Instructions for access to the Exposome Data Warehouse (EDW)

EDW can be accessed by either loading our PostgreSQL dump file into your private environment or using the EDW API. In this document we will show how to access data in both ways.

Lots of tutorials and further information are available through our github repo:

https://github.com/hms-dbmi/exposomeDW_public

EDW access via PostgreSQL dump file

Data Download

All EDW files including are available as a compressed binary PostgreSQL dump file (edwDump.bak). Note: the file was generated using pg_dump utility specifying the “custom” compression format (-Fc flag).

The file can be downloaded via

Data Dryad
https://doi.org/10.5061/dryad.fv93qg5

Amazon S3
https://s3.amazonaws.com/edwpostgres/edwDump.bak

Data Restore

Download file into your private server environment that is connected to a PostgreSQL database. Using the pg_restore utility the user can simply use the command

pg_restore -Fc -h ‘hostname’ -p ‘port number’ -U ‘user id’ edwDump.bak

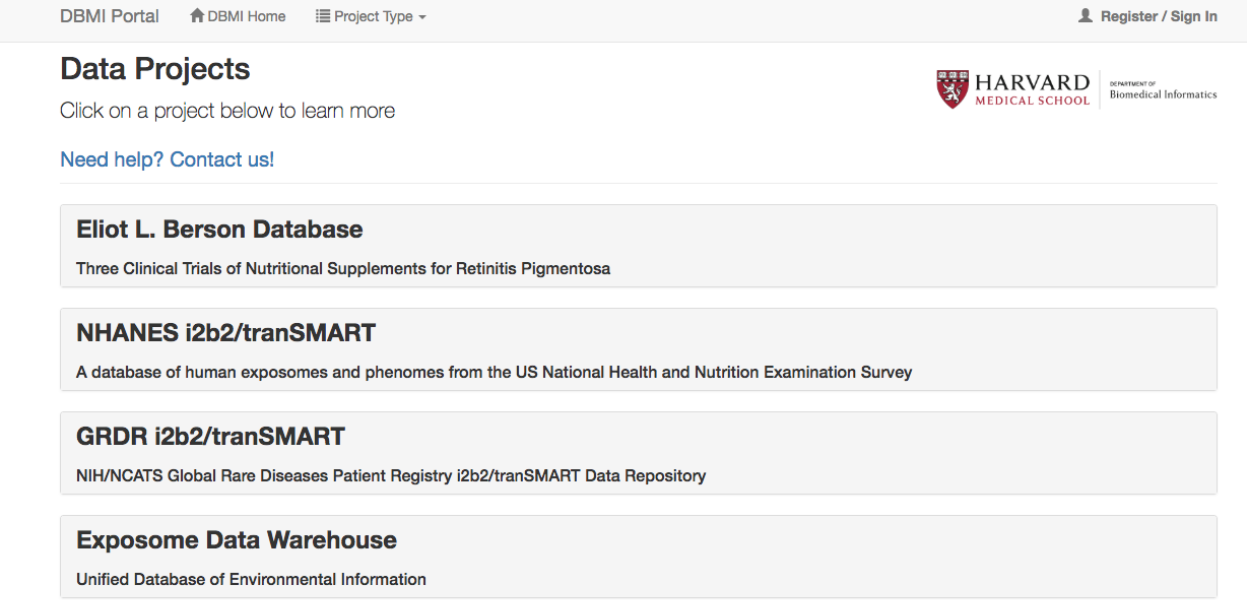
(replace quoted values with your actual values)

