Installation Guide for Toolbox App on the Local PC (Laptop/Desktop)

(Linux & Windows only)

Table of Contents

This installation document was created by Erdal Ayan, Software Developer for the Toolbox development in scope of the DFG project GND for Cultural Data – *GND4C* in Germany. The GND4C-Toolbox allows to query for person and building objects from GLAM metadata in the GND and score the response for likeliness of a match. In this document the reader will find the installation process of the FastAPI/Streamlit-based Toolbox app in a local machine (only with the Linux or Windows OS) step by step.

Please strictly follow the described steps without skipping any step and in case of error registration please contact to me via my email (erdal_ayan@yahoo.com)

Linux Installation Guide: (If you are a Windows user, please scroll down.)

Step 1: Install PostgreSQL:

Start your Terminal.
Update package list and install PostgreSQL via following commands:

sudo apt update
sudo apt install postgresql postgresql-contrib

PostgreSQL should start automatically. If not, start it via following command:

sudo service postgresql start

Step 2: Install PGAdmin:

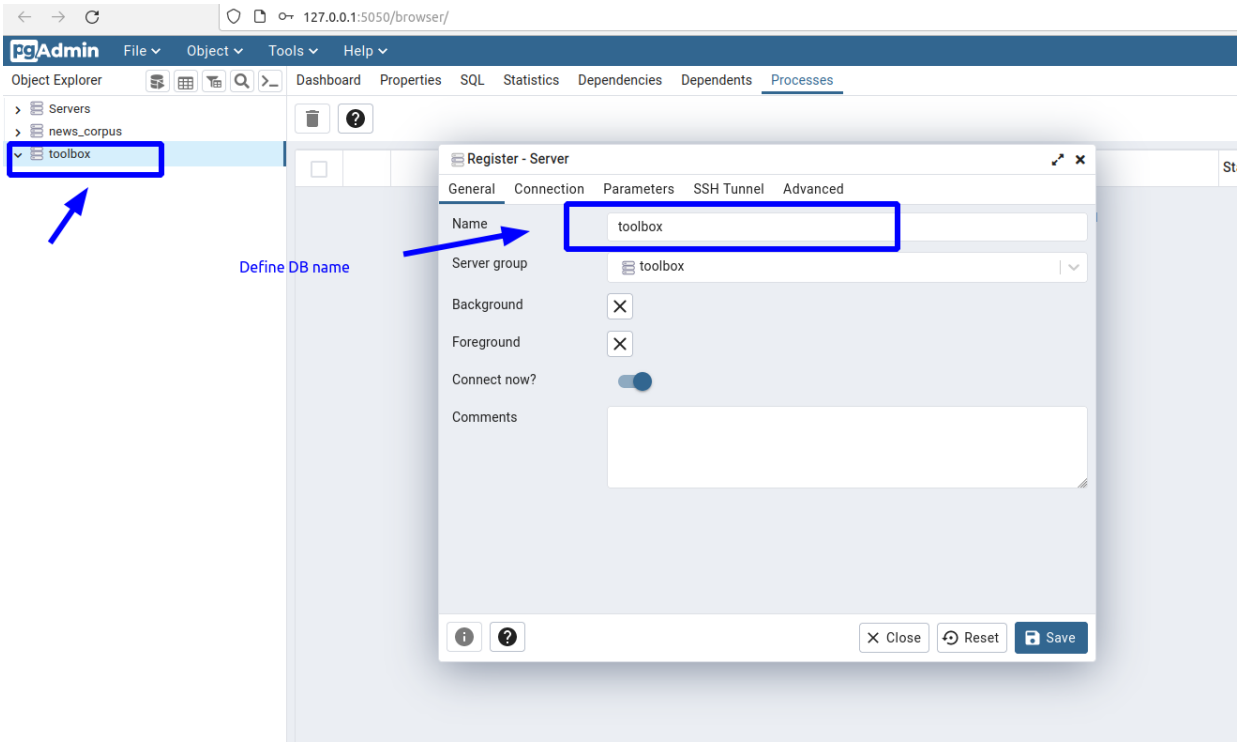
Download the latest version of PGAdmin from the official website:

https://www.pgadmin.org/download/

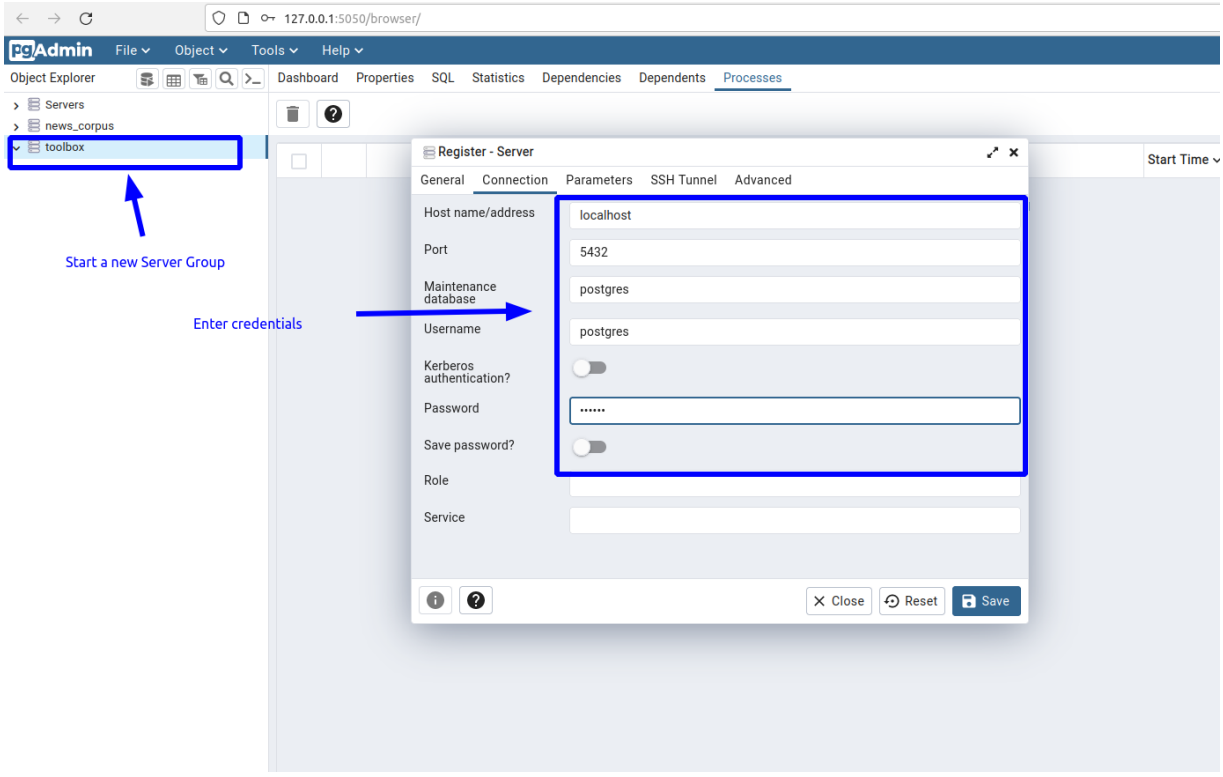
Follow the installation instructions here: https://computingforgeeks.com/how-to-install-pgadmin-4-on-ubuntu/?expand_article=1.

Step 3: Create a PostgreSQL database and user:

Launch your PGAdmin and log in using the PostgreSQL superuser credentials.
Create a new server group (toolbox), user (postgres), password (postgres) and a new database (postgres) using PGAdmin. After this configuration, the settings should look like the screenshots below.



Screenshot 1



Screenshot 2

Step 4: Install Conda:

Download the Conda installer for Linux from the official website:

https://docs.conda.io/projects/conda/en/latest/user-guide/install/linux.html

Follow the installation instructions.

Step 5: Install VSCode:

Download the VSCode installer for Linux from the official website:

https://code.visualstudio.com/docs/setup/linux

Follow the installation instructions.

