: <<http://hdl.handle.net/20.500.11959/1201>>

Acesso aberto a dados de pesquisa no Brasil: dataverse: documentação de instalação e configuração – Debian9

Resumo

Este relatório apresenta os resultados do projeto de pesquisa **Rede de Dados de Pesquisa Brasileira (RDP Brasil)**, desenvolvido pela Universidade Federal do Rio Grande do Sul (UFRGS) em parceria com a Universidade Federal do Rio Grande (FURG), sob coordenação executiva da Rede Nacional de Ensino e Pesquisa (RNP) e do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT). O relatório apresenta os procedimentos de instalação de uma versão do Dataverse com seus pré-requisitos e ferramentas de registro como o Handle. O relatório foi desenvolvido com base em uma Virtual Machine (VM) de dois processadores, com 4 Giga de RAM e disco rígido de 100Giga utilizando o Ubuntu 18.04 Server. São descritos todos os passos para implementação, podendo ser utilizado outra versão do Linux. Existe um capítulo à parte para a implementação do Handle.

Palavras-chave:

Instalação Dataverse. Debian 9. Dados abertos. Dados de pesquisa. Compartilhamento de dados. Tecnologias.

Sumário

DATAVERSE	6
1.1 Sobre o projeto Dataverse	6
1.2 Sobre a versão 4 do Dataverse	7
1.2.1 Recomendações para instalação do Dataverse:	8
1.2.2 Componentes opcionais	8
1.2 Formatos de Citação	9
INSTALAÇÃO	10
2.1 Requisitos de Hardware	10
Requisitos de Software	10
2.2.1 Pré-Requisitos de Software para Instalação	10
2.2.1.1 Linux	10
2.2.1.2 Unzip	10
2.2.1.3 VIM e NANO	10
2.2.1.4 WGET	10
2.2.1.5 GIT	11
2.2.1.6 LYNX	11
2.2.1.7 CURL	11
2.2.1.8 Java	11
2.2.1.9 Dataverse (download)	12
2.2.1.10 Glassfish	12
2.2.1.11 Postgtesql	15
2.2.1.12 Instalando jq	19
2.2.1.13 ImageMagick	19
2.2.1.14 Projeto R	20
INSTALAÇÃO DO DATAVERSE	21
3.1 Instalando o Dataverse (Versão DVInstall)	21
3.2 Ativando o R Server no Dataverse	22
Para testar o Dataverse acesse http://numero_IP/	23
DATAVERSE PARA MODO HANDLE	23
4.1 Para o DOI:	23
4.2 Para Handle:	24
5 HANDLE SERVER	27
5.1 HANDLE COMO SERVIDOR PRIMÁRIO	27
5.2 HANDLE COMO SERVIDOR SECUNDÁRIO	28

5.3 CONCLUINDO A INSTALAÇÃO DO HANDLE	30
5.4 REQUISITANDO UM PREFIXO	30
4 USANDO O DATAVERSE	32
REFERÊNCIAS	34

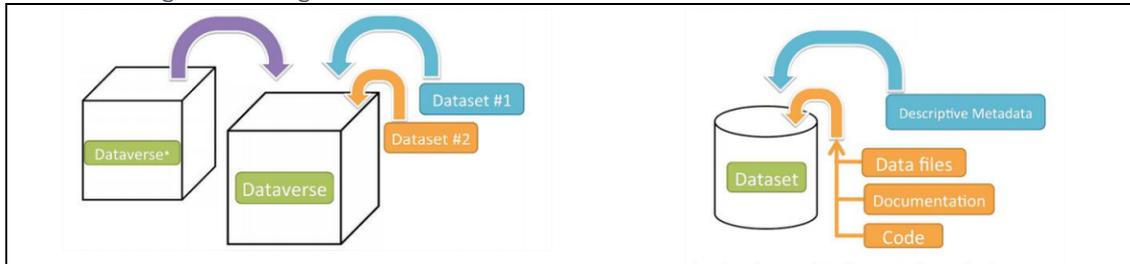
1 DATAVERSE

1.1 Sobre o projeto Dataverse

O Dataverse é um aplicativo da Web de código-fonte aberto para compartilhar, preservar, citar, explorar e analisar dados de pesquisa. Ele facilita a disponibilização de dados para outras pessoas e permite que você replique o trabalho de outras pesquisas com mais facilidade. Pesquisadores, periódicos, autores de dados, editores, distribuidores de dados e instituições afiliadas recebem crédito acadêmico e visibilidade na web.

Um repositório Dataverse é a instalação do software, que hospeda vários arquivos virtuais chamados Dataverses. Cada Dataverse contém conjuntos de dados e cada conjunto de dados contém metadados descritivos e arquivos de dados (incluindo documentação e código que acompanham os dados). Como método de organização, os Dataverses também podem conter outros Dataverses.

Figura 1 - Diagrama semântico do Dataverse e Dataset no Dataverse 4.0



Fonte: Quingley (2016).