EDW access via API

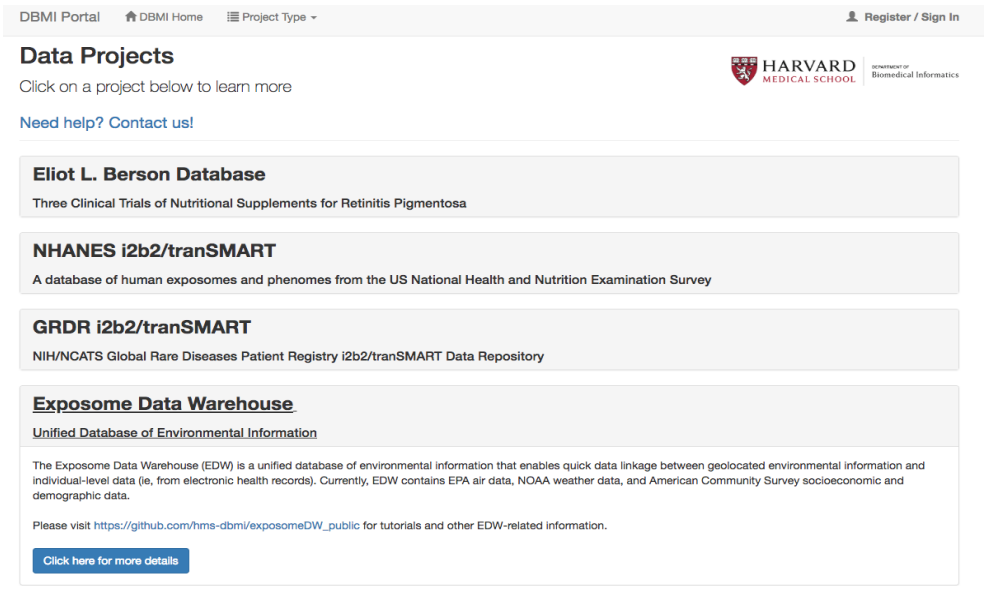
These instructions will give you step by step instructions for obtaining a JWT authentication token for the API. For further details on how to use the API go to the github repo and look for the API tutorial in the “Jupyter_Notebook_Demos” subdirectory.

Accessing the JWT Token

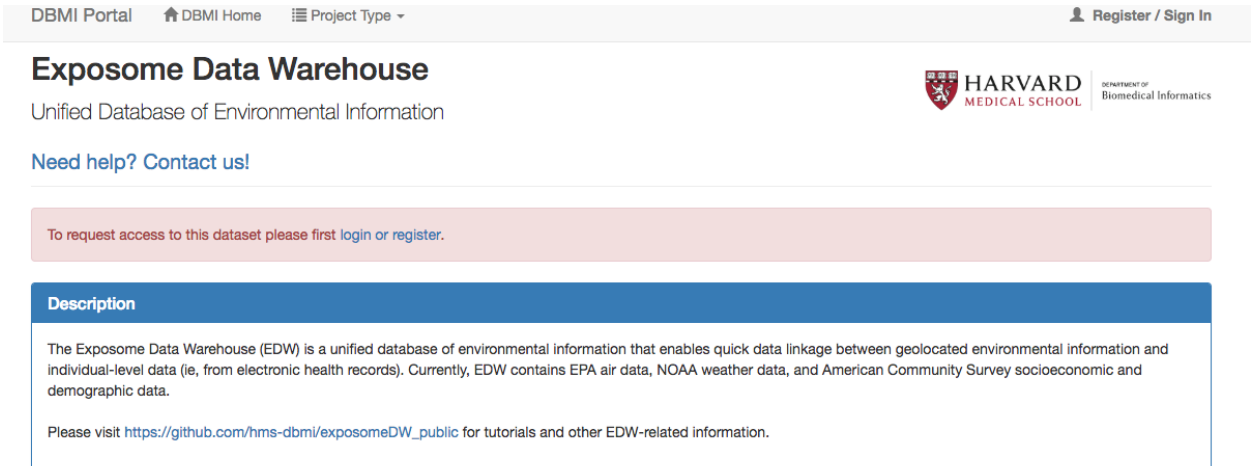
1. Go to https://portal.dbmi.hms.harvard.edu



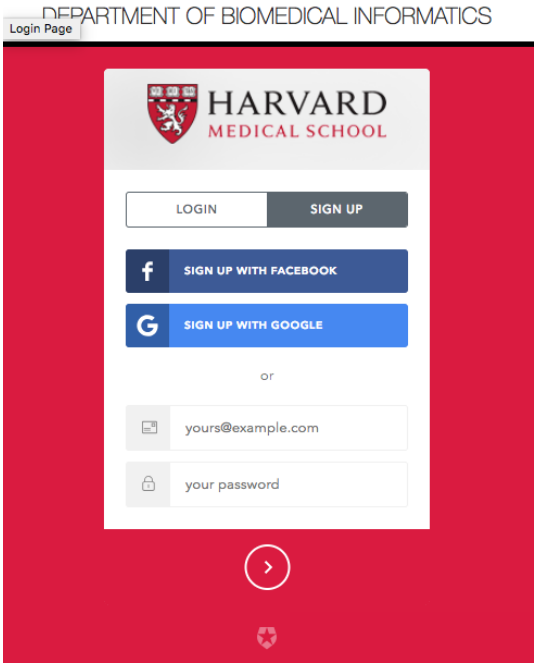
2. Click on *‘Exposome Data Warehouse’* then click on *‘Click here for more details’*