Step 6: Create a Conda virtual environment for Streamlit:

Open a Terminal Enter the command:

conda create --name streamlit python=3.9

Activate the environment in your Terminal:

conda activate streamlit

Step 7: Install Streamlit and psycopg2 within the virtual environment:

Install Streamlit and psycopg2 using pip:

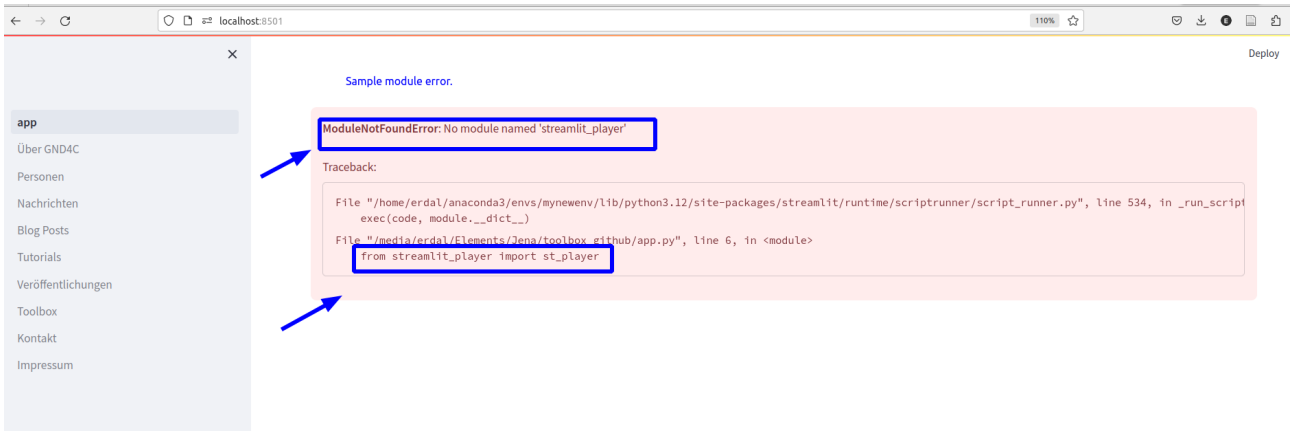
Open a terminal and enter the command:

pip install streamlit psycopg2

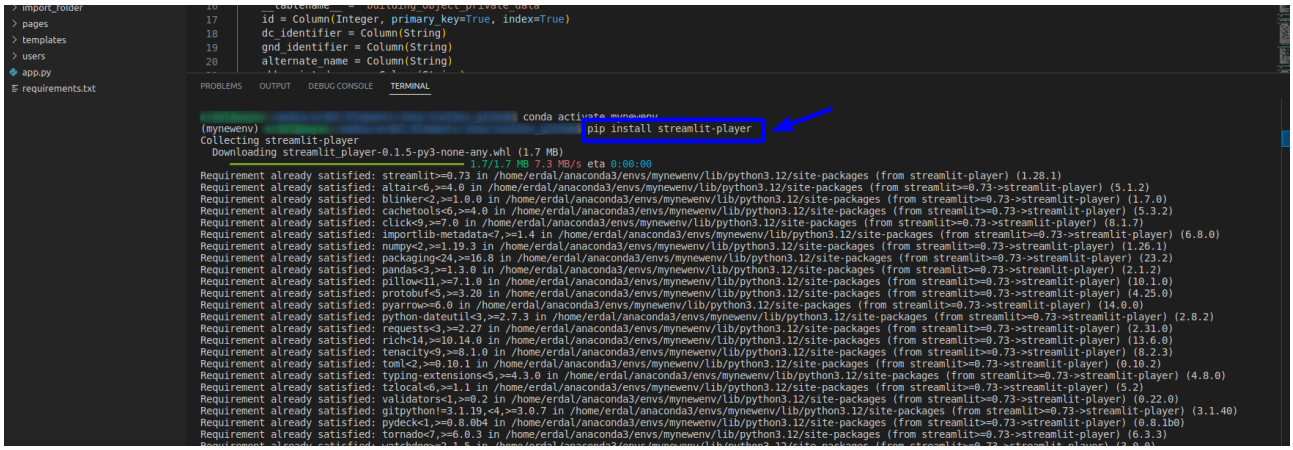
Also install other requirements via the following command:

pip install -r requirements.txt

In case there are still errors with the uninstalled Python libraries/modules as in Screenshot 3 below, please attempt to install the module manually by following the instructions[1] in screenshot 4.



Screenshot 3: Sample Module Error Registration



Screenshot 4: Sample solution for Module error registration

Step 8: Run FastApi Backend:

If you have not switched off your VSCode you do not have to restart your Terminal, otherwise start your VSCode and open the folder of the app in your VSCode. Start a “New Terminal”. Enter the following commands in your Terminal:

conda activate streamlit

uvicorn main:app --reload

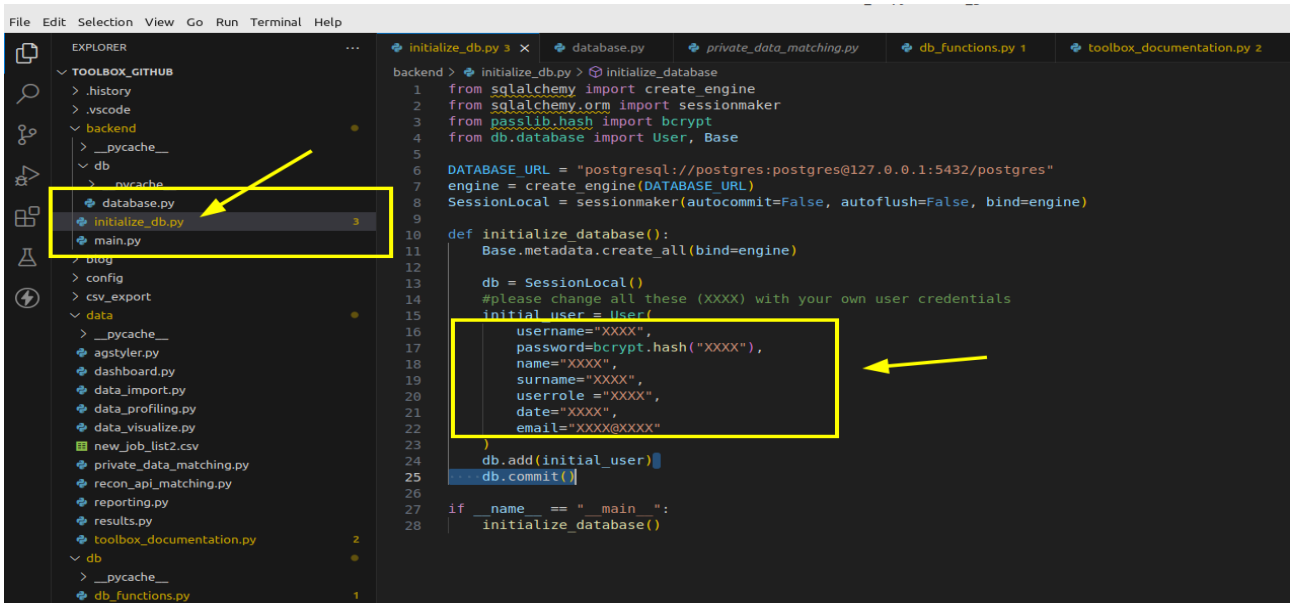
1 Please pay attention: Each module may require different installation command (pip install [name-of-library]) so please first check the installation command with the PyPi site of the library (look for an example here: https://pypi.org/project/streamlit-player/)

Step 9: Create a User in the Database:

Start your VSCode and open the folder of the app in your VSCode. Start a “New Terminal”. Enter the following commands in your Terminal:

cd backend

#Update the db connection and user credentials in the file “initialize_db.py” and “save” the file as in the screenshot below



Screenshot 5: Updating “initialize_db.py”

#Run the following command in your Terminal

python initialize_db.py

Step 10: Run Toolbox (Streamlit) app:

Start your VSCode and open the folder of the app in your VSCode. Start a “New Terminal”. Enter the following commands in your Terminal:

conda activate streamlit

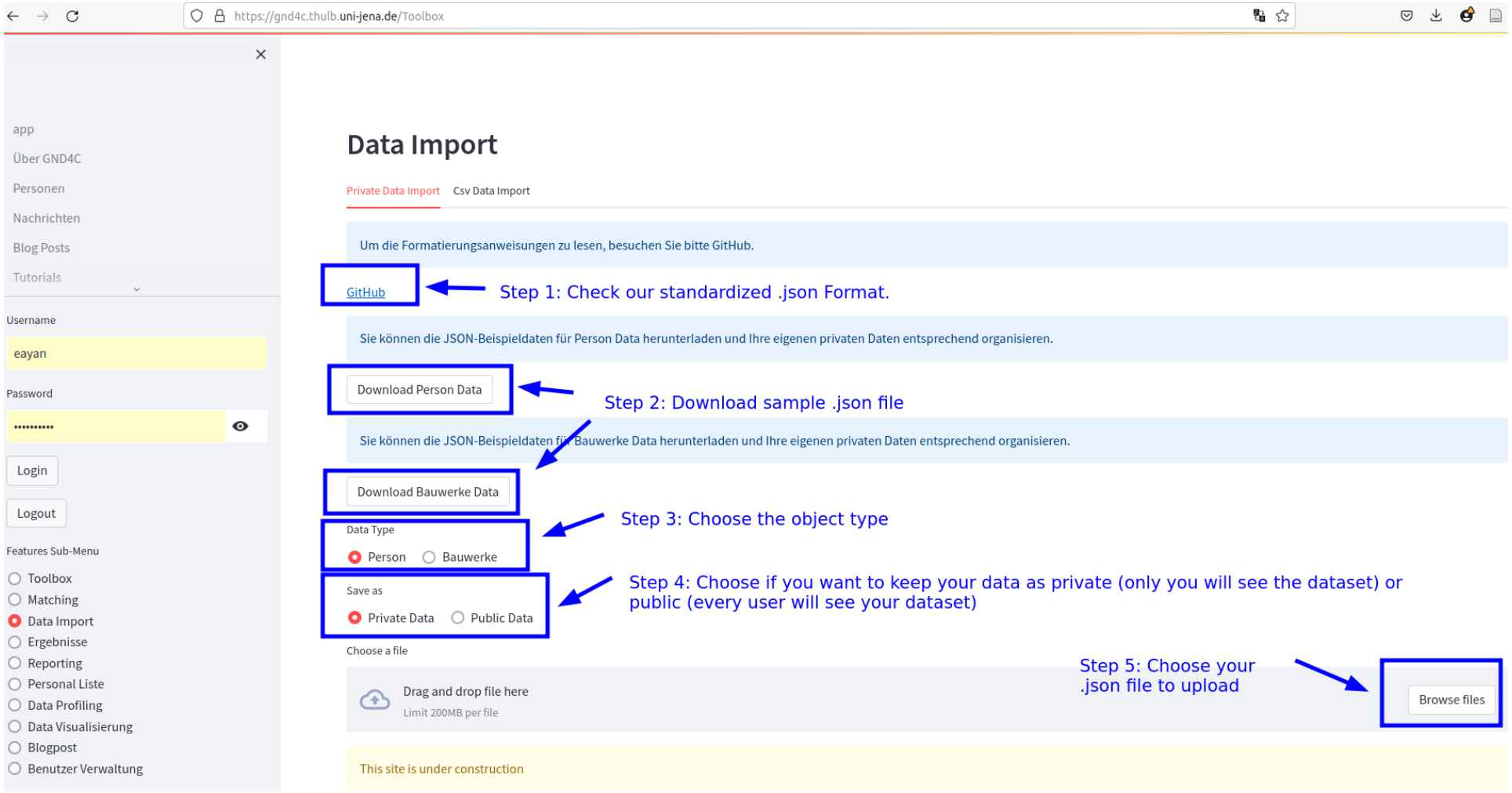
streamlit run app.py

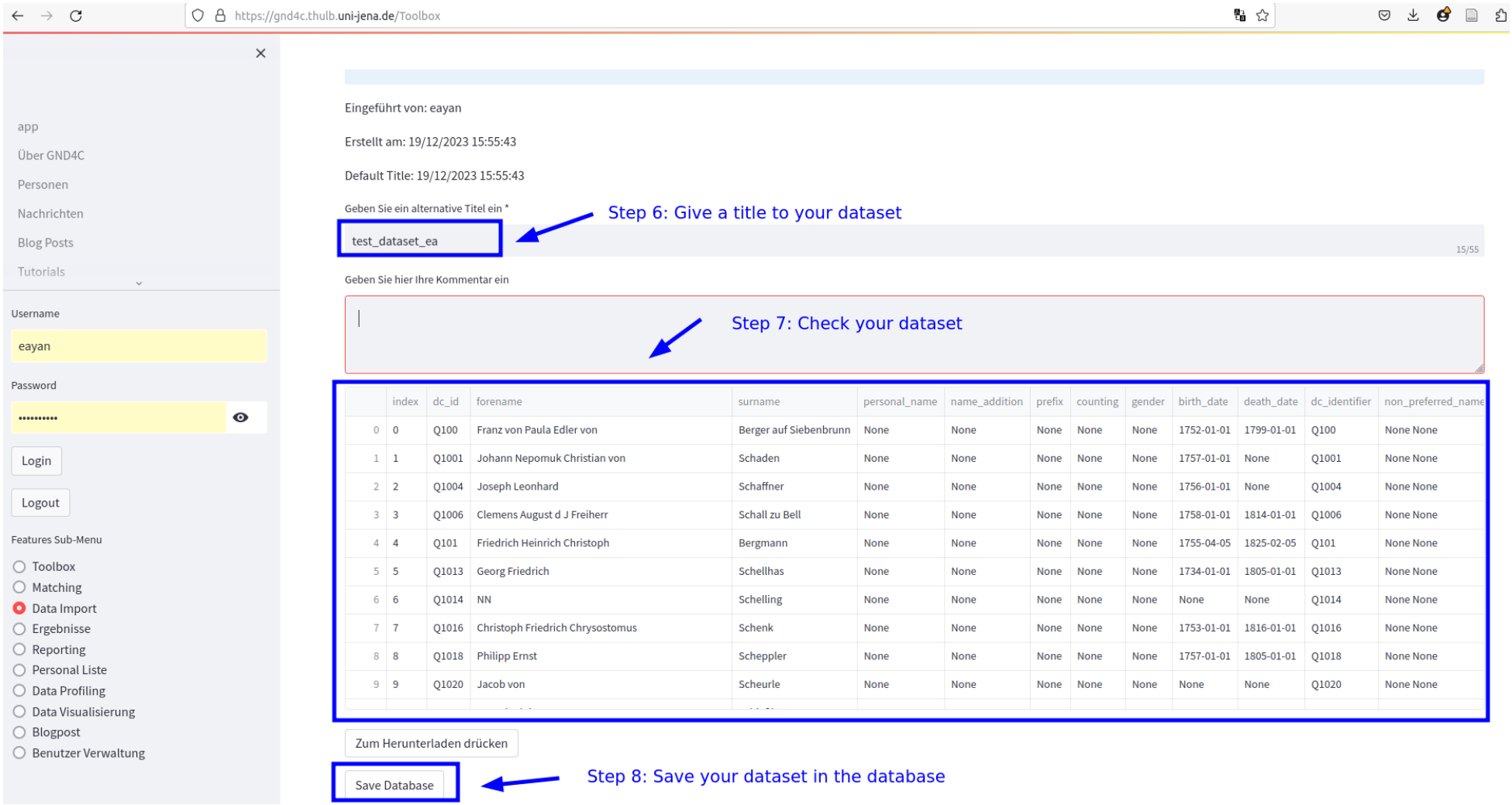
Step 11: Import Your Dataset:

When you manage to start your app and if everything works fine, you canimport your dataset into the Toolbox environment/Postgres database in order to create querying and matching events. Please first use the user-credentials you created above and log into the Toolbox environment and follow the steps defined below.

Attention: If you do not import any valid data set as shown in the instructions, the app will not work properly and give you any results but rather error registrations only.

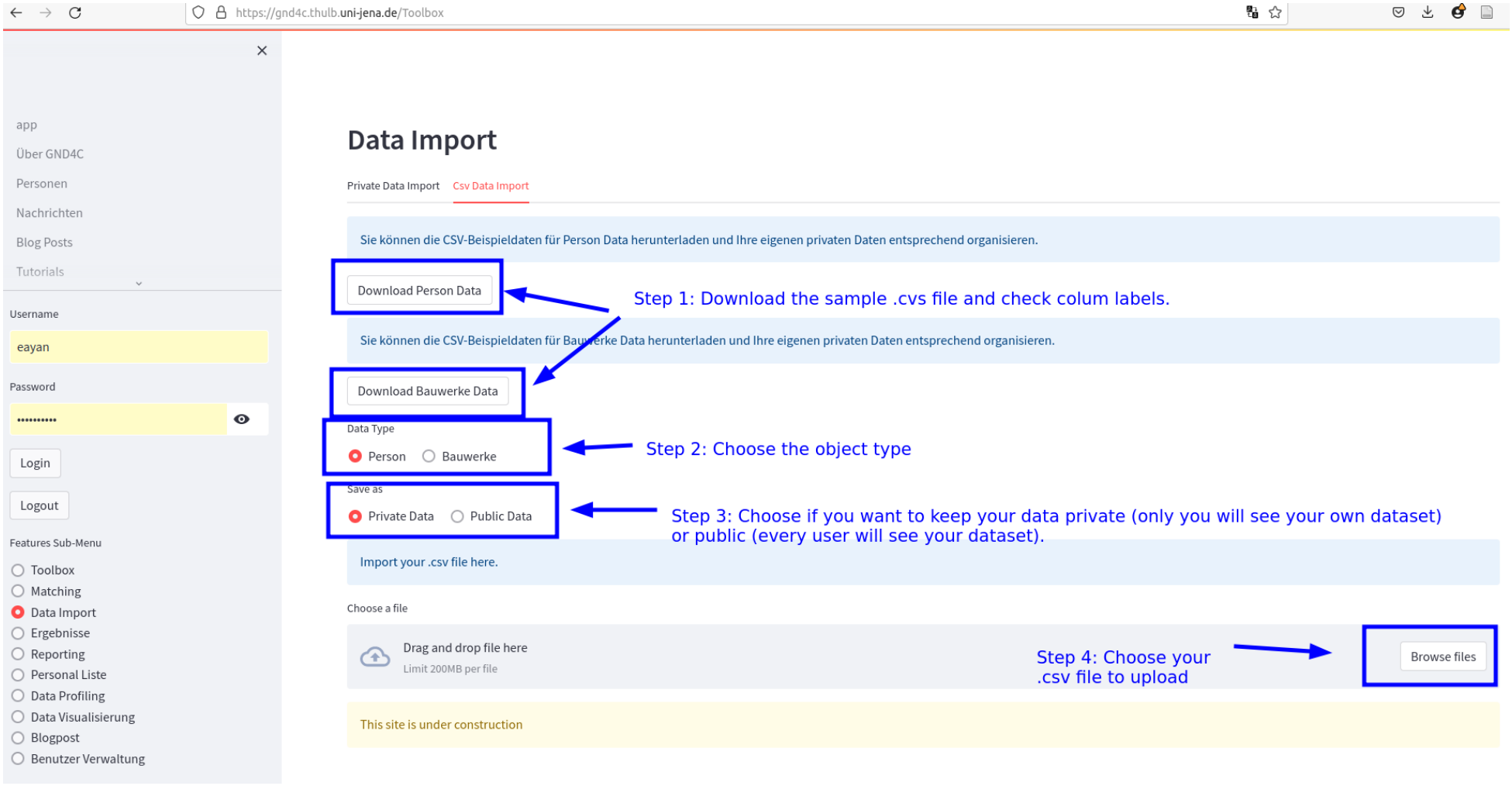
If you have your dataset as .json, please follow the steps to import .json dataset