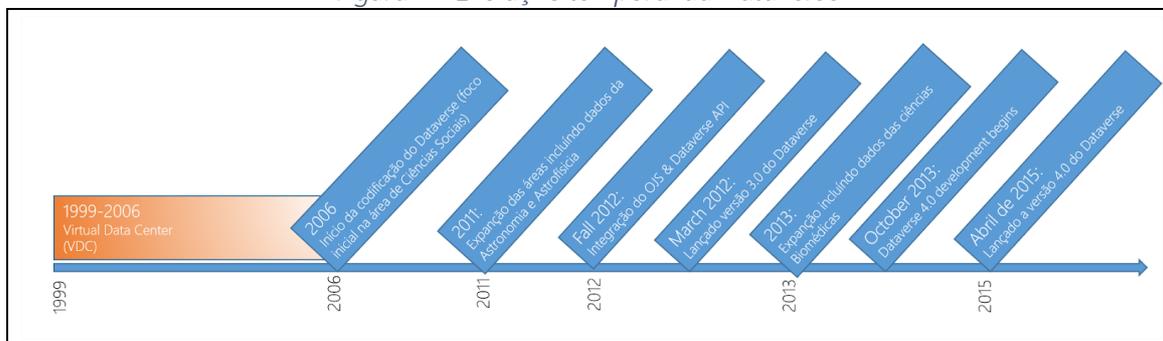
A percepção central por trás do Dataverse é automatizar grande parte do trabalho do arquivista profissional, fornecer serviços e distribuir crédito ao criador de dados. Antes do Dataverse, os pesquisadores eram forçados a escolher entre receber crédito por seus dados, controlando a distribuição, mas sem garantias de preservação a longo prazo, ou mantendo garantias de preservação a longo prazo, enviando-o para um arquivo profissional, mas sem receber muito crédito. O Dataverse quebra essa má escolha: colocamos um Dataverse (um arquivo virtual) em seu site que tem aparência, marca e URL de seu site, juntamente com uma citação acadêmica para os dados que lhe dão

total crédito e visibilidade na Web. No entanto, essa página do seu site é servida por um repositório Dataverse, com apoio institucional e garantias de preservação a longo prazo.

1.2 Histórico

O software Dataverse está sendo desenvolvido no Instituto Harvard para Ciências Sociais Quantitativas (IQSS), juntamente com muitos colaboradores e colaboradores em todo o mundo. O Dataverse foi construído com base em na experiência do projeto anterior de Virtual Data Center (VDC), que abrangeu o período de 1999 a 2006 como uma colaboração entre o Data Center Harvard-MIT (agora parte do IQSS) e a Biblioteca da Universidade de Harvard.

Figura 2 - Evolução temporal da Dataverse



Fonte: Adaptado pelos autores de Quingley (2016).

1.2 Sobre a versão 4 do Dataverse

O Dataverse é um aplicativo desenvolvido em Web Java Enterprise Edition (EE) fornecido como um arquivo de WAR (arquivo da Web).

1.2.1 RECOMENDAÇÕES PARA INSTALAÇÃO DO DATAVERSE:

- Linux: RHEL / CentOS é altamente recomendado, já que todo o desenvolvimento e QA acontece nesta distribuição, entretanto, este relatório foi feito em Ubuntu 18.03 server.
- Glassfish: um servidor de aplicativos Java EE onde o aplicativo Dataverse (arquivo war) deve ser implantado.
- PostgreSQL: um banco de dados relacional.
- Solr: um mecanismo de pesquisa. Um esquema específico do Dataverse é fornecido.

- e) Servidor SMTP: para enviar e-mails de redefinições de senha e outras notificações.
- f) Serviço de identificador persistente: Suporte DOI e ou Handle para o uso em produção, sendo necessário uma autoridade DOI ou Handle.net registrada.

1.2.2 COMPONENTES OPCIONAIS

Existem vários componentes opcionais que você pode escolher instalar ou configurar, incluindo:

- a) R, rApache, Zelig e TwoRavens: Tabular Data Exploration.
- b) Integração com Dropbox.
- c) Apache: um servidor web que pode “inverter proxy” em aplicações Glassfish e reescrever o tráfego HTTP.
- d) Shibboleth: um sistema de autenticação descrito no Shibboleth . seu uso com o Dataverse requer o Apache.
- e) OAuth2: um sistema de autenticação descrito no login do OAuth: ORCID, GitHub, Google .
- f) Geoconnect: um sistema que permite aos usuários criar mapas a partir de arquivos geoespaciais, descritos no Geoconnect.

1.2 Formatos de Citação

Autor(es) , **Ano da publicação** , **Título do Dataset** , **Nome do repositório ou arquivo** , **versão** , **Identificador Global persistente**

Campo	Descrição
Autor(es)	Nome dos autores
Ano de publicação	Data da publicação dos dados de pesquisa
Título do Dataset	Nome do Dataset definido pelo(s) autor(es)
Nome do repositório	
Versão	Identificação da versão dos dados disponibilizados

Identificador Global persistente	Número do DOI ou Handle utilizado para identificar os dados de pesquisa
---	---

2 INSTALAÇÃO

2.1 Requisitos de Hardware

Uma instalação básica do Dataverse é executada corretamente em hardware modesto, com recomendações mínimas de uma única máquina virtual com dois processadores de 2,8 GHz, 8 GB de RAM e 50 GB de disco.

