

ACESSO ABERTO A DADOS DE PESQUISA NO BRASIL

Dataverse: ativação do HANDLE



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
1.1	Universidade Federal do Rio Grande do Sul <ul style="list-style-type: none">• Rene Faustino Gabriel Junior	29/10/2019
1.2	Universidade Federal do Rio Grande do Sul (ajustes no HANDLE) <ul style="list-style-type: none">• Rene Faustino Gabriel Junior	10/12/2019

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; AZAMBUJA, Luís Alberto Barbosa; CAREGNATO, Sônia Elisa; PASSOS, Paula Caroline Schifino Jardim; PAVÃO, Caterina Groposo; BORGES, Nunes Borges; ROCHA, Rafael Port da; VANZ, Samile Andrea de Souza. Acesso aberto a dados de pesquisa no Brasil: ativação do Handle. 2020. Disponível em: <http://hdl.handle.net/20.500.11959/1266>

Rene Faustino Gabriel Junior
Luís Alberto Barbosa Azambuja
Victor Andrews Garcia Lima

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: Dataverse. Dados abertos. Dados de pesquisa. Compartilhamento de dados. Tecnologias.

Sumário

1	DATAVERSE	Erro! Indicador não definido.
1.1	Sobre o projeto Dataverse	Erro! Indicador não definido.
1.2	Sobre a versão 4 do Dataverse	Erro! Indicador não definido.
1.2.1	Recomendações para instalação do Dataverse:	Erro! Indicador não definido.
1.2.2	Componentes opcionais	Erro! Indicador não definido.
1.2	Formatos de Citação	Erro! Indicador não definido.
2	INSTALAÇÃO	Erro! Indicador não definido.
2.1	Requisitos de Hardware	Erro! Indicador não definido.
2.2	Requisitos de Software	Erro! Indicador não definido.
2.2.1	Pré-Requisitos de Software para Instalação	Erro! Indicador não definido.
2.2.1.1	Linux	Erro! Indicador não definido.
2.2.1.2	Unzip	Erro! Indicador não definido.
2.2.1.3	LYNX	Erro! Indicador não definido.
2.2.1.4	Java	Erro! Indicador não definido.
2.2.1.5	Dataverse (dOWNLOAD)	Erro! Indicador não definido.
2.2.1.6	Glassfish	Erro! Indicador não definido.
2.2.1.7	Postgtesql	Erro! Indicador não definido.
2.2.1.8	Instalando jq	Erro! Indicador não definido.
2.2.1.9	ImageMagick	Erro! Indicador não definido.
2.2.1.10	Projeto R	Erro! Indicador não definido.
3	INSTALAÇÃO DO DATAVERSE	Erro! Indicador não definido.
3.1	Instalando o Dataverse	Erro! Indicador não definido.
3.2	Ativando o R Server no Dataverse	Erro! Indicador não definido.
4	DATAVERSE PARA MODO HANDLE	Erro! Indicador não definido.
4.1	Para o DOI:	Erro! Indicador não definido.
4.2	Para Handle:	6
5	HANDLE SERVER	Erro! Indicador não definido.
5.1	HANDLE COMO SERVIDOR PRIMÁRIO	7
5.2	HANDLE COMO SERVIDOR SECUNDÁRIO	8
5.3	CONCLUINDO A INSTALAÇÃO DO HANDLE	10
5.4	REQUISITANDO UM PREFIXO	10
4	USANDO O DATAVERSE	Erro! Indicador não definido.
	REFERÊNCIAS	Erro! Indicador não definido.

1 HANDLE SYSTEM

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 Handle. Através deste é possível criar e publicar Dataverses e Datasets não sendo, entretanto, possível citá-los ou garantir sua preservação.

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:

- DataCite (<https://www.datacite.org>),
- Handle.Net (<https://www.handle.net>).

1.1 Prefix do DOI ou Handle

Na geração automática do DOI ou Handle pelo Dataverse pode-se utilizar um sufixo para identificar o servidor ou repositórios com a atribuição de um sufixo do identificador persistente.

