An Ecologist’s Guide to BIIGLE

This is the living version of the document where users can suggest edits to improve the Ecologist’s Guide to BIIGLE. You can find the finalized, original version on Zenodo here: LINK

# Introduction

Hello and welcome to our BIIGLE manual. This document is intended to help ecologists get started using BIIGLE to annotate their image and video data. The manual was originally written to support new members of the Deep-Sea Conservation Research Unit as well as undergraduate students at the University of Plymouth. It is a collection of information that we have found helpful to document in our experience of setting up and using BIIGLE. It is not intended as a replacement for the official BIIGLE manual available here https://biigle.de/manual. We intend this to be a living document that others can contribute to here https://github.com/DeepSeaCRU/CRU-resources. It is fair to say our current instruction manual is bias toward image data (although video is included), and only provides instructions for setting up using Amazon Web Service as the host for your imagery. If you can provide instruction for setting up on other cloud services, please contribute. But first things first, you need to create a BIIGLE account. To create an account, click on the “sign up” button in the top right corner of the website (https://biigle.de/) homepage. Fill in the module with your details and affiliation, choose a password and, after agreeing to the privacy notice and the terms of use, click on sign up. To log in, click on the green “login” button and insert the email used to sign up, and the chosen password.

|  |
| --- |
| image 1 |

Happy annotating!

# Setting up on BIIGLE

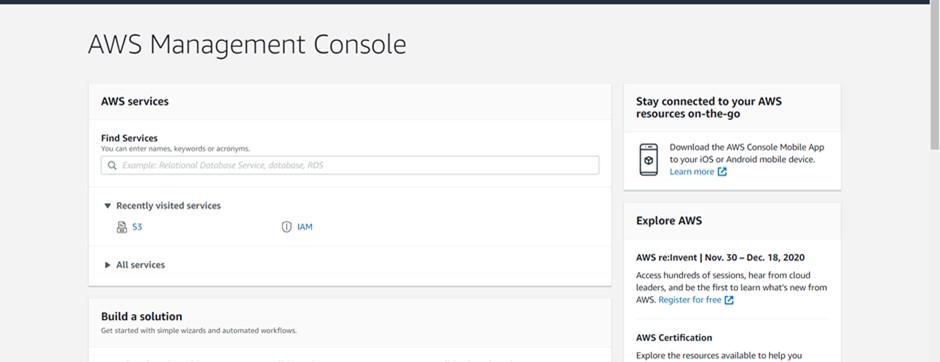
Getting set up on BIIGLE (and other annotation softwares) first requires that you host your imagery data, there are three options for this: via BIIGLE (upload files on homepage); local instances and remote volumes. See the BIIGLE manual for more details. Many institutes use a cloud-based storage solution to host their remote volumes. This is not the same as a google drive or equivalent where you might store or share data. It is a formal repository that enables your images to each have a unique URL. There are many types of cloud-based storage solutions that you may wish to consider, key ones we are aware of include Amazon Web Services (AWS), and Microsoft Azure. Each uses its own terminology to refer to its storage ‘containers’. AWS calls its containers ‘Buckets’, Microsoft Azure calls them ‘Blobs’; they are the same thing.

Below we describe the process for setting up and working with AWS. We invite others to contribute similar text for other cloud-storage options.

## How to get set up with an Amazon Web Services remote server

Create an Amazon Web Services account.

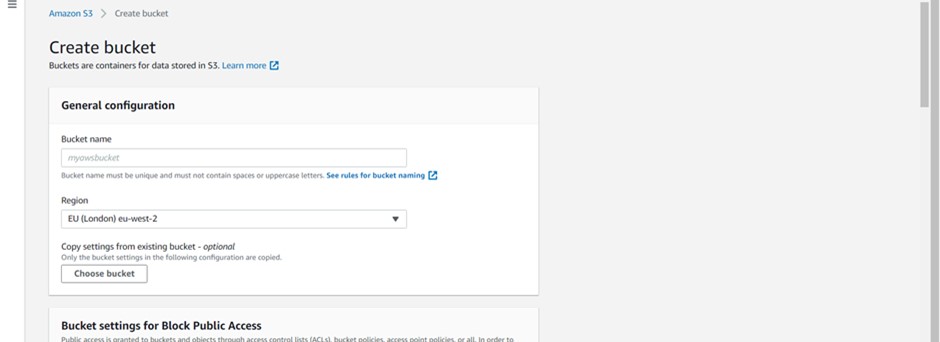
On the AWS management console type “S3” into the find services search box.



Select “S3”

Before you can upload data to Amazon S3, you must create a bucket in one of the AWS regions to store your data.