2.2 Requisitos de Software

2.2.1 PRÉ-REQUISITOS DE SOFTWARE PARA INSTALAÇÃO

2.2.1.1 LINUX

O Dataverse funciona em várias distribuições do Linux, a recomendação do desenvolvedor é para utilizar o RHEL / CentOS 7, entretanto os testes realizados no Laboratório no Centro de Documentação e Acervo Digital da Pesquisa (CEDAP) da Universidade Federal do Rio Grande do Sul (UFRGS) foi realizado no Ubuntu 18.04 Server, configurado para dois núcleos, 4 Giga de RAM e 100 Giga de disco, máquina virtualizada no VirtualBox.

Para instalação dos pacotes utilize o Super Usuário (su)

```
# sudo su
# sudo apt update # atualiza lista de pacotes
# sudo apt upgrade # atualiza pacotes
```

2.2.1.2 UNZIP

Desc

```
# apt-get install zip
```

2.2.1.3 VIM E NANO

Editores de arquivos

```
# apt-get install vim nano
```

2.2.1.4 WGET

```
# apt-get install wget
```

2.2.1.5 GIT

```
# apt-get install git
```

2.2.1.6 LYNX

Instalação de um browser texto para testar os recursos instalados

```
# apt-get install lynx
```

2.2.1.7 CURL

```
# apt-get install curl
```

2.2.1.8 JAVA

O Dataverse requer o Java SE 8 (8u74 / JDK 1.8.0u 74 ou superior).

Conectado como administrador, conta do su, instale a versão 8 do Java, observe que tem de ser a versão JDK.

```
# apt install default-jdk  
# java -version
```

Caso tenha mais de uma versão do Java instalado, pode-se definir o Java 8 como versão default.

```
# update-alternatives --config java
```

Para testar a versão corrente do java no sistema utilize:

```
# java -v
```

2.2.1.5 Docker

```
#sudo apt install apt-transport-https ca-certificates curl  
software-properties-common gnupg2
```

```
#curl -fsSL https://download.docker.com/linux/debian/gpg | sudo apt-key  
add -  
#sudo add-apt-repository "deb [arch=amd64]  
https://download.docker.com/linux/debian $(lsb_release -cs) stable"  
#sudo apt update  
#sudo apt install docker-ce
```

Para testar o Docker

```
#sudo systemctl status docker  
#docker -v
```

2.2.1.9 DATAVERSE (DOWNLOAD)

Alguns componentes e arquivos do Dataverse são necessário para instalação dos pré-requisitos. Baixe a versão do dvininstall.zip em uma pasta.

Diretório da versão para instalação do Dataverse /home/dataverse

Baixe o dvininstall.zip e o v4.11.tar.gz preferencialmente no diretório /home/dataverse e descompacte os arquivos

```
# useradd -m dataverse  
# cd /home/dataverse/  
# git clone https://github.com/IQSS/dataverse.git  
# wget  
  https://github.com/IQSS/dataverse/releases/download/v4.15.1/dvininstall  
.zip  
# wget https://github.com/IQSS/dataverse/archive/v4.11.tar.gz  
# unzip dvininstall.zip  
# unzip v4.11.tar.gz
```

2.2.1.10 GLASSFISH

Criar um usuário para o Glassfish

```
useradd -m glassfish
```

O parâmetro *-m* faz com que o sistema crie um diretório *glassfish4* na para /home/

Baixando a versão 4 do Glassfish na pasta /home/glassfish4

```
# su glassfish4
```

```
# cd /home/glassfish
# wget http://dlc-cdn.sun.com/glassfish/4.1/release/glassfish-4.1.zip
# unzip glassfish-4.1.zip
# sudosu
# mv glassfish4 /usr/local
# chown -R root:root /usr/local/glassfish4
# chown glassfish /usr/local/glassfish4/glassfish/lib
# chown -R glassfish:glassfish
  /usr/local/glassfish4/glassfish/domains/domain1
```



Uma vez que o Glassfish esteja instalado, é necessária uma versão mais atualizada da Weld Library (v2.2.10.SP1), para resolver um problema com a versão que vem com o glassfish 4.1

1.

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

Edite o arquivo domain.xml e altere a linha que contenha `-client` para `-server` em `<jvm-options>-client</jvm-options>`

```
# nano /usr/local/glassfish4/glassfish/domains/domain1/config/domain.xml
```

Ativando o GlassFish

```
/lib/jvm/java-11-openjdk-i386
```

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

```
Waiting for domain1 to start .....
Successfully started the domain : domain1
domain Location: /usr/local/glassfish4/glassfish/domains/domain1
Log File: /usr/local/glassfish4/glassfish/domains/domain1/logs/server.log
Admin Port: 4848
Command start-domain executed successfully.
```

...

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

¹ <https://github.com/IQSS/dataverse/issues/647>

```
sadmin osgi lb | grep 'Weld OSGi Bundle'  
216|Resolved | 1|Weld OSGi Bundle (2.2.10.SP1)
```

O arquivo certificado de autoridade (CA) do bundle de certificados do Glassfish contém certificados que expiraram em agosto de 2018, causando problemas com o login do ORCID. A data de expiração era 22 de agosto de 2018, isso pode ser visto com o comando para fixar o problema gerando um novo certificado:

```
# keytool -list -v -keystore  
  /usr/local/glassfish4/glassfish/domains/domain1/config/cacerts.jks
```