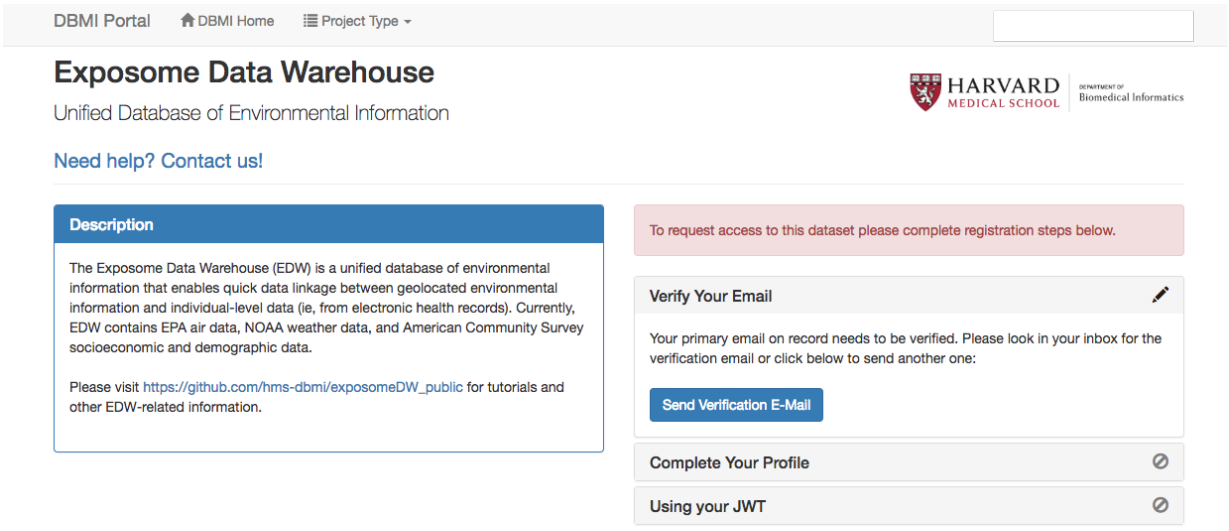
3. You will then be redirected to the “Exposome Data Warehouse” front page, there you will click on the “login or register” link.



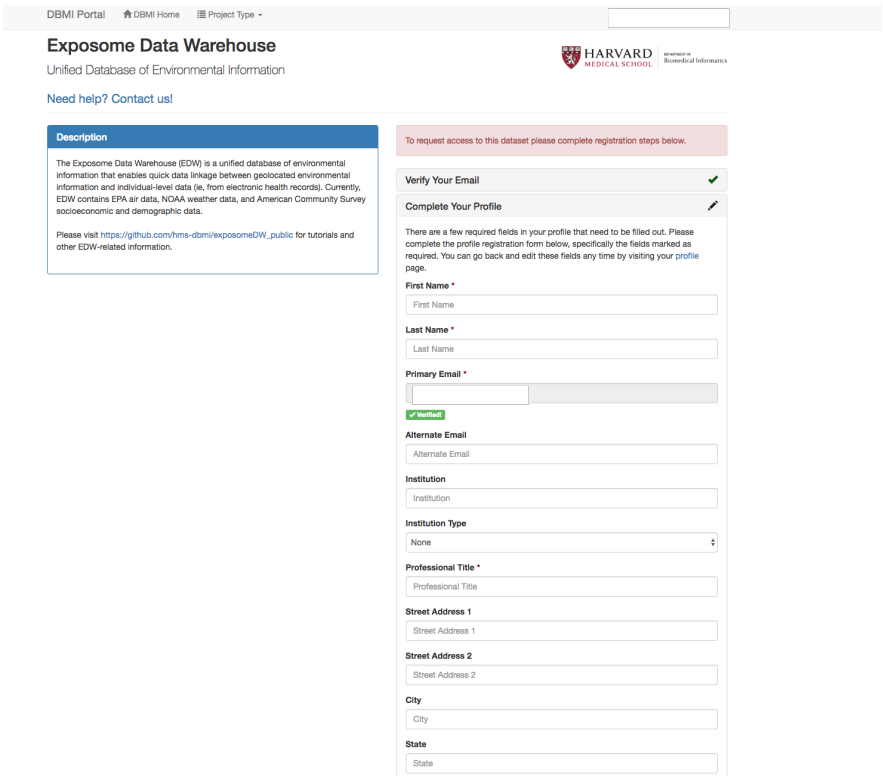
4. You will be redirected to a registration page where you can use sign up using your Google or Facebook login or via an email address. (If you use a Harvard or selected other institutional email address then page will be redirected to your institutional login page).

