Wildlife Insights Data Use & Citation Guide

Wildlife Insights provides access to one of the largest and most diverse camera trap datasets in the world. Before using data from Wildlife Insights in an analysis or for any other purpose, please read through this Data Use and Citation Guide to learn more about the data.

You can download data from your own project (known as a *private download*) or download public data from the Explore page (known as a *public download*). Anyone who downloads data from Wildlife Insights must first agree to the Terms of Use, which require a data user to provide attribution to the dataset creator/creators as required by the dataset license. Continue reading to learn more about the licenses available for Wildlife Insights projects and how to provide attribution.
NOTE: Many projects in Wildlife Insights are ongoing and are continually updated. If you download the same project(s) at different points in time, the resulting dataset may differ. If you plan to publish a scientific manuscript using data from Wildlife Insights, we recommend publishing the dataset used in an open-access public data repository found on the Explore page.

Table of Contents

Data available for download
Verifying data
How to verify whether a record has been reviewed
Downloading images from private downloads
Prerequisites
Downloading images using gsutil software
Viewing images from public downloads
Licenses
Citing data downloaded from Wildlife Insights
References:
In-text citations:

Data available for download

Every download from Wildlife Insights will include:

- **Projects.csv:** metadata about project methodology and objectives, including the type of project (sequence or image) and whether count was recorded in the project;
- **Cameras.csv**: metadata about the devices (cameras) used in the project;
- **Deployments.csv:** metadata about the placement of a camera, including start date, end date, coordinates and other camera settings;
- **Images.csv and (if applicable) Sequences.csv:** Data about the animals detected by the

camera traps are reported in one of two ways depending on how the data was recorded (denoted by *project_type* in the projects.csv). The download package will include both the images.csv and sequences.csv if the request includes sequence projects:
    - The **images.csv** contains data about each individual image, including species identifications and timestamp.
    - The **sequences.csv** contains data about each sequence of images. A sequence is a group of photos that are taken within some time frame (one minute) of each other. All photos in a sequence are tagged as a single detection of an animal or animal group. Data about a sequence of images including species identifications and date/time. The sequences.csv does not include records of every image nor does it include links to images within the sequence.
- **Data use & citation quick guide (**this document);
- **Data Dictionary**: definitions for each field in the files provided; and
- **Wildlife Insights Terms of Use and Privacy Policy**.

Note: Wildlife Insights places certain restrictions on data to protect sensitive or private information, which includes locations of sensitive species, images of humans, and embargoed projects.Public downloads will never include:

- *Data from embargoed deployments (project embargoes are applied separately to deployments within the project);*
- *Exact locations of deployments capturing sensitive species (Read more about how sensitive species data are protected by Wildlife Insights.);*
- *Images of humans;*
- *Images whose identifications have not been confirmed by the data provider.*

Verifying data

All users contributing data to Wildlife Insights retain ownership of their data and are responsible for reviewing and editing the identifications made by the computer vision model in their project(s). While Wildlife Insights provides tools for users to process and edit data as needed, Wildlife Insights cannot guarantee the accuracy of identifications available to the public.

How to verify whether a record has been reviewed

Every download package from Wildlife Insights includes a record of the most recent identification associated with an image. The record will detail the identification (class, order, family, genus and species), individual animal details (age, sex), and the name of the most recent identifier, along with other metadata.

In many cases, the most recent identifier recorded will be the name of a user associated with that project. This means that the image has been reviewed and/or edited by that user. This information can be found in the images.csv file in the column titled *identified_by.* In the example below, the reviewer's name is *Nicole Flores*.

In other cases, the most recent identifier in the *identified_by* column is *Computer Vision*. This means that the image has been uploaded and has passed through the

F	M
identified_by	common_name
Nicole Flores	blank
Nicole Flores	Tiger
Computer Vision	Margay
Computer Vision	Human
Nicole Flores	Dark-winged Trumpeter
Nicole Flores	Dark-winged Trumpeter
Nicole Flores	Human

computer vision model, but has not been reviewed by a user. In the example to the right, you can see the Computer Vision model has predicted there is one image of a Margay and one image of a Human. These identifications have **not** been confirmed by a user and may be incorrect.

You can assess the accuracy of the computer vision results by referencing the images.csv file and the column titled *cv_confidence.* Read more about interpreting computer vision results on our About AI page.

