

TerraByte Client

The TerraByte Client allows you to download data from our portal-server hosted on Compute Canada (Arbutus). For that purpose, the server exposes an API that can be accessed through HTTP-calls from the outside.

To interact with the portal, I created the TerraByte Client that does most of the heavy lifting for you. It looks like this:

The screenshot shows the TerraByte Client application window. The window title is "TerraByte Client". It has a tabbed interface with three tabs: "EAGL-I Data", "Fielddata", and "Info". The "EAGL-I Data" tab is selected. The interface contains several input fields and checkboxes. The "Start Date" is set to "17.04.2020" and the "End Date" is "15.12.2021". The "Min. age in days" is "0" and the "Max. age in days" is "1000". The "Plants" section has checkboxes for "Barley", "BarnyardGrass", "Black Bean", "CanadaThistle", and "Canola". The "Plant ID" field is empty. The "Min. Resolution (px)" is "0" and the "Max. Resolution (px)" is "4000". Below these fields are four checkboxes: "Single Plant Images" (checked), "Multiple Plant Images", "Bounding Box Images", and "JSON Files". To the right of these checkboxes is an "Archived Queries" dropdown menu. At the bottom of the window, there is a "Result:" section that says "Not logged in". Below this are six buttons: "Check Query", "Get Sample", "Download", "Set Credentials", "Set Sample Path", and "Set Download Path".

There are a couple of simple controls here, which I will explain in more detail.

Basic Workflow

A typical workflow of getting data would look something like this:

- 1) Adjust the search parameters to the data you are interested in. For example all Canolas of April 2021.
- 2) Using the "Check Query" button in the client get a first feel for how much data is available under these search parameters. Note, that if you select multiple options from "Single Plant Images, Multiple Plant Images, etc..." that the number of files is the sum of the different type of files selected.
- 3) Before you start a long download, you should check whether the images found are really what you expected. By hitting "Sample" a small number of images will be downloaded into your sample-folder (see below)

- 4) Once satisfied with the results you can press “Download” to get all the images of your query. Depending on your query the total download will take minutes (hundreds to thousands of images) up to hour (tens of thousands of images), days or even weeks (millions of images). No matter the number of images you plan to download the images are being streamed in. Meaning you will see the first images in your download-folder (see below) already after a couple of seconds usually.

What do the other buttons do?

Credentials

If you want to change your user credentials hit this button. Provide your username and your password in the prompts popping up.

The client program will remember this next time you start it. So, typically you must provide your credentials only once and never again. **WARNING:** If you copy your program folder (for example, because you want to use it on several systems) your user-credentials are copied with it. Do not distribute the program to others in that way, as you will give away your user-credentials too! New users should always get a fresh client from our GitHub page.

Sample Path

When you click this button, a dialog opens that allows you to pick a directory on your file system. Whenever you hit the “Sample” button the samples will be downloaded into that directory.

Download Path

Works the same way as the “Sample” button, but for your actual downloads. If you download several types of data, you might want to change the download location in between these downloads otherwise the images will be all mixed in the same folder and you might have a hard time separating them again.

Filter-Options

Currently you can use the following filters on the database:

Start Data/End Date

This is the at which the images had been taken. You can click the little calendar icon for a mouse-convenient input of a date.

Min/Max age in days

This filters by the age of the plant. We define the age by the number of days that have elapsed between seeding and imaging. For example a plant seeded at 1st of May and imaged on 8th of May has an age of 7 days.

Plants

You can select one or multiple plants you are interested in here. If you select nothing in this box all plants will be returned.

Plant-ID

You can extract each individual plant’s ID from the JSON files provided. If you type something into this field you will only get plants corresponding to this ID.

Min/Max Resolution

Subimages (i.e. single plant images) vary in resolution, depending on the distance and angle between the camera and the plant. By using this filter you can define a range of minimum/maximum width the images (in pixels) should have. As all of them are in a square aspect ratio, it also defines the same range for height of the image.