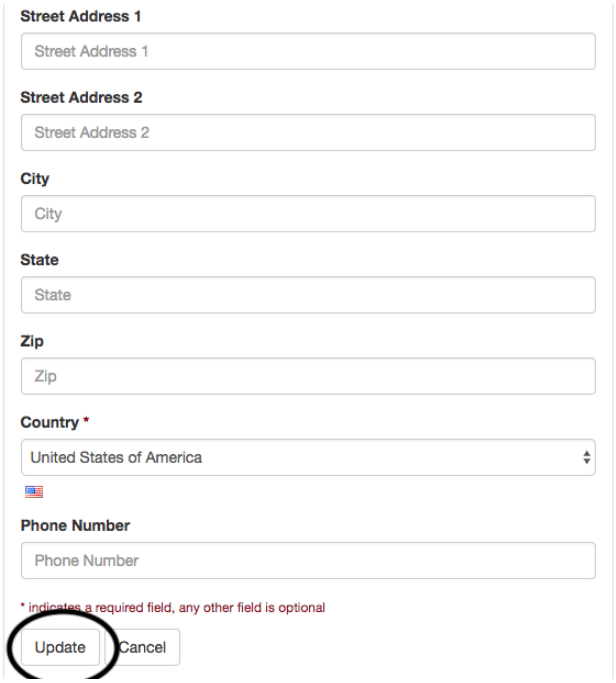


5. Follow the instructions for signing up and then you will come to the landing page below. Here you will verify your email.

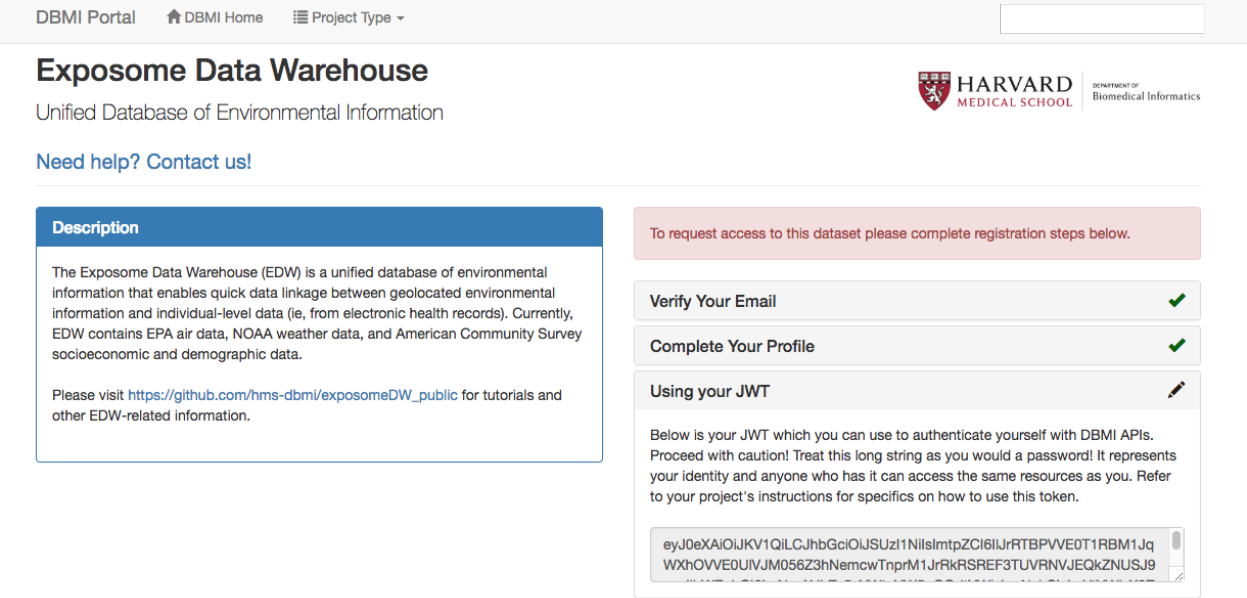


5. After you have verified your email, you will log back in and fill out the profile information and then click the update button.





