

Title : **Replicating FDP for Amsterdam UMC in myDRE environment**

Created by : **Priyanka Ojha** [Priyanka Ojha \(0000-0002-6844-6493\) - ORCID](#)

Created on : 2024/10/08

1. Introduction

I had the task to try out the reference implementation of FDP and also try to import data into it.

This involved working on 2 projects namely FDP and SeMPyRO

I started with reproducing FDP on myDRE environment.

My setup consisted of myDRE which is a managed service with Azure under the hood. I used a ubuntu22 machine as a base and installed docker on it and followed the steps mentioned below in that order :

1. To install docker on <https://mydre.org/> , I followed these instructions :

[Installing docker on Ubuntu 22 \(mydre.org\)](#)

<https://docs.docker.com/engine/install/ubuntu/#install-using-the-repository>

[Ubuntu | Docker Docs](#)

The Ubuntu release was Jammy in my case.

2. Then I installed Docker compose via this command : `sudo apt-get install docker-compose-plugin`
3. Check the version to ensure Docker compose is installed successfully : `docker-compose version`

```
mydre.org@dws2156srv3:~$ docker-compose version
docker-compose version 1.29.2, build unknown
docker-py version: 5.0.3
CPython version: 3.10.12
OpenSSL version: OpenSSL 3.0.2 15 Mar 2022
```

4. Create a folder fdp in root directory : `sudo mkdir fdp`
5. Give the read write and execute rights to it : `sudo chmod a+rwX fdp`
6. There I copied 2 files (docker-compose.yml and application.yml) I had created with the help of [Local Deployment – FAIR Data Point 1.16 documentation](#) in the fdp folder.
7. Go to the FDP folder via this command : `cd fdp`
8. Give this command to start the docker container : `docker-compose up -d`

9. To check if the containers are started correctly : you can use the following commands to check it : *docker ps*. It lists all the containers that are started. There should be 4 services running like below :

```
./mnt/fdp$ docker ps
```

CONTAINER ID	PORTS	IMAGE	NAMES	COMMAND	CREATED	STATUS
a29e7589199c		mongo:4.0.12		"docker-entrypoint.s..."	19 minutes ago	Up 19 min
utes 0.0.0.0:27017->27017/tcp			fdp-mongo-1			
c82572db3213		fairdata/fairdatapoint:1.16		"/bin/sh -c 'java -j..."	19 minutes ago	Up 19 min
utes			fdp-fdp-1			
618d8255e74e		metaphacts/blazegraph-basic:2.2.0-20160908.003514-6		"/entrypoint.bash"	19 minutes ago	Up 19 min
utes 0.0.0.0:8080->8080/tcp, :::8080->8080/tcp			fdp-blazegraph-1			
28069eb2a841		fairdata/fairdatapoint-client:1.16		"docker-entrypoint..."	19 minutes ago	Up 19 min
utes 0.0.0.0:80->80/tcp, :::80->80/tcp			fdp-fdp-client-1			

```
root@ubuntu:/mnt/fdp$
```

10. If any of the services is not started, their might be an issue. Most probably, its related to port 80 being occupied. So you can also check for it before starting the docker container and kill the process and then can resume again from step # 8 above.

11. Command to check if port 80 is occupied is :

```
fdp$ sudo lsof -i:80
```

12. Command to kill the process occupying port 80 is :

```
/fdp$ sudo kill 648
```

Do this at your own discretion.

13. You can also check for the similar issue here : [Issue accessing FAIR Data Point after first deployment · Issue #94 · FAIRDataTeam/FAIRDataPoint \(github.com\)](#)

