GERENCIAR PACOTES BAGIT v2

CRIAR PACOTES BAGITS POR DATASETS

```
sudo su - glassfish
curl -H X-Dataverse-key:$API_TOKEN
http://localhost:8080/api/datasets/:persistentId/?persistentId=doi:10.48331/SCIELODATA.RJ1HZO
export ID=19300
export API_TOKEN='YOUR API TOKEN'
curl -X POST -H "X-Dataverse-key: $API_TOKEN"
http://localhost:8080/api/admin/submitDatasetVersionToArchive/$ID/1.0
```

CRIAR PACOTES BAGITS DE TODOS OS DATASETS

Essa execução é feito uma vez por mês. Os pacotes bagits são gerados no diretório /preservacao/dataverse que é um ponto de montagem nfs em storage-scielo-232.scielo.org:/archivematica_transfersource

Antes de executar crie a pasta correspondente ao mês:

```
mkdir /preservacao/dataverse/exportacao-mes-ano/
Exemplo:
mkdir /preservacao/dataverse/exportacao-agosto-2026/
cd /preservacao/dataverse/
 mv doi-10-48331-scielodata-* exportacao-agosto-2026/
```

Depois de entrar na pasta:

```
export API_TOKEN='MEU TOKEN'
curl -X POST -H "X-Dataverse-key: $API_TOKEN"
'http://localhost:8080/api/admin/archiveAllUnarchivedDatasetVersions'
```

Uma vez que os pacotes são gerados precisamos organizar cada um dos datasets em pastas individuais. Para isso precisamos executar o script processar-dataverse.sh

```
cd /preservacao/dataverse
mv doi-10-48331-scielodata-* exportacao-mes-ano/
Exemplo:
mv doi-10-48331-scielodata-* exportacao-setembro-2026/
Depois:
sh -x processar-dataverse.sh <mes> <ano> /preservacao/dataverse/exportacao-mes-ano/
Exemplo:
./processar-dataverse.sh setembro 2026 /preservacao/dataverse/exportacao-setembro-2026/
```

Ele criará automaticamente:

```
/preservacao/dataverse/scielodata-agosto-2026/
```

e ao final, algo como:

```
scielodata-agosto-2026/
├── relatorio-processamento.txt
├── scielodata-agosto-2026-01/
│   ├── doi-10-48331-scielodata-0ppvrc/
│   │   └── doi-10-48331-scielodata-0ppvrc.v1.0/
│   │       ├── doi-10-48331-scielodata-0ppvrc.v1.0.zip
│   │       └── doi-10-48331-scielodata-0ppvrc_datacite.v1.0.xml
│   │
│   ├── doi-10-48331-scielodata-wrtkuf/
│   │   ├── doi-10-48331-scielodata-wrtkuf.v1.0/
│   │   │   ├── doi-10-48331-scielodata-wrtkuf.v1.0.zip
│   │   │   └── doi-10-48331-scielodata-wrtkuf_datacite.v1.0.xml
│   │   └── doi-10-48331-scielodata-wrtkuf.v1.1/
│   │       ├── doi-10-48331-scielodata-wrtkuf.v1.1.zip
│   │       └── doi-10-48331-scielodata-wrtkuf_datacite.v1.1.xml
│   │
│   └── ...
│
├── scielodata-agosto-2026-02/
│   └── ...
│
```

```
└── scielodata-agosto-2026-03/
    └── ...
```

TROUBLESHOOTING

Version was already submitted for archiving.

Quando você envia para criar o pacote em uma data versão, não é possível criar o mesmo sem uma intervenção. O erro que dá é:

```
{"status":"ERROR","message":"Version was already submitted for archiving."}
```

Para enviar novamente siga os passos:

```
sudo su - postgres
\c dvndb
SELECT
    id, versionnumber, minorversionnumber, versionstate,
    archivetime, archivalcopylocation, externalstatuslabel, archivenote
FROM datasetversion
WHERE dataset_id = '13150'
ORDER BY versionnumber DESC, minorversionnumber DESC;
```

Veja que o resultado foi:

```
dvndb-# ORDER BY versionnumber DESC, minorversionnumber DESC;
  id  | versionnumber | minorversionnumber | versionstate | archivetime |                archivalcopylocation                | ex
alstatuslabel | archivenote
------+---------------+--------------------+--------------+-------------+----------------------------------------------------+---
--------------+-------------
 2209 |             1 |                  0 | RELEASED     |             | {"status":"failure","message":"Bag not transferred"} |
              |
(1 row)

dvndb=#
```

Vamos agora atualizar o registro para limpar esta mensagem:

```
UPDATE datasetversion
SET archivetime = NULL,
    archivalcopylocation = NULL,
    externalstatuslabel = NULL,
    archivenote = NULL
WHERE id = 2209;
```

Resultado:

```
dvndb-# ORDER BY versionnumber DESC, minorversionnumber DESC;
  id  | versionnumber | minorversionnumber | versionstate | archivetime | archivalcopylocation | externalstatuslabel | archivenote
------+---------------+--------------------+--------------+-------------+----------------------+---------------------+-------------
 2209 |             1 |                  0 | RELEASED     |             |                      |                     |
(1 row)

dvndb=#
```

Agora vamos executar novamente o comando:

```
sudo su - glassfish
export ID=13150
export API_TOKEN='minha api'
curl -X POST -H "X-Dataverse-key: $API_TOKEN"
http://localhost:8080/api/admin/submitDatasetVersionToArchive/$ID/1.0
```

Outra forma para limpar é:

```
export API_TOKEN=xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx
export SERVER_URL=https://demo.dataverse.org
export PERSISTENT_IDENTIFIER=doi:10.5072/FK2/7U7YBV
export VERSION=1.0

curl -H "X-Dataverse-key: $API_TOKEN" -X DELETE
"$SERVER_URL/api/datasets/:persistentId/$VERSION/archivalStatus?persistentId=$PERSISTEN
T_IDENTIFIER"
```

Link: https://guides.dataverse.org/en/latest/api/native-api.html#delete-the-archival-status-of-a-dataset-by-version

Erro {"status":"failure","message":"Bag not transferred"}

Este erro aconteceu quando eu tentei criar os bagit e o usuário não tinha permissão para escrever no diretório /preservacao/dataverse. Uma vez que deu erro alguns pacotes ficaram com o status : {"status":"failure","message":"Bag not transferred"}

1 Primeiro: verifique antes de atualizar (modo seguro)

Antes de alterar qualquer coisa, rode este SELECT para confirmar exatamente **quais registros serão afetados**:

```
SELECT
    id, dataset_id, versionnumber, versionstate, archivetime, archivalcopylocation
FROM datasetversion
WHERE archivalcopylocation::text LIKE '%"Bag not transferred"%';
```

Isso lista **apenas as versões com erro de arquivamento**, garantindo que você não vai limpar dados corretos.

2 Atualização segura (modo em lote)

Depois de confirmar, execute:

```
UPDATE datasetversion
SET
    archivetime = NULL,
    archivalcopylocation = NULL,
    externalstatuslabel = NULL,
    archivenote = NULL
WHERE archivalcopylocation::text LIKE '%"Bag not transferred"%';
```

Esse comando:

- **limpa os campos de arquivamento** que ficaram marcados com erro,
- sem afetar versões bem-sucedidas ou DRAFTs normais.

Feito isso, faço novamente o processo de criação dos pacotes.

Revision #4
Created 7 August 2026 14:10:49 by Rondineli G. Saad
Updated 6 October 2026 20:08:34 by Rondineli G. Saad