Como exemplo, o prefixo do Handle 20.500.12401 quando gerado o sufixo seria apresentado como `https://hdl.handle.net/20.500.12401/dv/59`, sendo o 59 o número de registro. Com a inserção do sufixo, o identificador ficaria `https://hdl.handle.net/20.500.12401/dv/59`, sendo o “dv/” o identificador do repositório.

Pode-se utilizar outros sufixos como:

- “dv.”
- “dv_”
- Entre outros

Para definir o sufixo no Dataverse, deve-se utilizar via bash o comando abaixo:

```
curl -X PUT -d 'dv/' localhost:8080/api/admin/settings/:Shoulder
```

2 INSTALAÇÃO DO HANDLE SERVER

O instalador do Handle está disponível em http://handle.net/download_hnr.html

Recomenda-se utilizar sempre a última versão.

Para instalar no servidor principal, 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
```

O Handle como servidor primário é necessário quando não existe outra instalação do prefixo, sendo esse o responsável pelo banco de dados das resoluções dos identificadores persistentes. Caso esteja instalando mais de um servidor de Handle, esse deve ser o secundário.

2.1 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

2.2 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**

2.3 CONCLUINDO A INSTALAÇÃO DO HANDLE

Após instalar o servidor primário, é necessário re-encriptar 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-encriptado pelo script mencionado acima.

2.4 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.

3 INSTALANDO O HANDLE NO DATAVERSE

Os parâmetros utilizados no JVM e no Config do Dataverse são:

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 o Dataverse, as opções são executadas por meio de requisições 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
```

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 '20.500.11959'
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 :IdentifierGenerationStyle com o comando:

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

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 listar 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 criá-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'
```

4 TESTE DO HANDLE

O Handle pode ser testado por linha de comando diretamente do servidor, isso é possível com a execução de uma arquivo de lote processado via hdl-batch

Para testar, crie uma pasta no servidor dentro de /hs com o nome /hs/cmd

```
mkdir /hs/cmd
```

Acesse esta página

```
cd /hs/cmd
```

Dentro dela crie o arquivo para ser processado como no exemplo abaixo (file: cmd.handle):

```
AUTHENTICATE PUBKEY:300:0.NA/<handle nr>
/hs/svr_1/admpriv.bin|<senha da chave>
HOME <ip>:2641:TCP
0.NA/<handle nr>
CREATE <handle nr>/pgtest
100 HS_ADMIN 86400 1110 ADMIN 200:111111111111:0.NA/<handle nr>
3 URL 86400 1110 UTF8 https://<pagina que será direcionado o
handle>
7 EMAIL 86400 1110 UTF8 dadosabertos@rnp.br
9 DESC 86400 1110 UTF8 Repositório de dados de pesquisa RNP
```

O primeiro bloco realiza a autenticação no servidor, a segunda faz a criação do registro Handle.

Exemplos de valores atribuídos

- <handle vr> 20.500.12401
- <senha da chave>: (senha utilizada para criar o servidor Handle)
- <pagina que será direcionado o handle> ex: <http://dadosderede.rnp.br>
- <ip> ex: 138.121.71.160
 - IP do servidor local do Handle

Para executar o comando digite:

```
/hs/handle-9.0.1/bin/hdl-genericbatch cmd.handle
```

Resultado do comando:

```
Batch(cmd) process started ...
Batch process prints log on stdout ...
Start Time: Wed Feb 19 21:26:50 BRT 2020
Saving bootstrap handles to: /root/.handle/bootstrap_handles
==>SUCCESS[2]: home/unhome:0.NA/20.500.11959
==>SUCCESS[3]: create:20.500.11959/pgtest
==>FAILURE[4]: create:20.500.11959/FK2/7: Error(101): HANDLE
ALREADY EXISTS
Successes/Total Entries: 1/2
Batch File Lines: 3
Finish Time: Wed Feb 19 21:26:57 BRT 2020
This batch took 6 seconds to complete at an average speed of
0.29784065524944153 operations/second
Batch process finished
```