Sobrescreva o certificado CA do glassfish com o arquivo que veio com o java do sistema operacional e reinicie o Glassfish:

```
#cp /usr/lib/jvm/java-8-openjdk-amd64/jre/lib/security/cacerts  
  /usr/local/glassfish4/glassfish/domains/domain1/config/cacerts.jks  
# /usr/local/glassfish4/bin/asadmin stop-domain  
# /usr/local/glassfish4/bin/asadmin start-domain
```

Para testar o Glassfish

```
# Lynx http://localhost:8080
```

Transformando o Glassfish em um serviço do Linux

```
#!/bin/sh  
# chkconfig: 2345 80 01  
# description: GlassFish App Server  
set -e  
  
ASADMIN=/usr/local/glassfish4/bin/asadmin  
GF_USER=glassfish  
  
case "$1" in  
  start)  
    echo -n "Starting GlassFish server: glassfish"  
    # Increase file descriptor limit:  
    ulimit -n 32768  
    # Allow "memory overcommit":  
    # (basically, this allows to run exec() calls from inside the  
    # app, without the Unix fork() call physically hogging 2X  
    # the amount of memory glassfish is already using)
```

```

    echo 1 > /proc/sys/vm/overcommit_memory
    LANG=en_US.UTF-8; export LANG

    sudo -u $GF_USER $ASADMIN start-domain domain1
    ;;
stop)
    echo -n "Stopping GlassFish server: glassfish"

    sudo -u $GF_USER $ASADMIN stop-domain domain1
    echo "."
    ;;
*)
    echo "Usage: /etc/init.d/glassfish {start/stop}"
    exit 1
esac
exit 0

```

2.2.1.11 POSTGRES

Para instalar o Postgres no Ubuntu digite:

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

Depois de instalado, altere as permissões de acesso no arquivo:
postgresql.conf

```
# nano /etc/postgresql/10/main/postgresql.conf
```

Alterando a linha:

```
#listen_addresses = 'localhost'
para
listen_addresses = '*' #libera para todas as conexões ou
listen_addresses = 'localhost' #somente para servidor local

```

Para permitir acesso, altere também o arquivo pg_hba.conf

```
nano /etc/postgresql/10/main/pg_hba.conf
```

Alterar a linha de:

```
host    all             all             127.0.0.1/32    md5
```

Para:

```
host    all          all          127.0.0.1/32      trust
```

....

....

2.2.1.8 SOLR

Você não deve executar o Solr como root. Crie um usuário chamado **solr** um diretório para instalar o Solr em:

```
# useradd -m solr
```

Crie uma pasta para instalação do SLOR com o usuário solr, e atribua os privilégios do grupo para esse usuário (chown)

```
# mkdir /usr/local/solr
# chown solr /usr/local/solr
```

Torne-se o solr usuário e, em seguida, baixe e configure o Solr:

```
# su solr
# cd /usr/local/solr
# wget https://archive.apache.org/dist/lucene/solr/7.3.0/solr-7.3.0.tgz
# tar xvzf solr-7.3.0.tgz
# cd solr-7.3.0
# cp -r server/solr/configsets/_default server/solr/collection1
```

Utilizando o arquivo "dvininstall.zip" baixado na etapa de pré-requisitos. Extraia ele em /home/dataverse/dvininstall se ainda não o tiver feito. Então copie os arquivos nos seguintes diretórios.

```
cp /home/dataverse/dvininstall/schema.xml
   /usr/local/solr/solr-7.3.0/server/solr/collection1/conf
cp /home/dataverse/dvininstall/solrconfig.xml
   /usr/local/solr/solr-7.3.0/server/solr/collection1/conf
```



O Dataverse requer uma mudança no jetty.xml que vem junto com o Solr. Edite: /usr/local/solr/solr-7.3.0/server/etc/jetty.xml, aumentando requestHeaderSize de 8192 para 102400

Configurando o arquivo jetty.xml do SolR

```
nano /usr/local/solr/solr-7.3.0/server/etc/jetty.xml
```

Deve-se alterar a linha:

```
<Set name="requestHeaderSize"><Property  
  name="solr.jetty.request.header.size" default="8192" /></Set>  
para  
<Set name="requestHeaderSize"><Property  
  name="solr.jetty.request.header.size" default="102400" /></Set>
```

O Solr vai avisar sobre precisar aumentar o número de descritores de arquivos e processos máximos em um ambiente de produção, mas ainda vai rodar com os padrões. O dataverse já aumenta esses padrões para os níveis recomendados ao adicionar a linha **ulimit -n 65000** no script de inicialização, mas para maior eficiência, coloque o seguinte no arquivo `/etc/security/limits.conf`:

```
nano /etc/security/limits.conf
```

Insira os parâmetros antes do fim do arquivo:

```
#@student -maxlogins 4  
solr soft nproc 65000  
solr hard nproc 65000  
solr soft nofile 65000  
solr hard nofile 65000  
root soft nproc 65000  
root hard nproc 65000  
root soft nofile 65000  
root hard nofile 65000  
# End of file
```

Reinicializando o Glassfish

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

Para testar no navegador digite:

```
# Lynx http://localhost:8080/solr
```

Para ativar o solr, deve-se trocar de usuário root para solr

```
# su solr
# cd /usr/local/solr/solr-7.3.0
# bin/solr start
# bin/solr create_core -c collection1 -d server/solr/collection1/conf/
# exit
```



