Sharing Data with The Harvard Dataverse Repository

Danny Brooke,
Program Manager, Product Development

Sonia Barbosa
Manager of Data Curation, Harvard Dataverse Repository
Manager of the Murray Research Archive

Institute for Quantitative Social Science, Harvard University

Agenda

- Dataverse Project Background
- Dataverse Project APIs
- Demo
- Questions

enabling better, bigger, faster, more collaborative science

enabling better, bigger, faster, more collaborative science

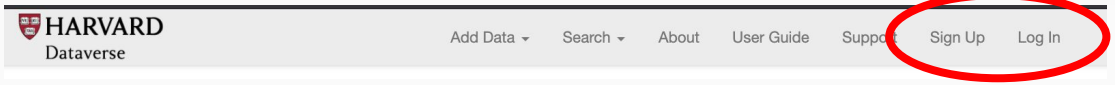
Dataverse: Open Source Research Data Repository Software

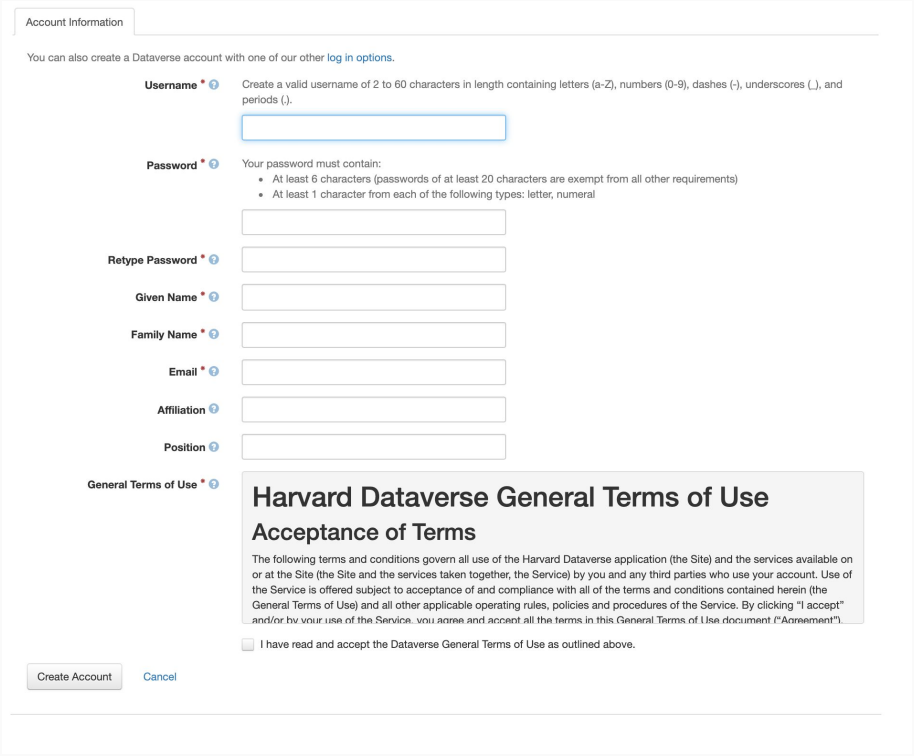
Dataverse Installations Worldwide

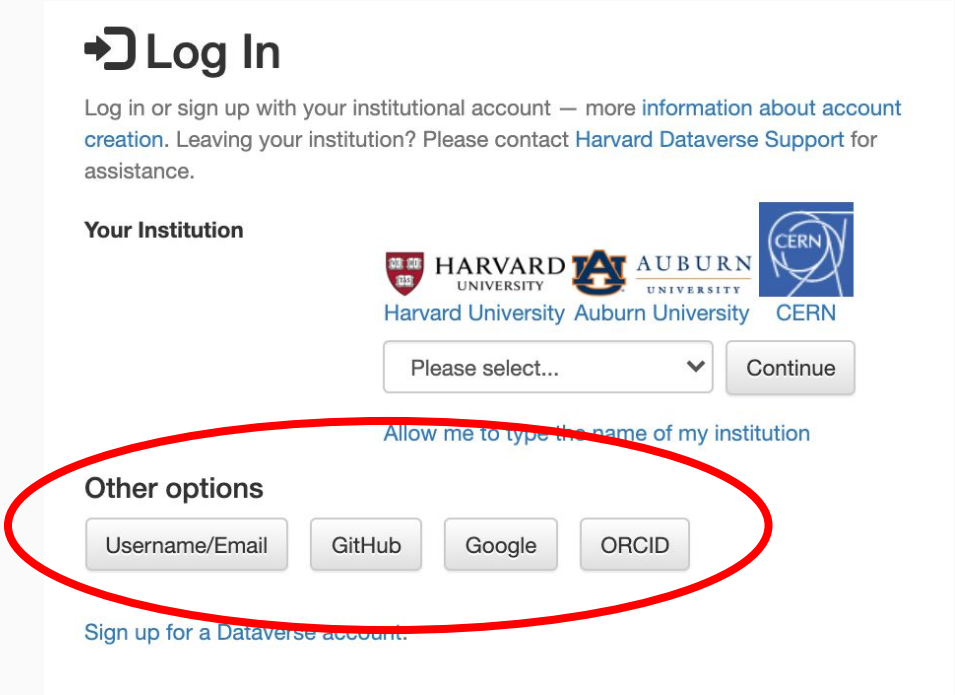
Depositing your Data Into the Harvard Dataverse Repository (in a few easy steps!)

- Create an Account (in the UI)
- Create an API Token (in the UI)
- Create a Dataverse collection (optional, but recommended!)
- Create a Dataset
- Add File(s)
- Publish your Dataset (and Dataverse collection)