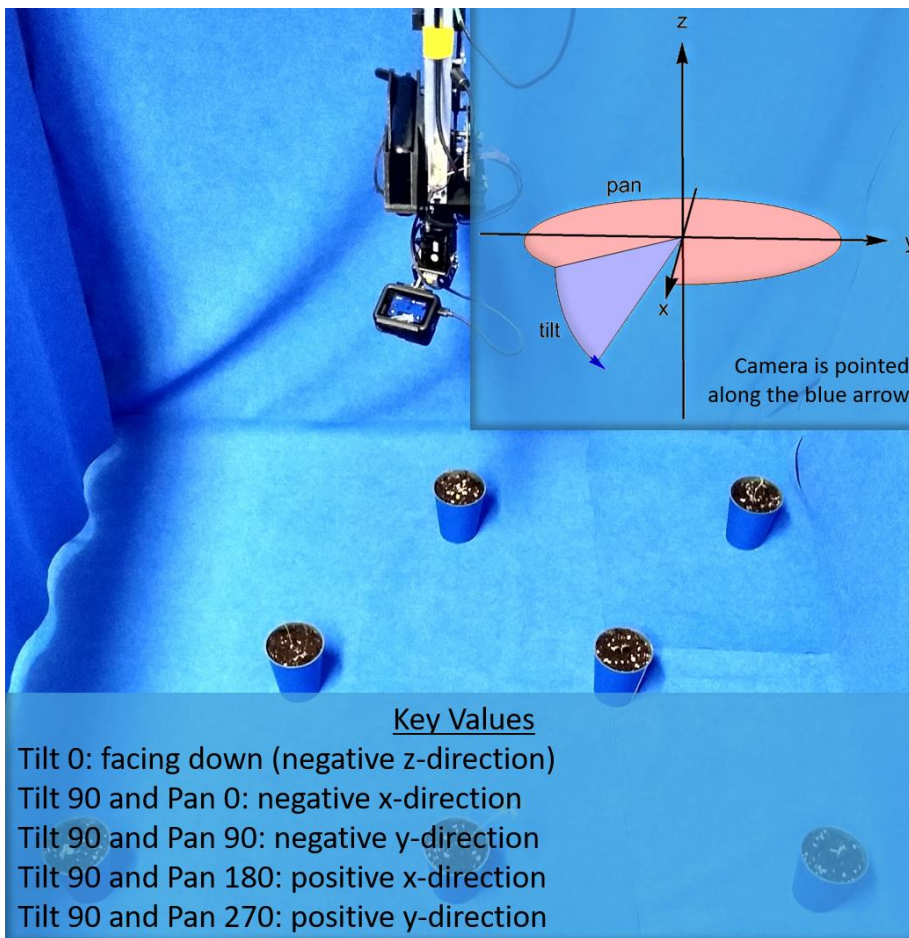
Archives

Under "Archived Queries" you can find a (short) list of archives that we currently provide on the server. These archives are collections of images and data we think you might consider interesting and thus have them pre-zipped lying on the server. The download speed of these archives is faster than defining your own datasets through the filters.

If you have an idea on what a good dataset would be to keep archived on the server, let us know by sending an email to: m.beck@uwinnipeg.ca

Appendix A

The JSON-files contain metadata information about the images available. The following sections explain the different fields in the JSON metadata files.



Here is a typical entry in the multi_plant collection of the MongoDB running on the server:

```
_id: "20200420114722_pose0"
```

```

__v:0
bb_file_name:"20200420114722-pose0-bb.jpg"
bb_images:"2020April/2020_04_20/bb_images"
camera:"GoPro"
camera_pose:Object
  x:0
  y:4.5861309700971065
  z:0
  polar_angle:270
  azimuthal_angle:73.30075576600639
date:"2020-4-20"
date_time:2020-04-20T16:47:22.000+00:00
file_name:"20200420114722-pose0.jpg"
horizontal_res:4000
institute:"UofW"
lens:"Hero 7 Black"
room:"RC"
single_plants:Array
  0:Object
    _id:60aacc6c4d0fb429c835d421
    plant_id:"falconv020"
    label:"WildBuckwheat"
    scientific_name:"Fallopia convolvulus"
    subimage_file_name:"WildBuckwheat202004201147220.jpg"
    date_planted:"2020-02-27"
    planting_date:2020-02-27T00:00:00.000+00:00
    age:53
    position_id:0
    x_min:0.3615
    x_max:0.6625
    y_min:0.44633333333333336
    y_max:0.8476666666666667
  1:Object
  2:Object
  singles:"2020April/2020_04_20/singles"
  source_file_path:"2020April/2020_04_20/plain_images"
  tags:Array
    0:"OnBoardLight"
time:"11:47:22"
version:2
vertical_res:3000

```

Of specific importance are the camera_pose object with its subfields:

```

x:0
y:4.5861309700971065
z:0
polar_angle:270
azimuthal_angle:73.30075576600639

```

These describes the camera position and orientation during the time the image was taken. The following graphic illustrates this further. Keep in mind that (unintuitively) polar_angle corresponds to the pan-angle and azimuthal_angle corresponds to the tilt-angle.

The next important entry is the array of single_plant objects.

```
single_plants:Array
  0:Object
    _id:60aacc6c4d0fb429c835d421
    plant_id:"falconv020"
    label:"WildBuckwheat"
    scientific_name:"Fallopia convolvulus"
    subimage_file_name:"WildBuckwheat202004201147220.jpg"
    date_planted:"2020-02-27"
    planting_date:2020-02-27T00:00:00.000+00:00
    age:53
    position_id:0
    x_min:0.3615
    x_max:0.6625
    y_min:0.44633333333333336
    y_max:0.8476666666666667
  1:Object
  2:Object
    singles:"2020April/2020_04_20/singles"
    source_file_path:"2020April/2020_04_20/plain_images"
    tags:Array
      0:"OnBoardLight"
```

Here the x_max, x_min, y_max, y_min correspond to where the upper left and lower right coordinates of the cropped out image lie in the original image (either given in a pixel value, or in a fraction of the full resolution. In the above example it is a fraction of the full resolution)

More information about the fields being used and their interpretation can be found in our paper at: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0243923>