Para que o solr seja inicializado cada vez que o servidor for reiniciado, é preciso criar um serviço dentro do Linux. Veja também o diretório de instalação do solr (ex: /usr/local/solr/solr-7.3.1)

Como administrador, copie o conteúdo abaixo em um arquivo dentro da pasta /etc/init.d

nano /etc/init.d/solr

```
#!/bin/sh
# Starts, stops, and restarts Apache Solr.
#
# chkconfig: 35 92 08
# description: Starts and stops Apache Solr
SOLR_DIR="/usr/local/solr/solr-7.3.1"
SOLR_COMMAND="bin/solr"
SOLR_ARGS="-m 1g"
SOLR_USER=solr
case $1 in
  start)
    echo "Starting Solr"
    # Increase file descriptor limit:
    ulimit -n 65000
    cd $SOLR_DIR
    sudo -u $SOLR_USER $SOLR_COMMAND start $SOLR_ARGS
    ;;
  stop)
    echo "Stopping Solr"
    cd $SOLR_DIR
    sudo -u $SOLR_USER $SOLR_COMMAND stop $SOLR_ARGS
    ;;
  restart)
    $0 stop
    sleep 1
    $0 start
    ;;
  *)
    echo "Usage: $0 {start|stop|restart}" >&2
    exit 1
    ;;
)
```

```
esac
```

```
# sudo su  
# cd /etc/init.d/  
# nano solr  
# ifconf
```

2.2.1.12 INSTALANDO JQ

O “jq” é uma ferramenta de linha de comando para analisar a saída JSON usada pelo script de instalação do Dataverse. <https://stedolan.github.io/jq> explica várias maneiras de instalá-lo, mas um método relativamente simples é descrito abaixo. Por favor, note que você deve baixar a versão de 64 ou 32 bits com base em sua arquitetura. No exemplo abaixo, a versão de 64 bits está instalada. Confirmamos que é executável e em nosso \$PATH verificando a versão (1.4 ou superior deve estar bem):

```
# cd /usr/bin  
# wget http://stedolan.github.io/jq/download/Linux64/jq  
# chmod +x jq  
# jq --version  
→ jq-1.4
```

2.1.1.13 IMAGEMAGICK

O ImageMagick® é utilizado para criar, editar, compor ou converter imagens de bitmap. Ele pode ler e gravar imagens em vários formatos (mais de 200), incluindo PNG, JPEG, GIF, HEIC, TIFF, DPX , EXR , WebP, Postscript, PDF e SVG. Use o ImageMagick para redimensionar, inverter, espelhar, girar, distorcer, distorcer e transformar imagens, ajustar cores de imagem, aplicar vários efeitos especiais ou desenhar texto, linhas, polígonos, elipses e curvas de Bézier.

O ImageMagick é um software gratuito fornecido como uma distribuição binária pronta para ser executada ou como código-fonte que você pode usar,

copiar, modificar e distribuir tanto em aplicativos abertos quanto proprietários. Ele é distribuído sob uma licença derivada do Apache 2.0 .

```
# apt-get install imagemagick
```

2.2.1.14 PROJETO R

Acessar o R

Instale-os seguindo os procedimentos normais de instalação do pacote R.

Por exemplo, com os seguintes comandos R:

```
# sudo apt-key adv --keyserver keyserver.ubuntu.com --recv-keys  
E298A3A825C0D65DFD57CBB651716619E084DAB9  
# sudo add-apt-repository 'deb  
https://cloud.r-project.org/bin/linux/ubuntu bionic-cran35/'  
# sudo apt update  
# sudo apt install r-base  
# sudo apt install r-base-dev
```

Para iniciar o R:

```
# sudo -i R
```

Com o R iniciado, instale as bibliotecas

```
install.packages("R2HTML", repos="https://cloud.r-project.org/",  
lib="/usr/lib/R/library" )  
install.packages("rjson", repos="https://cloud.r-project.org/",  
lib="/usr/lib/R/library" )  
install.packages("DescTools", repos="https://cloud.r-project.org/",  
lib="/usr/lib/R/library" )  
install.packages("Rserve", repos="https://cloud.r-project.org/",  
lib="/usr/lib/R/library" )  
install.packages("haven", repos="https://cloud.r-project.org/",  
lib="/usr/lib/R/library" )
```

2.2.1.11 Apache Proxy

O Apache proxy permite um redirecionamento da porta 8080 para 80 e ou para a porta 443.

```
apt-get install apache2
sudo a2enmod proxy
sudo a2enmod proxy_http
sudo a2enmod proxy_balancer
sudo a2enmod lbmethod_byrequests
```

Editar o arquivo 000-default.conf

```
sudo nano /etc/apache2/sites-available/000-default.conf
```

Editar o conteúdo para:

```
<VirtualHost *:80>
    ProxyPreserveHost On
    ProxyPass / http://127.0.0.1:8080/
    ProxyPassReverse / http://127.0.0.1:8080/
</VirtualHost>
```

Reiniciar o serviço do apache:

```
# service apache2 restart
```

3 INSTALAÇÃO DO DATAVERSE

3.1 Instalando o Dataverse (Versão DVInstall)

Na pasta que foi extraída de dvininstall.zip nos pré-requisitos e rode o script de instalação

```
# cd /home/dataverse/dvininstall/
# ./install
```