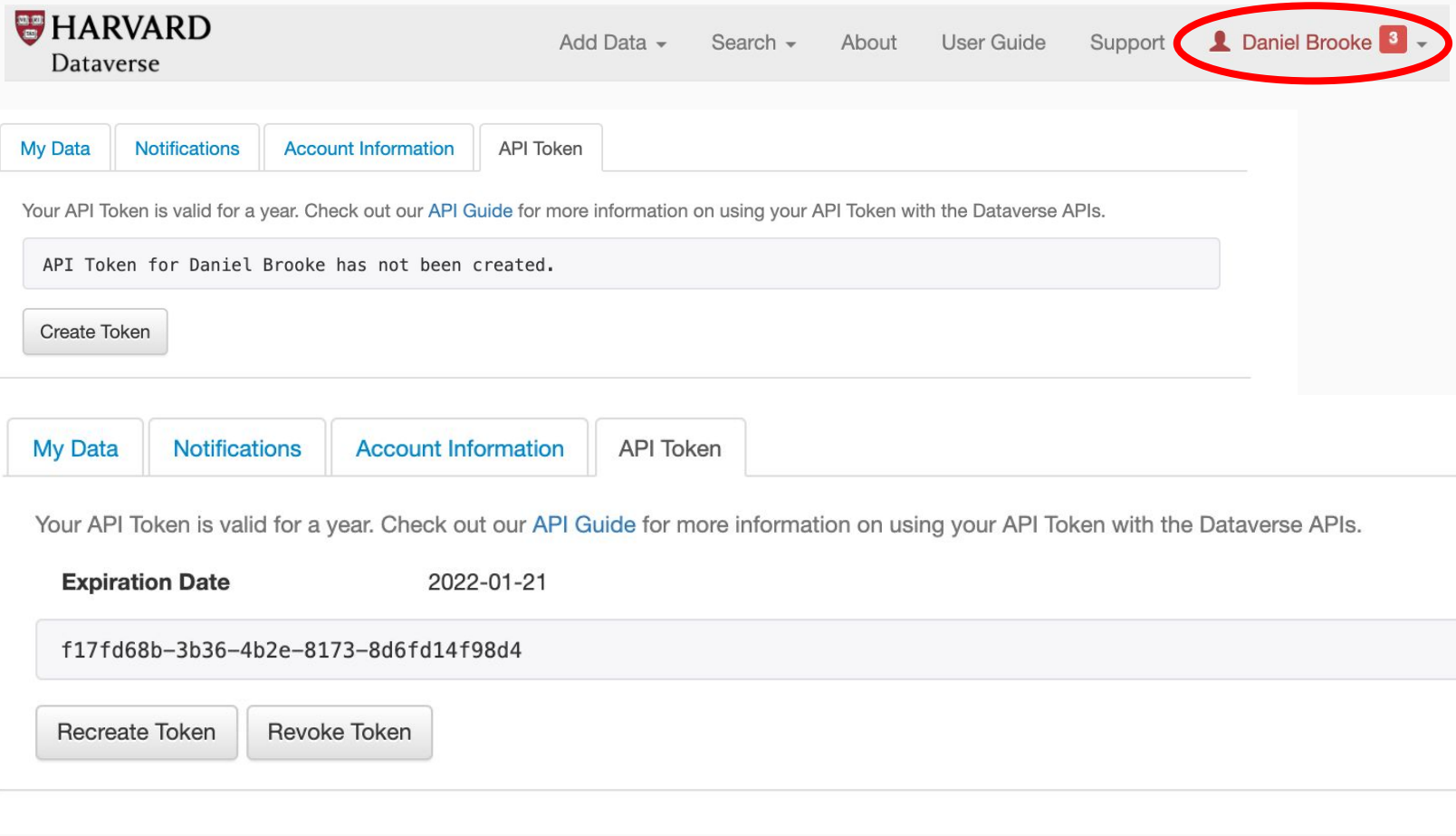
Step 1: Create an Account



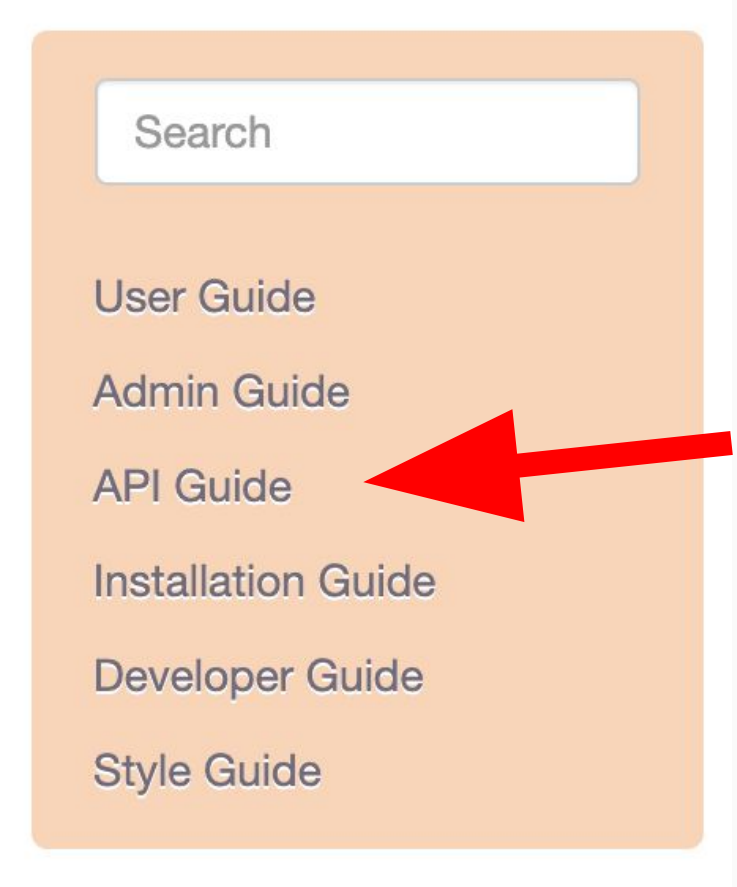




Step 2: Create an API Token



Dataverse Software API Guides



guides.dataverse.org

Step 3a: Create a Dataverse Collection: Json File

Download **dataverse-complete.json** file and modify it to suit your needs.

```
{
  "name": "Scientific Research",
  "alias": "science",
  "dataverseContacts": [
    {
      "contactEmail": "pi@example.edu"
    },
    {
      "contactEmail": "student@example.edu"
    }
  ],
  "affiliation": "Scientific Research University",
  "description": "We do all the science.",
  "dataverseType": "LABORATORY"
}
```