Click on “create bucket”



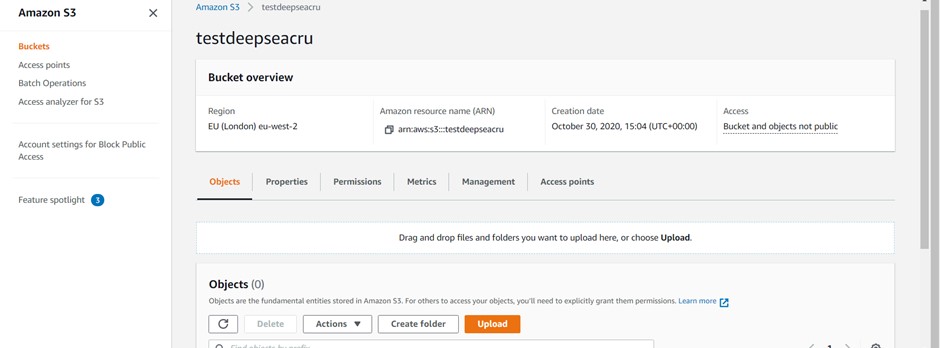
Name your bucket in a sensible way. BIIGLE recommends you use their personal random string generator [available here https://biigle.de/manual/tutorials/volumes/remote-volumes] to do this as this is hard to guess and keeps your data safer. But you can use whatever name you like.

Select your nearest region

Uncheck the “Block all public access” box as BIIGLE needs to see your folder.

Leave all other settings as default and click on “create bucket”

Once you have created your bucket, click on it to take you to this screen



*This next section may not be necessary but is included for completeness.*

Select the permissions tab

Scroll down to Cross-origin resource sharing (CORS) and click on edit.

The cross-origin policy is a security mechanism of web browsers that prevents malicious third parties from extracting sensitive information from your web pages. This includes cases like loading files from remote sources in BIIGLE. Although those files can be displayed, BIIGLE cannot access the raw data in the browser which is a requirement for some features of the image annotation tool. Cross-Origin Resource Sharing (CORS) is a mechanism to manually configure exceptions for the cross-origin policy. With a correct CORS configuration, BIIGLE can process files from remote sources just like regular files.

To set up CORS for the files of your remote source (in this example we are using AWS), you have to update the configuration of the webserver that serves the files.

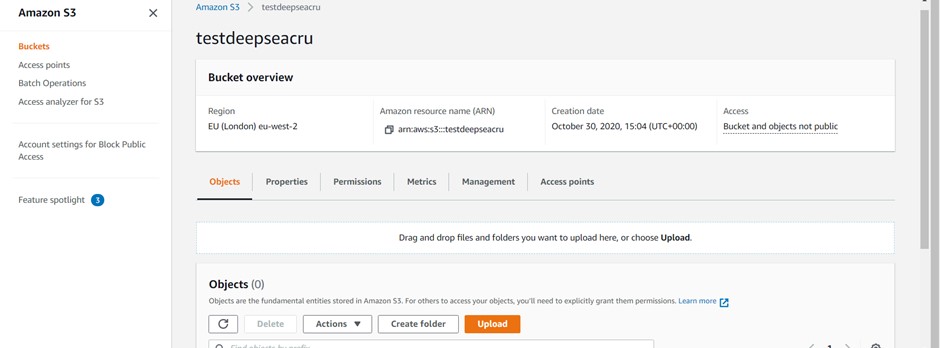
Copy and paste the script below into the text box. Then save changes.

[  
 {  
 "AllowedHeaders": [  
 "\*"  
 ],  
 "AllowedMethods": [  
 "GET"  
 ],  
 "AllowedOrigins": [  
 "\*"  
 ],  
 "ExposeHeaders": []  
 }  
]

*Continue from here if skipping the previous section*

The following section is repeated under the ‘Setting up your BIIGLE volumes’ section. We include here as it is important you test if your AWS bucket is set up correctly before moving on.

To test that your setup is working go back to the “Objects” tab.



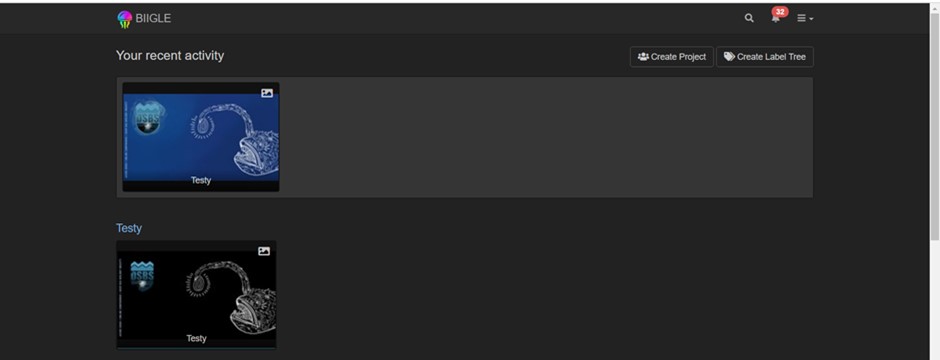
This is where you can upload files, so create a new folder, for example called test.