O script vai pedir algumas configurações. Eis os valores padrões de cada opção

- Fully Qualified Domain Name of your host: **Nome do Host**
- Internet Address of your host: **localhost**
- User Glassfish: **glassfish**

- Glassfish Directory: **/usr/local/glassfish4**
- Administrator email address for this Dataverse: **(e-mail válido)**
- SMTP (mail) server to relay notification messages: **Insira um SMTP válido (próprio ou externo)**
- Postgres Server Address: **[127.0.0.1]**
- Postgres Server Port: **5432**
- Postgres ADMIN password: **(senha do postgres)**
- Name of the Postgres Database: **dvndb**
- Name of the Postgres User: **dvnapp**
- Postgres user password: **(senha do usuario postgres)**
- Remote Solr indexing service: **LOCAL**
- Rserve Server: **localhost**
- Rserve Server Port: **6311**
- Rserve User Name: **rserve**
- Rserve User Password: **rserve**
- Administration Email address for the installation;

3.2 Ativando o R Server no Dataverse

```
# cd /home/dataverse/dataverse-4.11/scripts/r/rserve/
# ./rserve-setup.sh
```

Para testar o Dataverse acesse http://numero_IP/

4 DATAVERSE PARA MODO HANDLE

Identificadores persistentes são tanto requerimentos quanto uma parte integral da plataforma Dataverse. Por padrão e por conveniências de teste, o instalador configura um namespace de teste temporário no Doi através do EZID². Através deste é possível criar e publicar Dataverses e Datasets não sendo, entretanto, possível citá-los ou garantir sua preservação.

² Não está mais disponível para uso.

Para configurar propriamente um identificador para uma instalação de produção, devem ser adquiridos uma conta e um namespace através de um dos seguintes provedores de DOI ou HDL mediante a cobrança de taxaço:

- EZID (<http://ezid.cdlib.org>), (nã dispõe mais do serviço)
- DataCite (<https://www.datacite.org>),
- Handle.Net (<https://www.handle.net>).

Uma vez tendo a conta cadastrada e um namespace DOI ou Handle funcionando, as seguintes opções da JVM e da database devem ser configuradas:

4.1 Para o DOI:

JVM:

- doi.baseurlstring
- doi.username
- doi.password

Database:

- :DoiProvider
- :Protocol
- :Authority
- :Shoulder
- :IdentifierGenerationStyle (optional)
- :DataFilePIDFormat (optional)

4.2 Para Handle:

JVM:

- dataverse.handlenet.admcredfile (chave privada de administração)
- dataverse.handlenet.admprivphrase (senha utilizada na encriptação da chave)
- dataverse.handlenet.index (index do handle, valor padrão é 300)

Database:

- :Protocol
- :Authority
- :IdentifierGenerationStyle (optional)
- :DataFilePIDFormat (optional)

Configurando a Database, as opções da Database são configuradas através de requisiçoões http através da linha de comando. Para esse fim utiliza-se o CURL

Para listar as configurações atuais, execute:

```
# curl localhost:8080/api/admin/settings
```

Para editar alguma configuração, execute:

```
# curl -X PUT -d '(novo valor)' localhost:8080/api/admin/settings/:(configuração)
```

Primeiro é necessário informar o Dataverse que será utilizado handle em vez do DOI que vem por padrão. Para isso, edita-se o valor de : "Protocol" para hdl

```
curl -X PUT -d 'hdl' localhost:8080/api/admin/settings/:Protocol
```

Em seguida, é necessário informar o Dataverse o prefixo handle através da configuração :Authority

```
# curl -X PUT -d '[seu_prefixo]' localhost:8080/api/admin/settings/:Authority
```

Configurar o estilo de geração do identificador, uma vez que o HANDLE funciona com números e o dataverse vem configurado de modo a gerar uma string de caracteres aleatórios para o doi, deve-se mudar a opção :IdentifierGenerationStrategy com o comando:

```
# curl -X PUT -d 'sequentialNumber' localhost:8080/api/admin/settings/:IdentifierGenerationStrategy
```

Após, baixe o arquivo `createsequence.sql` no link http://guides.dataverse.org/en/latest/_downloads/createsequence.sql e através do usuário postgres execute-o na tabela `dvndb`, para gerar o contador da sequência na base de dados.

```
# sudo -u postgres -i
# wget http://guides.dataverse.org/en/latest/_downloads/createsequence.sql
# psql dvndb
# \i createsequence.sql
```

Configurando a JVM

Há dois modos de configurar as opções da JVM: via linha de comando através ou editando o arquivo domain.xml. A maior diferença entre um e outro é que editar o arquivo requer que o servidor seja reiniciado para que as alterações façam efeito. Por tanto, é preferível configurar via linha de comando. Caso seja necessário, o arquivo domain.xml fica no seguinte caminho:

```
/usr/local/glassfish4/glassfish/domains/domain1/config/domain.xml
```

Para acessar o aplicativo e configurar via linha de comando, é necessário estar no diretório /bin do servidor.

```
cd /usr/local/glassfish4/glassfish/bin
```

Dentro do diretório, é só executar o app asadmin passando os comandos como parâmetros.

Para listas as configurações atuais execute:

```
./asadmin list-jvm-options
```

Caso o sistema peça por senha, ela se encontra no arquivo local-password dentro da pasta glassfish/domains/domain1/config

Para alterar configurações já definidas, se faz necessário deletar a versão atual e depois cria-la novamente através dos comandos delete-jvm-options e create-jvm-options.

