

# OTU AIM Database Manual

## 1 Requirements for the program to work

1. Installed R version 4.3.0
2. Installed RStudio
3. Access to a PostgreSQL database

## 2 Starting and closing the application

You must do these steps every time before you can use the application.

1. Start the PostgreSQL Server
  - a. open windows cmd
  - b. Run following code:

```
cd "C:\Program Files\PostgreSQL\16\bin"
pg_ctl start -D "C:\Program Files\PostgreSQL\16\data"
```
2. Start RStudio  
Run code:

```
library(AIMdb)
run_app()
```
3. after finishing working close the server in the cmd again  

```
pg_ctl stop -D "C:\Program Files\PostgreSQL\16\data"
```

If you want to access the data base via pgAdmin you can use this Password for data base: ö+87=v

## 3 AIMdb usage

### 1.1 Data upload

You can upload new projects here. You just have to select an excel file, type in the project name and customer name and Load data. Don't use any spaces for the project and customer name. If the project you want to upload already exists and you only want to change some data, use the update function instead. If you upload the excel sheet with new columns or new meta data, the old data won't be overwritten and duplicates will occur.

### 1.2 Search data

Here you can search for any character type search term. It is not possible to search for specific dates or read counts. But you can limit this by using the filters below the result table. Inserting sequence data is possible. Even finding sequence sections is no problem. On the left side you can choose which attributes you want to show. After clicking the search button two result table appear. The first one shows all records which were found in the database. This means that it can also show duplicate rows, because the rows differ in attributes which are not selected in the left panel. If you are searching for a

specific Family like “Cecidomyiidae”, it will not only show the records where the consensus Family is Cecidomyiidae, but also where only the BOLD data base indicates Cecidomyiidae.

The second table below that summarizes the attributes for each BOLD unique BOLD BIN. There are 2 columns for every attribute. The first column shows in how many records the BIN is present of the whole records found for this search term and attribute. The second column simply shows these records.

In the upper right corner, you can filter both tables again for a specific search term. there are also filters for each column below the table.

### 1.3 Create plots

1. Choose which projects you want to include into your analysis
2. Choose the plot type bar chart and further configurations tools will open
3. Select the category of the x-axis and which attribute you want to count on the y-axis

→ The graph then shows the unique counts of e.g. BOLD BINs per category.

#### **Color variable:**

If you want to segment the bars further, you can use the colour variable to do so. But this time the graph will not only show the unique BINs, but the unique BINs per colour variable category stacked above each other.

#### **Count type:**

The count type can only be changed to reads count if the y-variable is numerical. If you do so, the reads count will be added together.

#### **Percent switch:**

The percent switch turns the y-scale into percentages.

#### **Turn on taxon filter:**

You can choose a taxonomy you want to filter and enter all values you want to be included in the field below separated by comma. (like this: BOLD:AAH2876,BOLD:AAH3617, BOLD:AAH3623)

### 1.4 Delete or create tables

You probably won't be using this much, because creating a table is not very useful if you can't connect it to the other ones. Changing the data base structure requires some SQL knowledge. If you want to do it, you should download pgAdmin and manage the database there.

Deleting a table is useful if you inserted wrong data or the data integrity is damaged. But be aware, that all data is lost then. If you want to delete tables, you firstly have to click “refresh tables” and select a table you want to delete. If you want to reset the database and delete all data, you can you the delete all Tables section.

### 1.5 View and edit tables

Here you can browse through an entire table in the database. If you want to change single entries, you can double click the cell and change it. Sadly, you have to do it cell by cell.

## 1.6 Update data

If you want to change data for more than single cells you can do so with the Update tool. **Don't try to change the id columns.**

1. choose table to change
2. choose column to change
3. Write new value you want to update
4. If you don't specify the WHERE-Clause, all entries in this column will be changed to your new value. If you only want to change it for specific rows you can use another column to specify in which rows you want the value to update.

This example updates all habitats to semagr, where the location name in the location\_name column is either Soller, Mühlbauer or E-Hof.

Example:

**Table**

location

**Column**

habitat

**New value**

semagr

**WHERE-clause**

location\_name IN ('Soller', 'Mühlbauer', 'E-Hof')

## 1.7 Bug fixing

If the application shut down, just restart R in RStudio. You can do that in the upper left corner by clicking on Session → Restart R. Afterwards you have to load the package via library(AIMdb) again and rerun the program run\_app().

## 1.8 Bugs

Bugs I am aware of but can not fix. Before the first time using the search tool after starting the application, you must open the graphs tab once. Otherwise, the graph tool will not work after executing one search command. If you forgot it just restart the app.