Click on your folder and upload a couple of images.

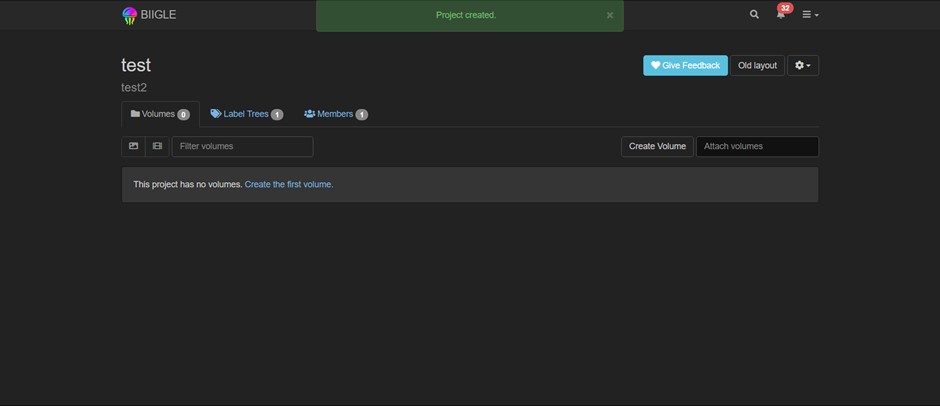
Click “upload”, click “add files”, add your files, scroll down to the access control list and check the “everyone (public access) read check box. A new check box will appear to confirm that you understand the implications of making your files public access, just check it, then scroll down to the bottom of the page and click “upload”. Once uploaded, you can check that the files will be accessible by BIIGLE by clicking on the object URL, if it has been uploaded correctly with public settings enabled, the image or video will open. If the file does not open, check the settings.

Now return to your folder and click on a newly uploaded image. In the overview tab you will see the object URL, you will need to copy this but only the bit up to and including your folder name. This is what BIIGLE will need to link up to.

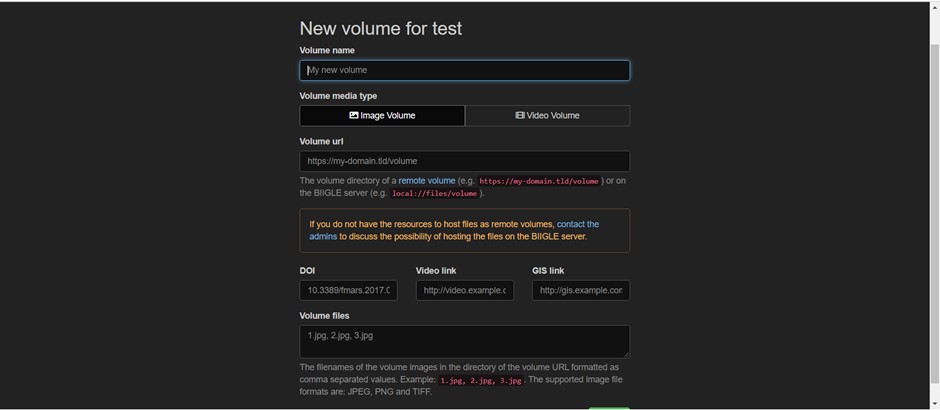
Now log into BIIGLE and go to your dashboard.



Click on create project



Click on create volume



Here you can name your volume, for example it might be a cruise name, a project name or other useful name.

In volume URL paste the URL for your Amazon S3 bucket and folder. In the volume files you will need a list of the names of all the images that are in the file. Below is an R Script to read a directory containing images, and place a comma separated list on the clipboard for you to paste into BIIGLE or a text file for later use.

setwd("F:/DSCRU-frames-main/renamed frames/DIVE295\_SCORPIO\_4")  
list.files() -> imgs  
 writeClipboard(  
 paste0(  
 unlist(paste(imgs,collapse = " , ")  
 ) ,  
 collapse = " , "))

### Enabling others to access your S3 AWS bucket when working in a team

## Formatting your data for use in BIIGLE

## Converting video file to upload to BIIGLE

## Recommendations for the file structure of your bucket and file-naming conventions

# Setting up projects and volumes in BIIGLE

## Recommendations for project / volume structuring in BIIGLE

## Make label trees

### To use a publicly accessible (standard) tree

### Creating your own tree from scratch

## Attaching a label tree to your project

# Make annotations

## Image annotation

## Video annotation

## Setting an annotation session

# Quality control and the largo tool

## Suggested best practice in annotation

# Downloading data and reformatting

## Get and use a report

### Image annotation report - CSV variant

# Using the BIIGLE API

## Intro to BIIGLE’s Application Programming Interface

## Basic requests

### Requesting a BIIGLE report

### Requesting the ID numbers of BIIGLE objects

## Using the API with R and Python

### Accessing label tree information

## Uploading annotations to BIIGLE

# Exporting BIIGLE files for use in YOLO

# Future updates to this manual