Downloading images from private downloads

If you have been granted access to a private Project as an Owner, Editor, and Viewer, individual images can be easily downloaded from the Wildlife Insights website by clicking on the image in the *Identify* or *Catalogued* tab and then clicking on the *Download* button.

You can also bulk download the images from a Google Cloud Platform bucket (i.e, folder) by following the instructions below.

Prerequisites

To bulk download images from a private project, you will need:

- **A user account on the Google Cloud Platform (GCP):** You need a Google account to access the Google Cloud Platform. If you don't have one, you can create a free Google account at https://accounts.google.com/signup using the email address you used for your Wildlife Insights account. Your Google account must be associated with the same email you use for your Wildlife Insights account to access a project's bucket. To confirm this email, navigate to your *account details* page by clicking your name in the right-hand corner of the screen in Wildlife Insights.
- **Permission to access the project:** To bulk download images from a private project, you must have access to a private Project as an ***Owner***, ***Editor***, or ***Viewer***. *Contributors* and *Taggers* do not have permission to download images.
- **Install gsutil:** gsutil is an easy-to-use, Python application that lets you access Cloud Storage from the command line and is the best option for batch downloads. Install gsutil following these Google Cloud Platform instructions: https://cloud.google.com/storage/docs/gsutil_install

Downloading images using gsutil software

Step 1 - Request a download

- Request a private download (if you have permission to access the project). You can filter the data in the project or subproject based on your criteria (e.g., specific species) if you only need a subset of images.
- Once a request has been processed, you'll receive an email from Wildlife Insights with a link to download the dataset. You will have 24 hours to download the data bundle and

the images from the GCP bucket. The expiration date for downloading images is the same as the expiration date for the download link in the email.

- Download the download bundle, including an images.csv file with information you'll need to download images in Step 3.

Step 2 - Authenticate with Google Cloud

- Open Command Prompt (Windows) or Terminal (Mac/Linux)
- Run the command:

  `gcloud auth login`

- A web browser will open automatically to a sign in page. Sign in with your Google account that has access to the project.
- Return to your Command Prompt (Windows) or Terminal (Mac/Linux) - you should see a message you have logged in successfully.

Step 3 - Download images

Once you've installed gsutil, you can download images using these commands in your Command Prompt (Windows) or Terminal (Mac/Linux). You will need to:

- Replace the curly brackets `{ }` values with the relevant values from the location column in the *images.csv* file. Open the images.csv file in your download bundle and note `{Project}`, `{Deployment_ID}`, and `{Image}` values in each image file path in the *location* column as pictured below.

- Replace the `{DOWNLOAD_FOLDER}` with your desired download location (e.g., `C:\wi_project`) or `.` to download to your current folder. Note: Windows uses `\` for directory separators. Mac/Unix uses `/` for directory separators.

To download all images in a project:

```
gsutil -m cp -r  gs://{PROJECT} {DOWNLOAD_FOLDER}
```

Examples:

Windows	`gsutil -m cp -r  gs://camera_trap_project_main C:\wi_project` `gsutil -m cp -r  gs://camera_trap_project_main .`
Mac/Unix	`gsutil -m cp -r  gs://camera_trap_project_main data/wi_project` `gsutil -m cp -r  gs://camera_trap_project_main .`

To download all images of one specific deployment:

```
gsutil -m cp -r  gs://{PROJECT}/deployment/{DEPLOYMENT_ID} {DOWNLOAD_FOLDER}
```

Examples:

Windows	`gsutil -m cp -r  gs://camera_trap_project_main/deployment/2019497 C:\wi_project`
	
	`gsutil -m cp -r  gs://camera_trap_project_main/deployment/2019497 .`

Mac/Unix	`gsutil -m cp -r  gs://camera_trap_project_main/deployment/2019497 data/wi_project`
	
	`gsutil -m cp -r  gs://camera_trap_project_main/deployment/2019497 .`

To download a single image, replace `{IMAGE}` with the full file path URL in the *images.csv* file under the *location* column:

```
gsutil -m cp -r  gs://{PROJECT}/deployment/{DEPLOYMENT_ID}/{IMAGE} {DOWNLOAD_FOLDER}
```

Examples:

Windows	`gsutil -m cp -r  gs://camera_trap_project_main /deployment/2019497/a4db8e24-1b5b-439d-a00d-a904ba6b82e2.jpg C:\wi_project`
	
	`gsutil -m cp -r  gs://camera_trap_project_main /deployment/2019497/a4db8e24-1b5b-439d-a00d-a904ba6b82e2.jpg .`

Mac/Unix	`gsutil -m cp -r  gs://camera_trap_project_main /deployment/2019497/a4db8e24-1b5b-439d-a00d-a904ba6b82e2.jpg data/wi_project`
	
	`gsutil -m cp -r  gs://camera_trap_project_main /deployment/2019497/a4db8e24-1b5b-439d-a00d-a904ba6b82e2.jpg .`

To download a subset of images from a project or subproject:

If you need to a subset of images from a project or subproject, rather than all images:

- Filter the data in the project or subproject based on your criteria (e.g., specific species) when requesting your data download.