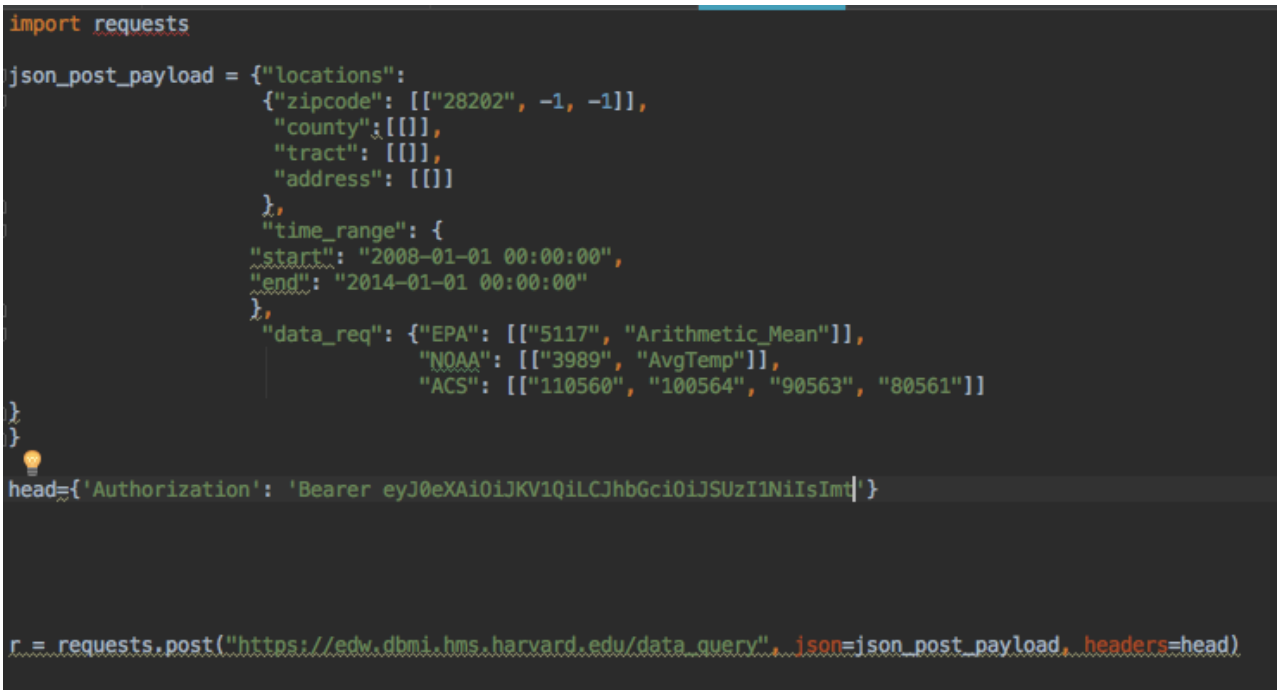
6. After you have filled out your profile information, you will see a dialog box containing your JWT token. You will use this JWT token in order to access the API.



7. Once you have your JWT token you can use it in the header of your REST API call e.g. if your token is eyJ0eXAiOiJKV1QiLCJhbGciOiJSUzI1NiIsImt

then your API call will have an **Authorization** key of the form (see Python code below)

Bearer eyJ0eXAiOiJKV1QiLCJhbGciOiJSUzI1NiIsImt



```
import requests

json_post_payload = {"locations":
                     {"zipcode": [["28202", -1, -1]],
                      "county":[[]],
                      "tract": [[]],
                      "address": [[]]
                     },
                     "time_range": {
                    "start": "2008-01-01 00:00:00",
                    "end": "2014-01-01 00:00:00"
                    },
                     "data_req": {"EPA": [["5117", "Arithmetic_Mean"]],
                                  "NOAA": [["3989", "AvgTemp"]],
                                  "ACS": [["110560", "100564", "90563", "80561"]]
}
}

head={'Authorization': 'Bearer eyJ0eXAiOiJKV1QiLCJhbGciOiJSUzI1NiIsImt'}



r = requests.post("https://edw.dbmi.hms.harvard.edu/data_query", json=json_post_payload, headers=head)
```