https://guides.dataverse.org/en/latest/api/native-api.html#id20

Step 3b: Create a Dataverse Collection: Execute a Curl Command

```
curl -H X-Dataverse-key:xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx -X POST https://demo.dataverse.org/api/dataverses/root --upload-file dataverse-complete.json
```

https://guides.dataverse.org/en/latest/api/native-api.html#id20

Step 4a: Create a Dataset: Json File

As a starting point, you can download **`dataset-finch1.json`** and modify it to meet your needs.

```
{
  "datasetVersion": {
    "metadataBlocks": {
      "citation": {
        "fields": [
          {
            "value": "Darwin's Finches",
            "typeClass": "primitive",
            "multiple": false,
            "typeName": "title"
          },
          {
            "value": [
              {
                "authorName": {
                  "value": "Finch, Fiona",
                  "typeClass": "primitive",
                  "multiple": false,
                  "typeName": "authorName"
                },
                "authorAffiliation": {
                  "value": "Birds Inc.",
                  "typeClass": "primitive",
                  "multiple": false,
                  "typeName": "authorAffiliation"
                }
              }
            ],
            "typeClass": "compound",
            "multiple": true,
            "typeName": "author"
```

https://guides.dataverse.org/en/latest/api/native-api.html#id37

Step 4b: Create a Dataset: Execute Curl Command, Note the PID

```
curl -H "X-Dataverse-key:xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx" -X POST "https://demo.dataverse.org/api/dataverses/root/datasets" --upload-file "dataset-finch1.json"
```

You should expect an HTTP 200 (“OK”) response and JSON indicating the database ID and Persistent ID (PID such as DOI or Handle) that has been assigned to your newly created dataset.

https://guides.dataverse.org/en/latest/api/native-api.html#id37

Step 5: Add Files: Execute Curl Command, Repeat as Needed

```
curl -H X-Dataverse-key:xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx -X POST -F file=@data.tsv
-F 'jsonData={"description":"My description.","directoryLabel":"data/subdir1","categor
ies":["Data"], "restrict":"false"}' "https://demo.dataverse.org/api/datasets/:persiste
ntId/add?persistentId=doi:10.5072/FK2/J8SJZB"
```

- A description of the file
- The “File Path” of the file, indicating which folder the file should be uploaded to within the dataset
- Whether or not the file is restricted

```
curl -H X-Dataverse-key:xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx -X POST -F file=@data.tsv
-F  "https://demo.dataverse.org/api/datasets/:persistentId/add?persistentId=doi:10.507
2/FK2/J8SJZB"
```

https://guides.dataverse.org/en/latest/api/native-api.html#id63

Step 6a: Publish Dataverse Collection

```
curl -H X-Dataverse-key:xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx -X POST https://demo.data
verse.org/api/dataverses/root/actions/:publish
```

https://guides.dataverse.org/en/latest/api/native-api.html#id40

Step 6b: Publish Dataset

```
curl -H "X-Dataverse-key: xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx" -X POST "https://demo.dataverse.org/api/datasets/:persistentId/actions/:publish?persistentId=doi:10.5072/FK2/J8SJZB&type=major"
```

https://guides.dataverse.org/en/latest/api/native-api.html#id53

Your Dataset has been Published!

Recap!

- Create an Account (in the UI) ✔️
- Create an API Token (in the UI) ✔️
- Create a Dataverse collection (optional, but recommended!) ✔️
- Create a Dataset ✔️
- Add File(s) ✔️
- Publish your Dataset (and Dataverse collection) ✔️

A few last things before the Demo

- Guides.dataverse.org is your friend!
- We've only scratched the surface of what Dataverse Software APIs can do
- We try to make everything you can do in the UI available through the APIs
- SWORD API is a standard implementation - we focused on the Native API today

Follow this link to the demo website!

https://demo.dataverse.org/dataverse/datafest2021

Thank You!

The Dataverse Project: dataverse.org

Harvard Dataverse: dataverse.harvard.edu

Test the features for Dataverse at: demo.dataverse.org

Contact support@dataverse.harvard.edu with questions, or ask the Dataverse Project Google Group:

https://groups.google.com/g/dataverse-community