- Copy the full file path url (under the *location* column in the images.csv file) of the images you wish to download into a single text file with no header and one file path URL per row. Name the file as `paths.txt`.
- Run the following command in gsutil, replacing the `{PATH_LIST_FOLDER}` with the folder you saved your `paths.txt` file in (e.g., `C:\wi_project`) or `i` if you saved it in your current folder.

For Windows:

```
type {PATH_LIST_FOLDER}\paths.txt | gsutil -m cp -I {DOWNLOAD_FOLDER}
```

For Mac/Unix:

```
cat {PATH_LIST_FOLDER}/paths.txt | gsutil -m cp -I {DOWNLOAD_FOLDER}
```

Examples:

Windows	*`type C:\wi_project\paths.txt	gsutil -m cp -I C:\wi_project\jaguar_images`*  *`type .\paths.txt	gsutil -m cp -I .\jaguar_images`*
Mac/Unix	*`cat data/wi_project/paths.txt	gsutil -m cp -I data/wi_project/jaguar_images`*  *`cat paths.txt	gsutil -m cp -I .`*

Viewing images from public downloads

When you request a public download, you can view thumbnails of the images from the public dataset on the Wildlife Insights website. Currently, you cannot download images from public downloads. To view images from a public download:

- Request a public download. You can filter the data in the projectbased on your criteria (e.g., specific species) if you only need a subset of images. Once a request has been processed, you'll receive an email from Wildlife Insights with a link to download the dataset.
- Download the download dataset bundle, including an images.csv *file.*
- In the images.csv file, find the direct link to an image thumbnail in the *location* column and open it in your web browser. If you are not already, you must log in with your Wildlife Insights user account. The image link will be formatted like this:

  `https://app.wildlifeinsights.org/download/{download_id}/project/{project_id}/data-files/{data_file_id}`

Note: Only images whose identifications have been confirmed by the data provider can be viewed.

Licenses

Creative Commons provides standardized licenses and guidelines that make it easy to share work and for users to provide attribution to the work. Each project in Wildlife Insights can be licensed under the following options (please read the Terms of Use for full details).

- Images (recorded data) can be licensed under **CC0, CC BY** or **CC BY-NC**.
- Metadata can be licensed under **CC0** or **CC BY**.

These licenses are described below:

- **Creative Commons Zero (CC0)** permits a user to share, adapt and modify the work, even for commercial purposes, without providing attribution (summary, full legal text);
- **Creative Commons Attribution 4.0 (CC BY),** permits a user to share and adapt material with appropriate attribution, including for commercial purposes (summary, full legal text);
- **Creative Commons Attribution-NonCommercial 4.0 (CC BY-NC)**, permits a user to share and adapt material with appropriate attribution, only for noncommercial purposes (summary, full legal text).

Citing data downloaded from Wildlife Insights

All projects licensed under **CC BY** or **CC BY-NC** require a data user to provide attribution. Wildlife Insights makes it easy to provide attribution by providing a list of data citations for each project in a download request. Projects in Wildlife Insights are also assigned an Archival Resource Key (ARK), which is a persistent, permanent link to a project and dataset. Please refer to the projects.csv file to view the recommended citations and licenses for the projects in this download.

Wildlife Insights suggests citing projects using the following format:

References:

> Author(s) (Year accessed from Wildlife Insights). Project name. DOI. Accessed via Wildlife Insights on dd-mm-yyyy.
>
> *Example:* Ahumada J, Schipper J (2020). Cafe Fauna. https://n2t.net/ark:/12345/bcd987 accessed via Wildlife Insights on 03-10-2020.

In-text citations:

> *Example:* Data used in this study were accessed from Wildlife Insights on *date* (Ahumada & Schipper, 2020)

For additional information on citations, Wildlife Insights recommends reading GBIF's citation guidelines.