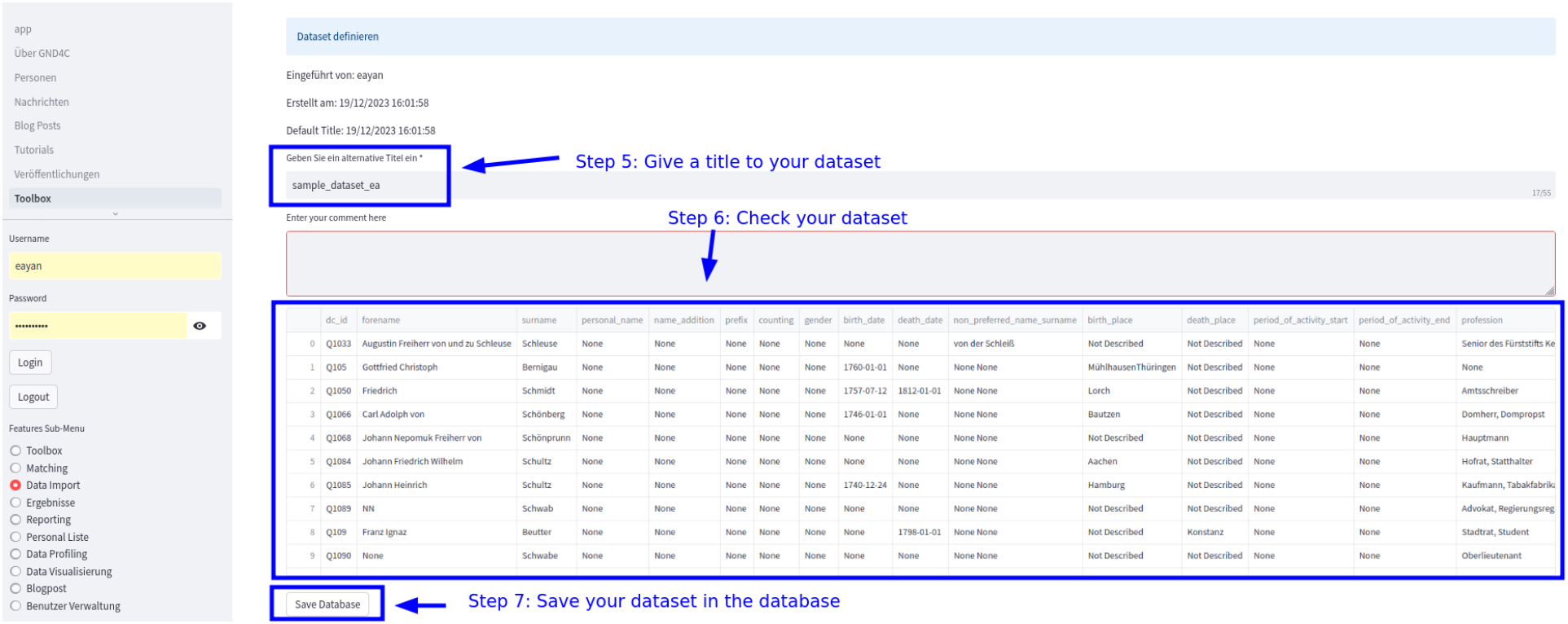




	index	dc_id	forename	surname	personal_name	name_addition	prefix	counting	gender	birth_date	death_date	dc_identifier	non_preferred_name
0	0	Q100	Franz von Paula Edler von	Berger auf Siebenbrunn	None	None	None	None	None	1752-01-01	1799-01-01	Q100	None None
1	1	Q1001	Johann Nepomuk Christian von	Schaden	None	None	None	None	None	1757-01-01	None	Q1001	None None
2	2	Q1004	Joseph Leonhard	Schaffner	None	None	None	None	None	1756-01-01	None	Q1004	None None
3	3	Q1006	Clemens August d J Freiherr	Schall zu Bell	None	None	None	None	None	1758-01-01	1814-01-01	Q1006	None None
4	4	Q101	Friedrich Heinrich Christoph	Bergmann	None	None	None	None	None	1755-04-05	1825-02-05	Q101	None None
5	5	Q1013	Georg Friedrich	Schellhas	None	None	None	None	None	1734-01-01	1805-01-01	Q1013	None None
6	6	Q1014	NN	Schelling	None	None	None	None	None	None	None	Q1014	None None
7	7	Q1016	Christoph Friedrich Chrysostomus	Schenk	None	None	None	None	None	1753-01-01	1816-01-01	Q1016	None None
8	8	Q1018	Philipp Ernst	Scheppler	None	None	None	None	None	1757-01-01	1805-01-01	Q1018	None None
9	9	Q1020	Jacob von	Scheurle	None	None	None	None	None	None	None	Q1020	None None

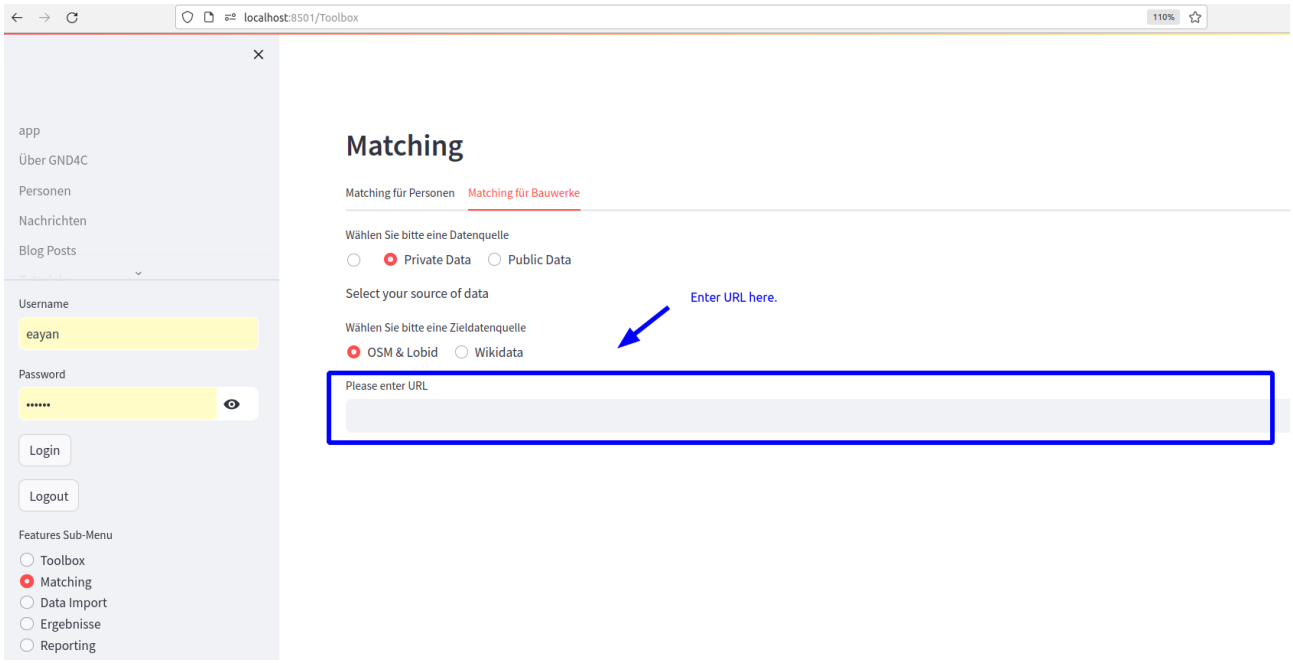
If you have your dataset as .csv, please follow the steps to import .csv dataset