Como o dataverse vem configurado para doi, inicialmente não há a necessidade de substituir(deletar e criar novamente) sendo possível, somente criar as configurações do HANDLE

```
./asadmin create-jvm-options  
'-Ddataverse.handlenet.admcredfile=/hs/svr_1/admpriv.bin'  
./asadmin create-jvm-options '-Ddataverse.handlenet.admprivphrase=SENHA DO  
ADMPRIV AQUI'  
./asadmin create-jvm-options '-Ddataverse.handlenet.index=300'
```

5 HANDLE SERVER

Instalando o Handle.net.

Crie a pasta hs na raiz e dentro desta baixe e extraia o zip do handle.net através do link no site https://www.handle.net/download_hnr.html

```
mkdir hs
cd hs
wget https://www.handle.net/hnr-source/handle-9.0.3-distribution.tar.gz
tar -xvzf handle-9.0.3-distribution.tar.gz
```

Em seguida, rode o script de instalação e siga os passos que serão descritos na seção abaixo (5.1) caso pretenda instalar o handle.net como servidor primário ou na seção (5.2) caso pretenda instalar como servidor secundário (*mirror*).

```
/hs/handle-9.0.3/bin/hdl-setup-server /hs/svr_1
```

5.1 HANDLE COMO SERVIDOR PRIMÁRIO

Will this be a "primary" server (ie, not a mirror of another server)?(y/n) [y]: **Y**

Will this be a dual-stack server (accessible on both IPv6 and IPv4)?(y/n) [n]: **N**

Through what network-accessible IP address should clients connect to this server? **SEU IP**

If different, enter the IP address to which the server should bind. **[enter]**

Enter the (TCP/UDP) port number this server will listen to [2641]: **[enter]**

What port number will the HTTP interface be listening to? [8000]: **[enter]**

Would you like to log all accesses to this server?(y/n) [y]: **[enter]**

("N" (Never), "M" (Monthly), "W" (Weekly), or "D" (Daily))?
[Monthly] : **[enter]**

Enter the version/serial number of this site [1]: **[enter]**

Please enter a short description of this server/site: **Breve
descrição do seu servidor**

Please enter the name of your organization: **Nome da
organização**

Please enter the name of a contact person for cc (optional)
[(none)]: **[enter]**

Please enter the telephone number of c or of cc (optional)
[(none)]: **[enter]**

Please enter the email address of c or of cc: **Insira um
email no formato xxxx@xxxxx.xxx**

Do you need to disable UDP services?(y/n) [n]: **[enter]**

Would you like to encrypt your private key?(y/n) [y]:
[enter]

Insira a senha do seu servidor primário

Would you like to encrypt your private key?(y/n) [y]:
[enter]

Insira novamente a senha do seu servidor primário

5.2 HANDLE COMO SERVIDOR SECUNDÁRIO

Para instalar o Handle como **servidor secundário**, apontando para um
servidor primário já existente e configurado:

Will this be a "primary" server (ie, not a mirror of
another server)?(y/n) [y]: **N**

Will this be a dual-stack server (accessible on both IPv6
and IPv4)?(y/n) [n]: **N**

Through what network-accessible IP address should clients connect to this server? **SEU IP**

If different, enter the IP address to which the server should bind. **[enter]**

Enter the (TCP/UDP) port number this server will listen to [2641]: **[enter]**

What port number will the HTTP interface be listening to? [8000]: **[enter]**

Would you like to log all accesses to this server?(y/n) [y]: **[enter]**

("N" (Never), "M" (Monthly), "W" (Weekly), or "D" (Daily))? [Monthly] : **[enter]**

Enter the address of a primary server: **IP DO SERVIDOR PRIMÁRIO**

Enter the port number of the same primary server [2641]: **[enter]**

Enter the version/serial number of this site [1]: **[enter]**

Please enter a short description of this server/site: **Breve descrição do seu servidor**

Please enter the name of your organization: **Nome da organização**

Please enter the name of a contact person for cc (optional) [(none)]: **[enter]**

Please enter the telephone number of c or of cc (optional) [(none)]: **[enter]**

Please enter the email address of c or of cc: **Insira um email no formato xxxx@xxxxx.xxx**

Do you need to disable UDP services?(y/n) [n]: **[enter]**

Would you like to encrypt your private key?(y/n) [y]: **N**
Não encriptar a chave privada do servidor secundário, pois quando o dataverse tentar acessar utilizando a senha do servidor primário ocorrerá erros.

Would you like to encrypt your private key?(y/n) [y]: **N**

5.3 CONCLUINDO A INSTALAÇÃO DO HANDLE

Após instalar o servidor primário, é necessário re-criptar o arquivo admpriv.bin usando o app hdl-keyutil. Para isso, utiliza-se o seguinte comando.

```
/hs/handle-9.0.3/bin/hdl-keyutil /hs/svr_1/admpriv.bin
```

Ao iniciar o script basta selecionar a opção de encriptar a chave utilizando a encriptação da versão 6 do handle. Isso evitará erros futuros com o dataverse.

Se instalar um servidor secundário, após concluir a instalação, é necessário copiar para a pasta onde o mesmo foi instalado o arquivo admpriv.bin do servidor primário já re-criptado pelo script mencionado acima.