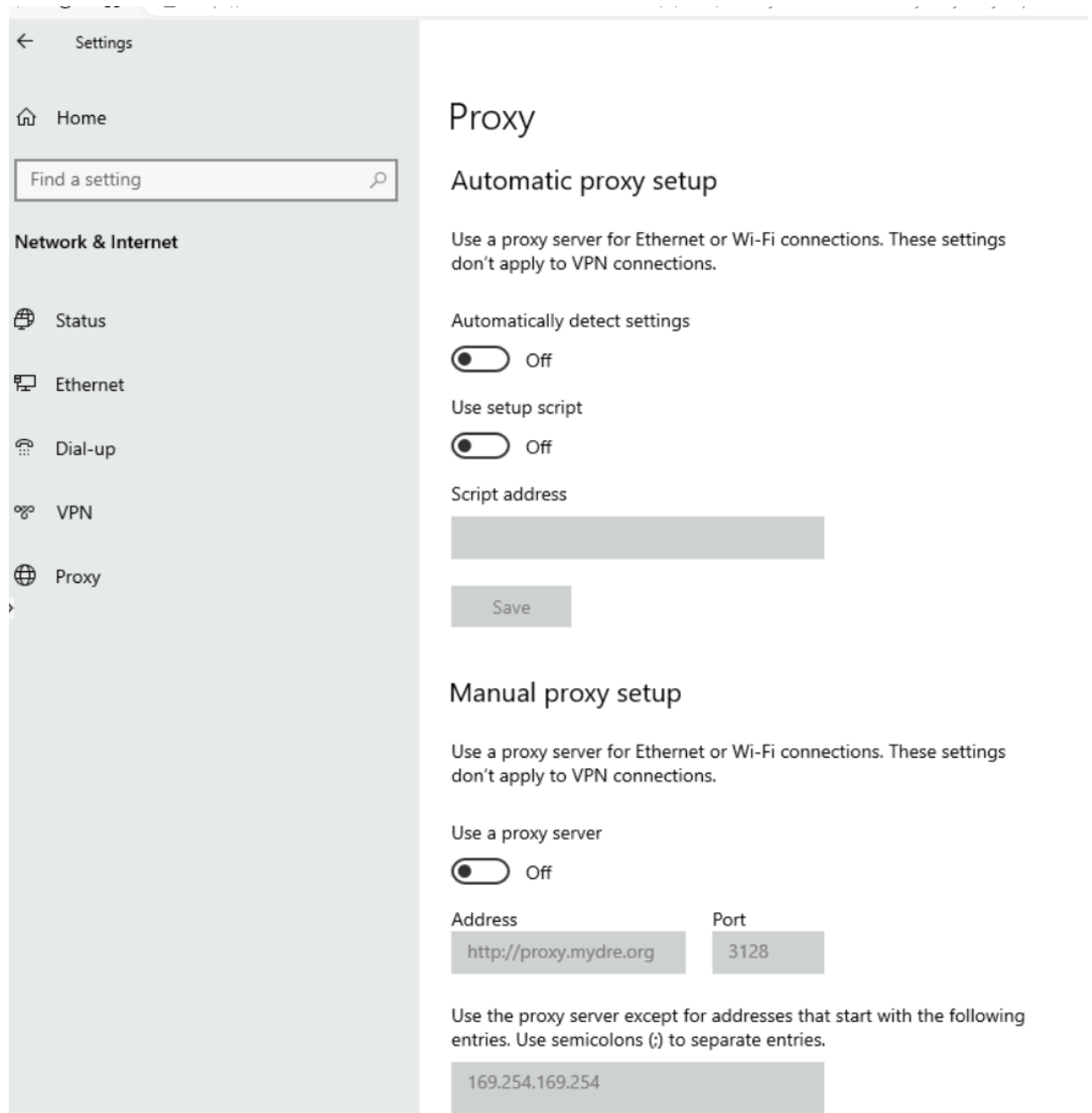
14. If all the services are running correctly, you should be able to access the FDP via this command : *curl http://localhost* and see the following output below.

```
root@ubuntu:/mnt/fdp$ curl http://localhost
{
  "@context": "https://w3id.org/fdp/fdp-o#FAIRDataPoint",
  "@type": "DataCatalog",
  "dcterms": {
    "accessRights": "http://localhost:8080#accessRights",
    "language": "http://id.loc.gov/vocabulary/iso639-1/en",
    "license": "http://rdflib.org/2011/07-15/rdflib-cc-by-nc-nd3.0",
    "repositoryIdentifier": "http://localhost:8080#identifier",
    "metadataIdentifier": "http://localhost:8080#identifier",
    "metadataIssued": "2024-05-13T11:12:18.986Z",
    "metadataModified": "2024-05-13T11:12:18.987Z",
    "schemaLabel": "My FAIR Data Point",
    "description": "Duis pellentesque, nunc a fringilla varius, magna dui porta quam, nec ultricies augue, vel varius leo viverra at. Donec scelerisque id ipsum id semper. Maecenas facilisis augue vel justo molestus vitae scelerisque. Nullam fermentum lectus nisi, id vulputate nisi congue nec. Morbi fermentum justo at juvenatis. Cras ullamcorper, justo vitae feugiat commodo, orci metus suscipit purus, quis sagittis turpis ante varius. Duis finibus dapibus ex, a hendrerit mauris efficitur at.",
    "hasVersion": "1.0",
    "publisher": "http://localhost:8080#publisher"
  }
}
```

15. If a different message is shown, you might have to wait a bit as some services take a bit of time and try again . If that doesn't help and you get a different message then shown above, then something is wrong and you need to check the logs via this command *docker-compose logs -f >> fdplogs.txt*
16. You can now via a browser from a windows machine in the same workspace access the URL via the IP address of the Linux host PC.

17. Make sure to white list the IP address of the Linux machine.
18. Also on the Windows machine in Azure workspace, disable the Proxy by following these steps :

Select the **Start** button, then select **Settings > Network & Internet > Proxy**



Once that is done, you should be able to access the FDP via the browser :

19. Open the Edge or chrome based browser, as it doesn't work on Internet Explorer. Type the IP address of Linux PC and the FDP website should be accessible as shown below:

20. Now, you can also import the Shacl File by following steps mentioned here :

<https://github.com/Health-RI/health-ri-metadata/issues/99#issuecomment-2176244905>

21. Now you can go to this [project](#) and try to run the notebook mentioned here and try importing the data to your FDP server that you configured above.

Conclusion :

FDP runs in myDRE environment, so the current task is complete. Having said that there are still a lot of open issues and I need to replicate the same with productions settings in a production environment. So I will post an update once I know more .

References

1. [FAIR Data Point | FAIR Data Points can be used to describe your data sets in a FAIR way, using standard metadata and make them available through simple WWW protocols.](#)
2. [FAIRDataTeam/FAIRDataPoint \(github.com\)](#)
3. [FAIR Data Point Reference Implementation Documentation – FAIR Data Point 1.16 documentation](#)
4. [Health-RI/SeMPyRO: pydantic data models for DCAT-AP v3 \(github.com\)](#)
5. [AmsterdamUMC/FDP_AMS at myDRE \(github.com\)](#)