app
Über GND4C
Personen
Nachrichten
Blog Posts
Tutorials
Veröffentlichungen
Toolbox
Username
eayan
Password
Login
Logout
Features Sub-Menu
Toolbox
Matching
Data Import
Ergebnisse
Reporting
Personal Liste
Data Profiling
Data Visualisierung
Blogpost
Benutzer Verwaltung
Dataset definieren
Eingeführt von: eayan
Erstellt am: 19/12/2023 16:01:58
Default Title: 19/12/2023 16:01:58
Geben Sie ein alternative Titel ein *
sample_dataset_ea
17/55
Step 5: Give a title to your dataset
Enter your comment here
Step 6: Check your dataset
dc_id forename surname personal_name name_addition prefix counting gender birth_date death_date non_preferred_name_surname birth_place death_place period_of_activity_start period_of_activity_end profession
0 Q1033 Augustin Freiherr von und zu Schleuse Schleuse None None None None None None None von der Schleiß Not Described Not Described None None Senior des Fürststifts Ke
1 Q105 Gottfried Christoph Bernigau None None None None None 1760-01-01 None None None MühlhausenThüringen Not Described None None None
2 Q1050 Friedrich Schmidt None None None None None 1757-07-12 1812-01-01 None None Lorch Not Described None None Amtsschreiber
3 Q1066 Carl Adolph von Schönberg None None None None None 1746-01-01 None None None Bautzen Not Described None None Domherr, Dompropst
4 Q1068 Johann Nepomuk Freiherr von Schönprunn None None None None None None None None None Not Described Not Described None None Hauptmann
5 Q1084 Johann Friedrich Wilhelm Schultz None None None None None None None None None Aachen Not Described None None Hofrat, Statthalter
6 Q1085 Johann Heinrich Schultz None None None None None 1740-12-24 None None None Hamburg Not Described None None Kaufmann, Tabakfabrik
7 Q1089 NN Schwab None None None None None None None None None Not Described Not Described None None Advokat, Regierungsreg
8 Q109 Franz Ignaz Beutter None None None None None None 1798-01-01 None None Not Described Konstanz None None Stadtrat, Student
9 Q1090 None Schwabe None None None None None None None None None Not Described Not Described None None Oberlieutenant
Save Database
Step 7: Save your dataset in the database

Step 12: Define Your OpenStreet Map URL Credentials:

If you attempt to query for buildings (Bauwerke) data against OpenStreet Map via a service like Geofabrik, you need to enter your URL with your own credentials.



Windows Installation Guide:

Step 1: Install PostgreSQL:

Download the PostgreSQL installer for Windows from the official website:

https://www.postgresql.org/download/windows/

Follow the installation instructions.

Step 2: Install PGAdmin:

Install PGAdmin:

Download the latest version of PGAdmin from the official website:

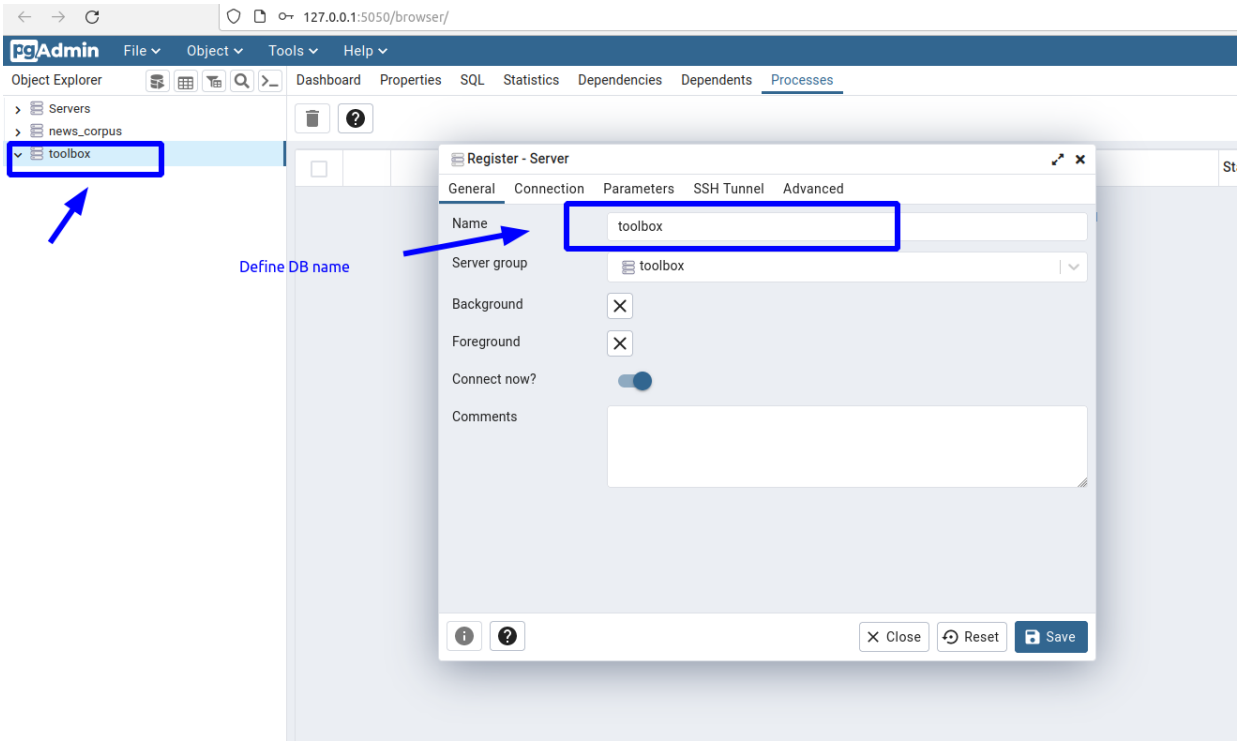
https://www.pgadmin.org/download/

Follow the installation instructions.

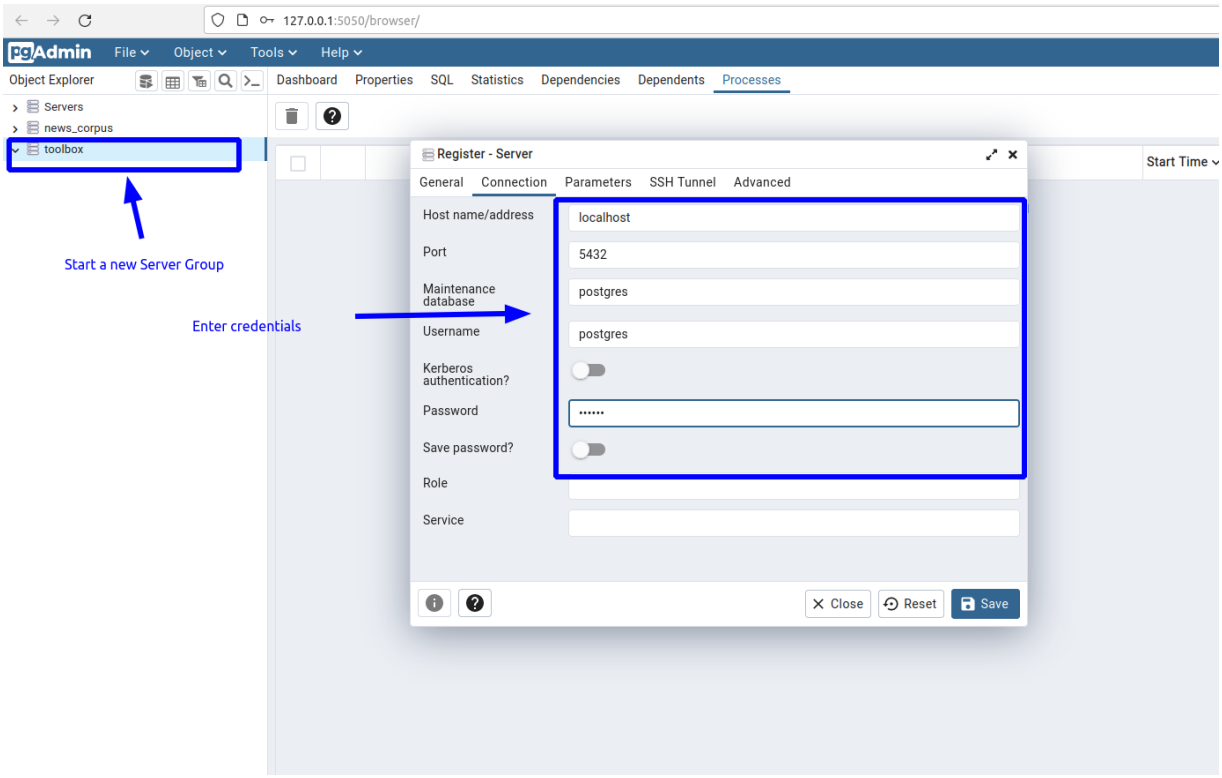
Step 3: Create a PostgreSQL database and user:

Launch your PGAdmin and log in using the PostgreSQL superuser credentials.

Create a new server group (toolbox), user (postgres), password (postgres) and a new database (postgres) using PGAdmin. After this configuration, the settings should look like the following screenshots below.



Screenshot 1



Screenshot 2

Step 4: Install Conda:

Download the Conda installer for Windows from the official website:

https://docs.conda.io/projects/conda/en/latest/user-guide/install/windows.html

Follow the installation instructions.

Step 5: Install VSCode:

Download the VSCode installer for Windows from the official website:

https://code.visualstudio.com/download

Follow the installation instructions.