5.4 REQUISITANDO UM PREFIXO

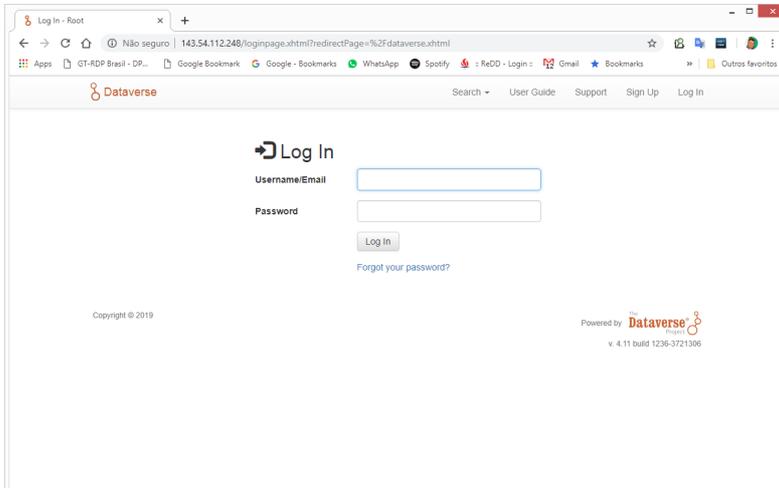
Vá ao site <http://hdl.handle.net/20.1000/111> e registre sua organização como uma provedora.

Em seguida vá ao site http://www.handle.net/prefix_request.html e faça o upload do arquivo sitebndl.zip que foi criado dentro da pasta onde foi instalado seu **servidor primário**.

Após receber por e-mail seu prefixo, edite o arquivo config.dct do servidor primário e do secundário (caso tenha criado um), substituindo todas as entradas onde esteja escrito YOUR_PREFIX pelo prefixo que recebeu por e-mail.

4 USANDO O DATAVERSE

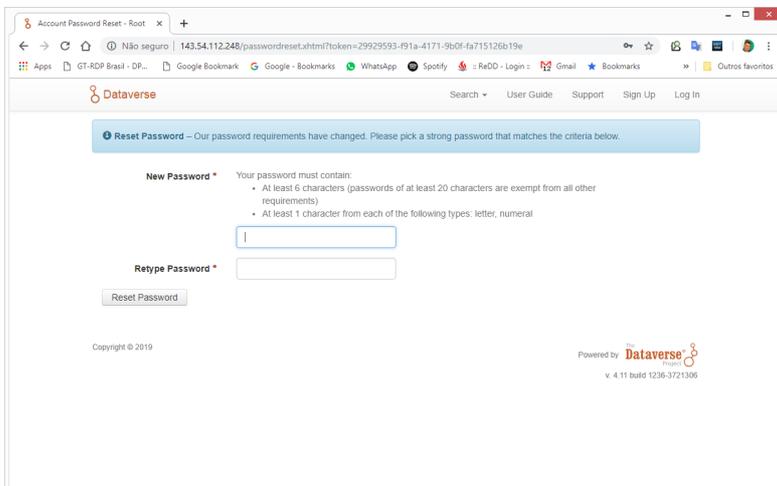
Para acessar o Dataverse, utilize o endereço <http://<número do ip do servidor>>



Se não foram alterados na instalação, os valores padrões de acesso são:

- username: **dataverseAdmin**
- password: **admin**

No primeiro acesso será necessário alterar a senha do superusuário



Tela principal do Dataverse

The screenshot shows the Dataverse interface for a dataset named "Root". The page is marked as "Unpublished". It displays a "Metrics" section with "0 Downloads" and buttons for "Contact", "Share", "Publish", and "Edit". Below this, there is a search bar with the text "Search this dataverse..." and buttons for "Find" and "Advanced Search". An "Add Data" button is also present. A summary box on the left lists "Dataverses (0)", "Datasets (0)", and "Files (0)". A central message states: "This dataverse currently has no dataverses, datasets, or files. You can add to it by using the Add Data button on this page." The footer includes "Copyright © 2019" and "Powered by The Dataverse Project v. 4.11 build 1236-3721306".

Configuração do DOI para o DataCite

```
Your Dataverse has been configured to use DataCite, to register DOI global
  identifiers in the test name space "10.5072" with the "shoulder" "FK2"
However, you have to contact DataCite (support@datacite.org) and request a test
  account, before you can publish datasets. Once you receive the account name and
  password, add them to your domain.xml, as the following two JVM options:
  <jvm-options>-Ddoi.username=...</jvm-options>
  <jvm-options>-Ddoi.password=...</jvm-options>
and restart Glassfish If this is a production Dataverse and you are planning to
  register datasets as "real", non-test DOIs or Handles, consult the "Persistent
  Identifiers and Publishing Datasets" section of the Installataion guide, on how
  to configure your Dataverse with the proper registration credentials.
```

REFERÊNCIAS

QUIGLEY, Elizabeth. The Dataverse Project. Disponível em: http://dataverse.org/files/dataverseorg/files/introduction_to_dataverse.pdf?m=1447352697. Acesso em 12/fev./2019.

KING, Gary. An Introduction to the Dataverse Network as an Infrastructure for Data Sharing. *Sociological Methods and Research*, n.36, p. 173–199, 2007.