Step 6: Create a Conda virtual environment for Streamlit in VSCode:

Start your VSCode.
Open a Terminal and enter the command:

conda create --name streamlit python=3.9

Activate the environment:
conda activate streamlit

Step 7: Install Streamlit and psycopg2 within the virtual environment:

Install Streamlit and psycopg2 using pip:

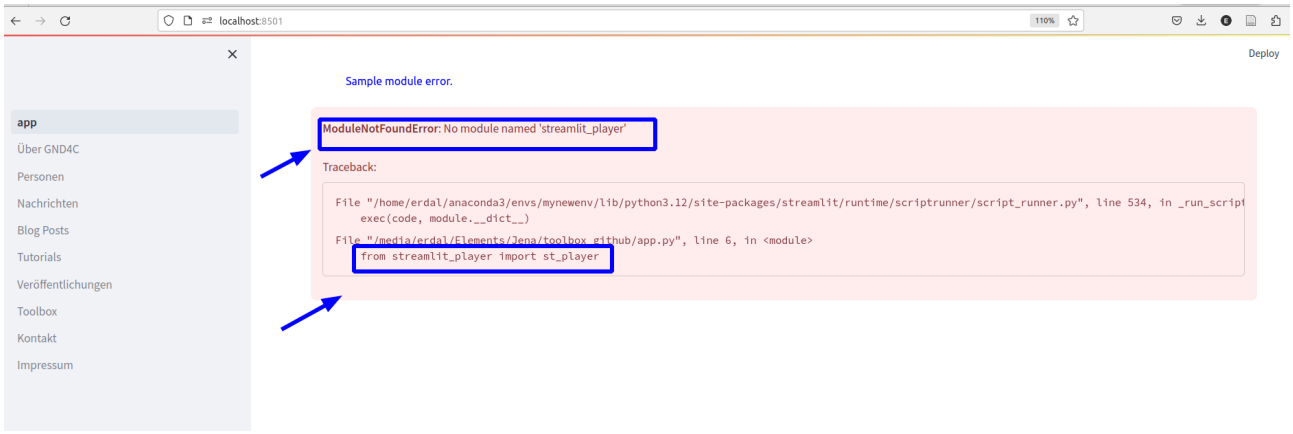
Open a terminal and enter the command:

pip install streamlit psycopg2

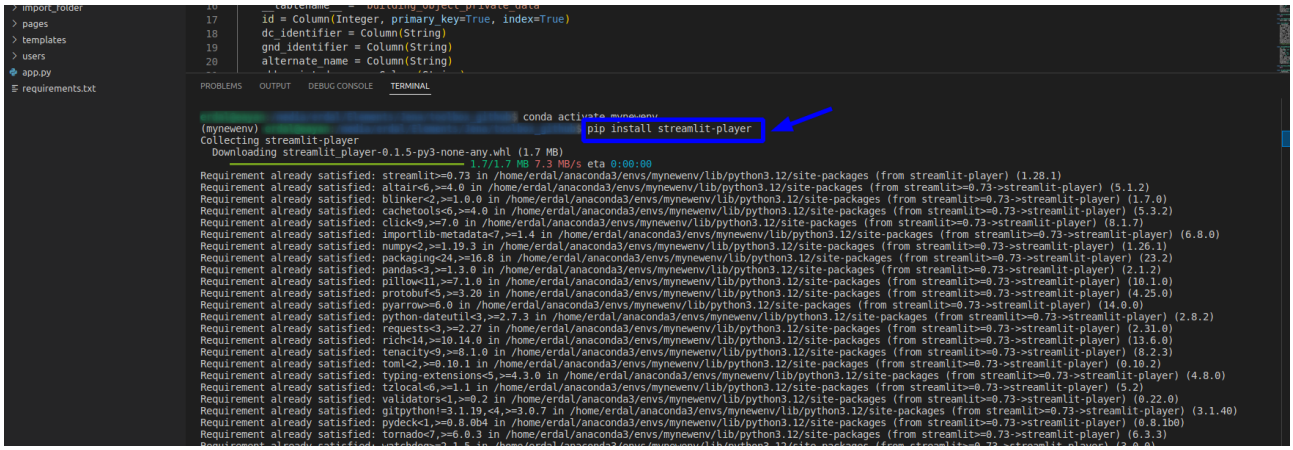
Also install other requirements via the following command:

pip install -r requirements.txt

In case there are still errors with the uninstalled Python libraries/modules as in screenshot 3 below, please attempt to install the module manually by following the instructions[2] in screenshot 4 below.



Screenshot 3: Sample Module Error Registration



Screenshot 4: Sample solution for Module error registration

Step 8: Run FastApi Backend:

If you have not switched off your VSCode you do not have to restart your Terminal, otherwise start your VSCode and open the folder of the app in your VSCode. Start a “New Terminal”. Enter the following commands in your terminal:

conda activate streamlit

uvicorn main:app --reload

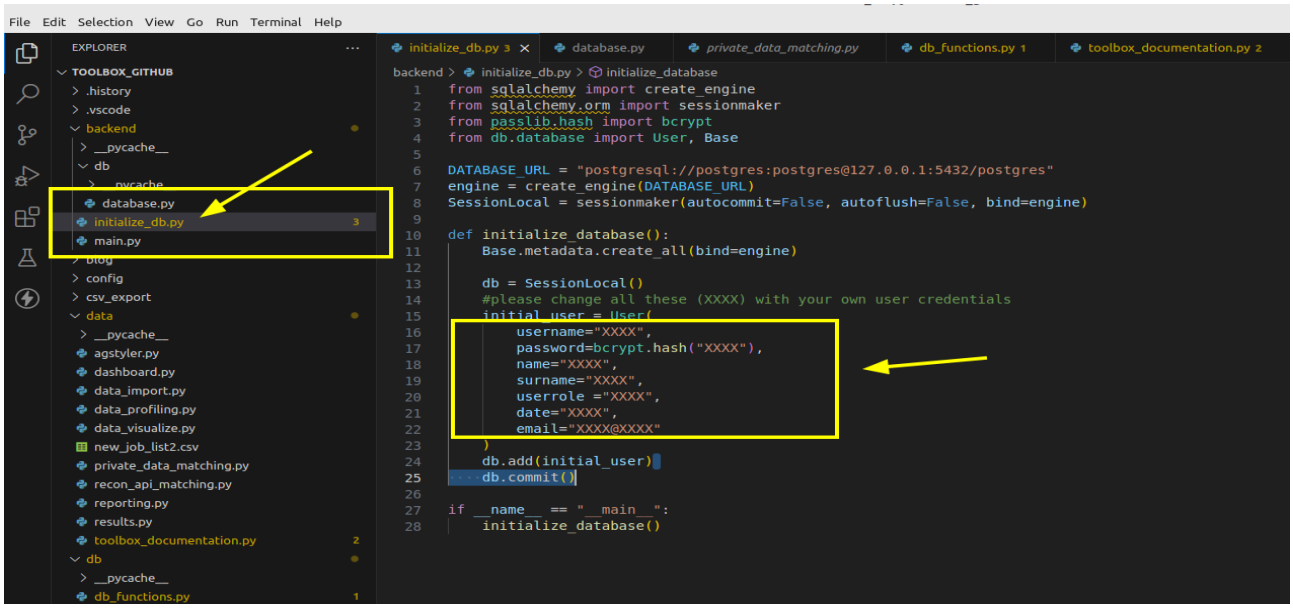
Step 9: Create a User in the Database:

2 Please pay attention: Each module may require different installation command (pip install [name-of-library]) so please first check the installation command with the PyPi site of the library (look for an example here: https://pypi.org/project/streamlit-player/)

Start your VSCode and open the folder of the app in your VSCode. Start a “New Terminal”. Enter the following commands in your terminal:

cd backend

#Update the db connection and user credentials in the file “initialize_db.py” and “save” the file as in the screen shot below



Screen shot 5: Updating “initialize_db.py”

#Run the following command in your Terminal

python initialize_db.py

Step 10: Run Toolbox (Streamlit) app:

Start your VSCode and open the folder of the app in your VSCode. Start a “New Terminal”. Enter the following commands in your terminal:

conda activate streamlit

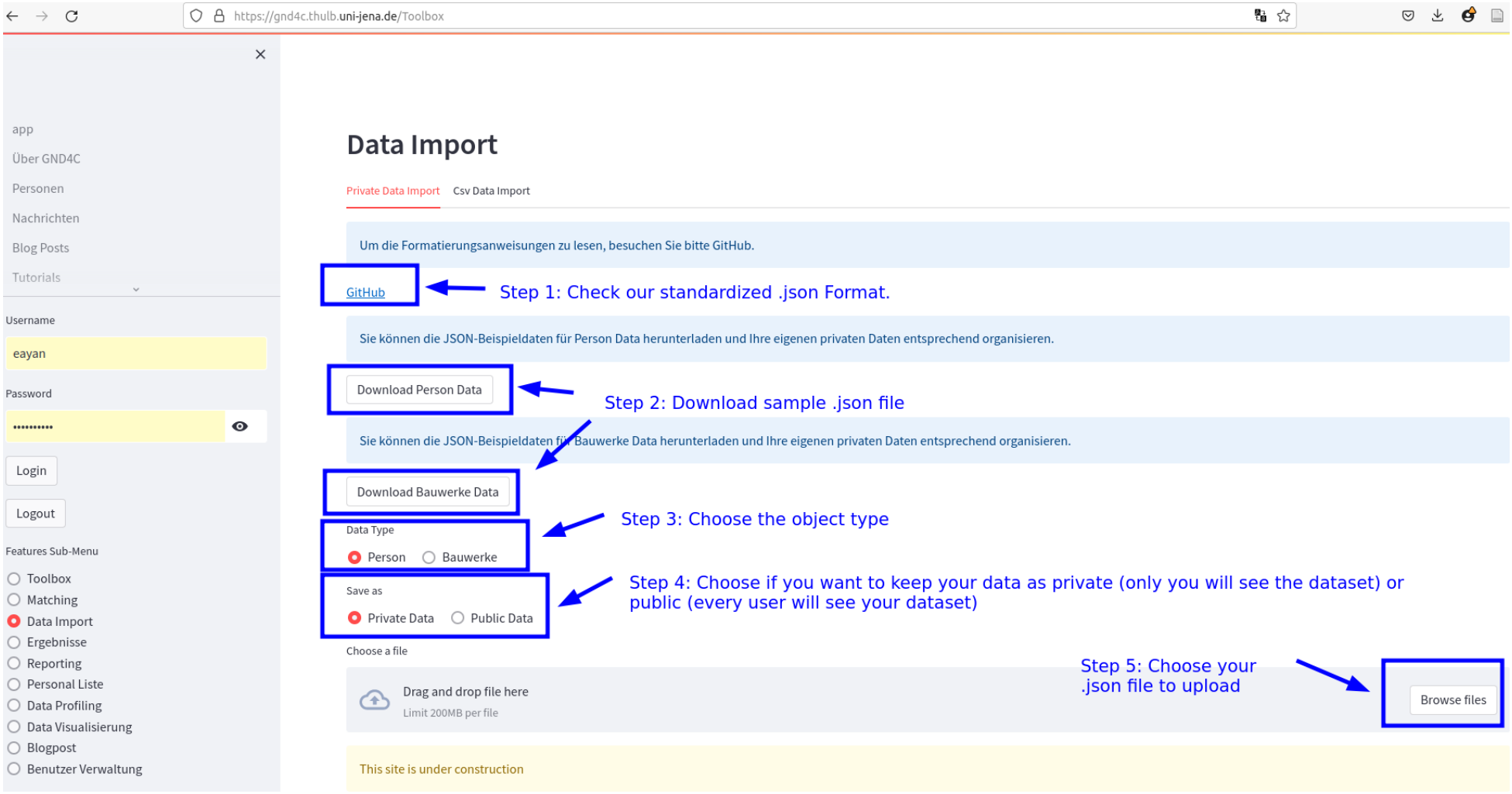
streamlit run app.py

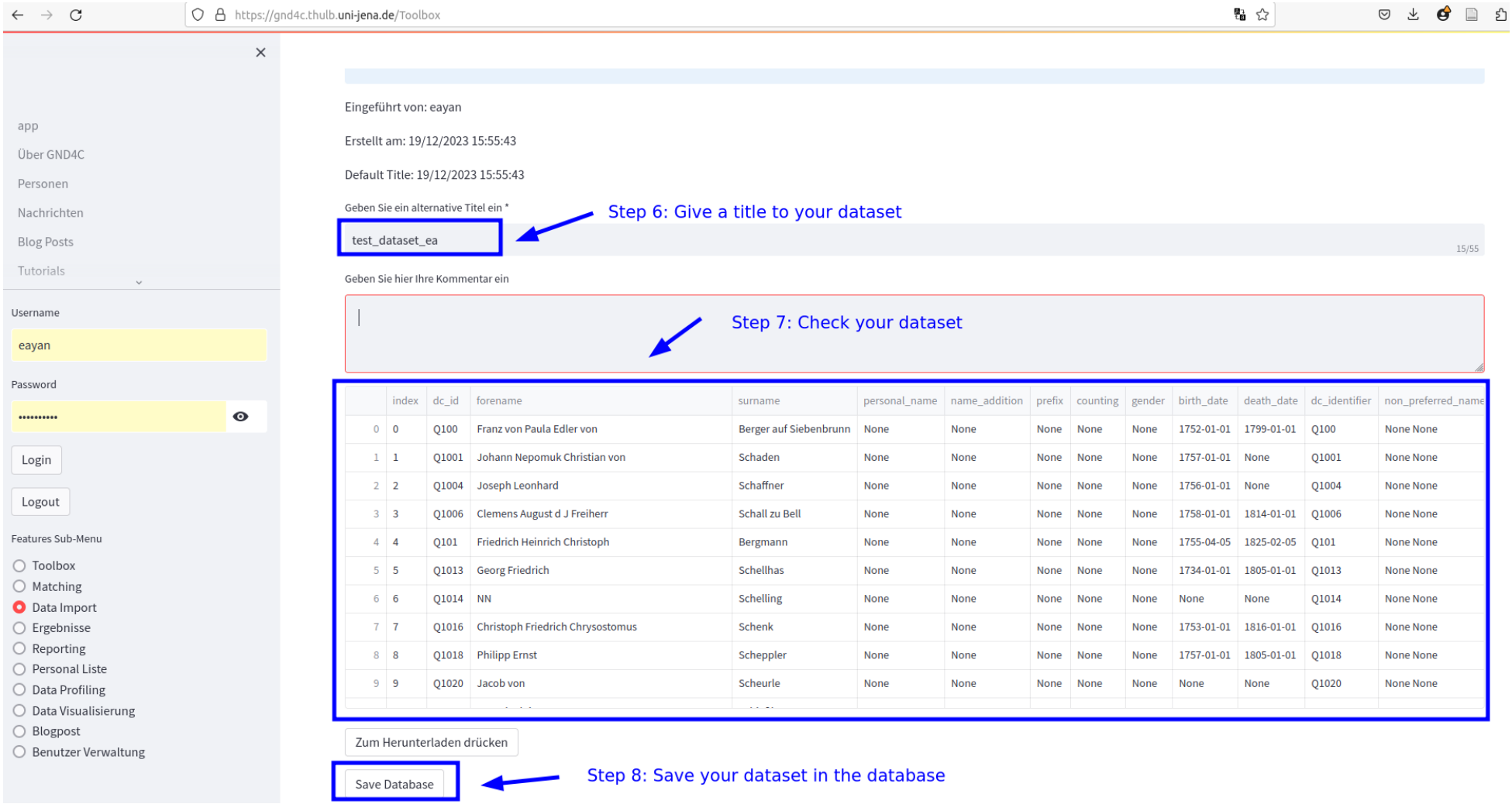
Step 11: Import Your Dataset:

When you manage to start your app and if everything works fine, you can import your dataset into the Toolbox environment/Postgres database in order to create querying and matching events. Please first use the user-credentials you created above and log in to the Toolbox environment and follow the steps defined below.

Attention: If you do not import any valid data set as shown in the instructions, the app will not work properly and give you any results but rather error registrations only.

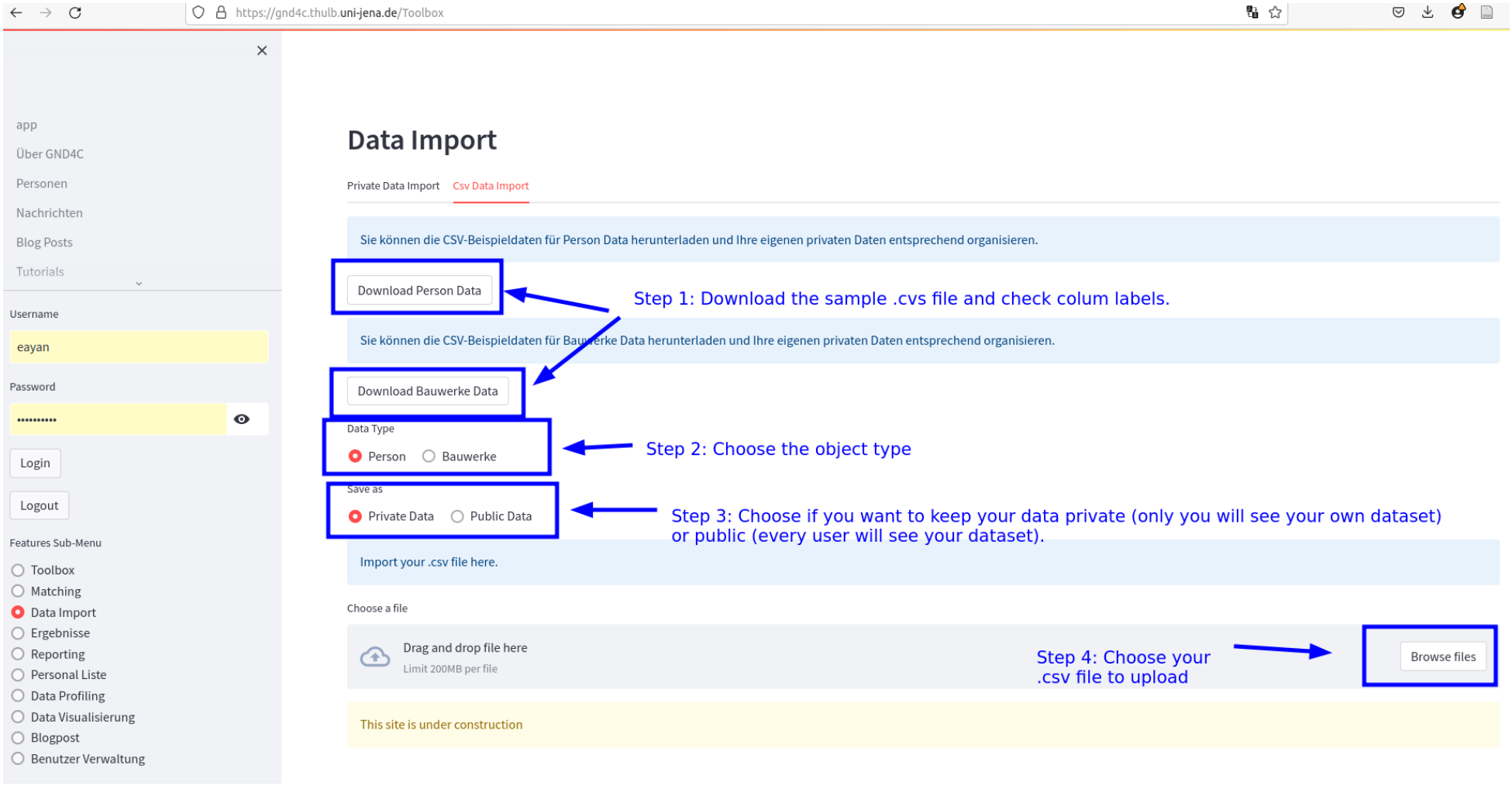
If you have your dataset as .json, please follow the steps to import .json dataset

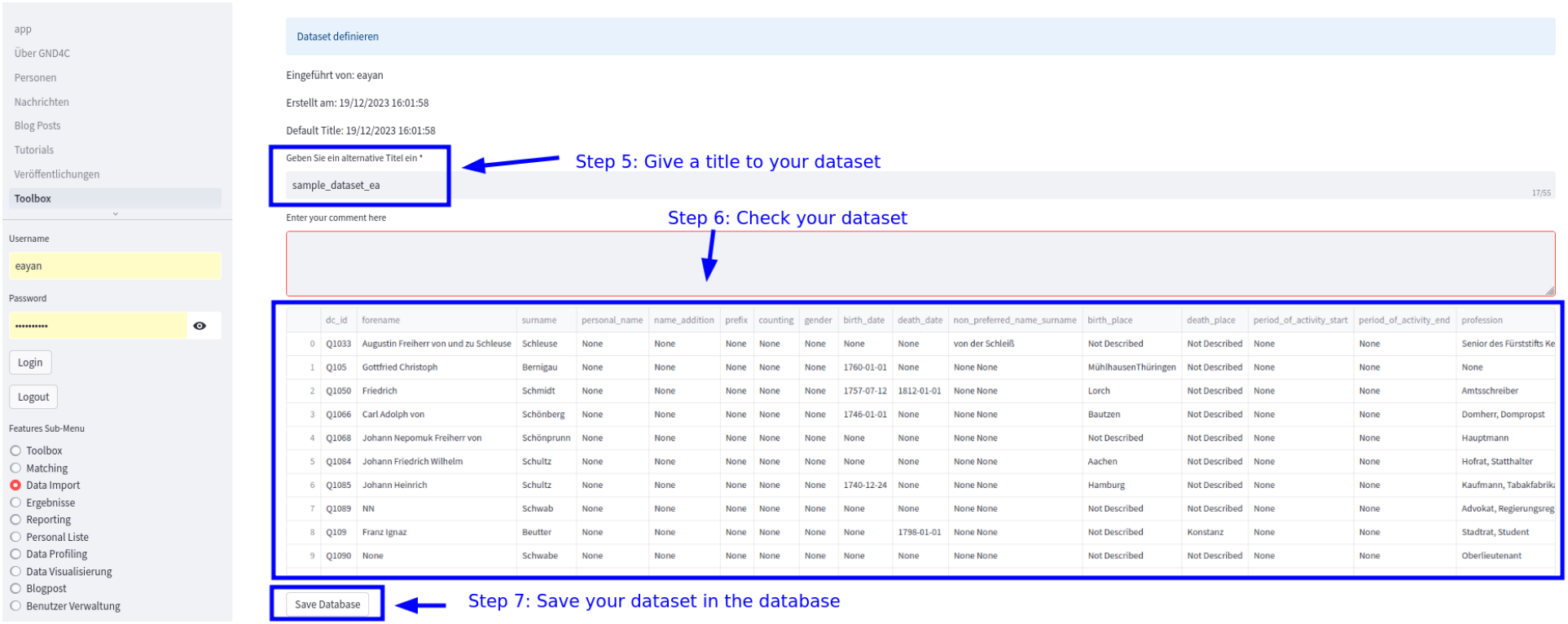




	index	dc_id	forename	surname	personal_name	name_addition	prefix	counting	gender	birth_date	death_date	dc_identifier	non_preferred_name
0	0	Q100	Franz von Paula Edler von	Berger auf Siebenbrunn	None	None	None	None	None	1752-01-01	1799-01-01	Q100	None None
1	1	Q1001	Johann Nepomuk Christian von	Schaden	None	None	None	None	None	1757-01-01	None	Q1001	None None
2	2	Q1004	Joseph Leonhard	Schaffner	None	None	None	None	None	1756-01-01	None	Q1004	None None
3	3	Q1006	Clemens August d J Freiherr	Schall zu Bell	None	None	None	None	None	1758-01-01	1814-01-01	Q1006	None None
4	4	Q101	Friedrich Heinrich Christoph	Bergmann	None	None	None	None	None	1755-04-05	1825-02-05	Q101	None None
5	5	Q1013	Georg Friedrich	Schellhas	None	None	None	None	None	1734-01-01	1805-01-01	Q1013	None None
6	6	Q1014	NN	Schelling	None	None	None	None	None	None	None	Q1014	None None
7	7	Q1016	Christoph Friedrich Chrysostomus	Schenk	None	None	None	None	None	1753-01-01	1816-01-01	Q1016	None None
8	8	Q1018	Philipp Ernst	Scheppler	None	None	None	None	None	1757-01-01	1805-01-01	Q1018	None None
9	9	Q1020	Jacob von	Scheurle	None	None	None	None	None	None	None	Q1020	None None

If you have your dataset as .csv, please follow the steps to import .csv dataset




app
Über GND4C
Personen
Nachrichten
Blog Posts
Tutorials
Veröffentlichungen
Toolbox
Username
eayan
Password
Login
Logout
Features Sub-Menu
Toolbox
Matching
Data Import
Ergebnisse
Reporting
Personal Liste
Data Profiling
Data Visualisierung
Blogpost
Benutzer Verwaltung
Dataset definieren
Eingeführt von: eayan
Erstellt am: 19/12/2023 16:01:58
Default Title: 19/12/2023 16:01:58
Geben Sie ein alternative Titel ein *
sample_dataset_ea
17/55
Step 5: Give a title to your dataset
Enter your comment here
Step 6: Check your dataset
dc_id forename surname personal_name name_addition prefix counting gender birth_date death_date non_preferred_name_surname birth_place death_place period_of_activity_start period_of_activity_end profession
0 Q1033 Augustin Freiherr von und zu Schleuse Schleuse None None None None None None None von der Schleiß Not Described Not Described None None Senior des Fürststifts Ke
1 Q105 Gottfried Christoph Bernigau None None None None None 1760-01-01 None None None MühlhausenThüringen Not Described None None None
2 Q1050 Friedrich Schmidt None None None None None 1757-07-12 1812-01-01 None None Lorch Not Described None None Amtsschreiber
3 Q1066 Carl Adolph von Schönberg None None None None None 1746-01-01 None None None Bautzen Not Described None None Domherr, Dompropst
4 Q1068 Johann Nepomuk Freiherr von Schönprunn None None None None None None None None None Not Described Not Described None None Hauptmann
5 Q1084 Johann Friedrich Wilhelm Schultz None None None None None None None None None Aachen Not Described None None Hofrat, Statthalter
6 Q1085 Johann Heinrich Schultz None None None None None 1740-12-24 None None None Hamburg Not Described None None Kaufmann, Tabakfabrik
7 Q1089 NN Schwab None None None None None None None None None Not Described Not Described None None Advokat, Regierungsreg
8 Q109 Franz Ignaz Beutter None None None None None None 1798-01-01 None None Not Described Konstanz None None Stadtrat, Student
9 Q1090 None Schwabe None None None None None None None None None Not Described Not Described None None Oberlieutenant
Save Database
Step 7: Save your dataset in the database

Step 12: Define Your OpenStreet Map URL Credentials:

If you attempt to query for buildings (Bauwerke) data against OpenStreet Map via a service like Geofabrik you need to enter your URL with your own credentials.