

# AI-Taxonomist

## Training & capacity building resources

Cos4Cloud Services

### About the AI-Taxonomist system and user guide

This is a system and user guide to [AI-Taxonomist](#), a Cos4Cloud service which enables citizen observatories, citizen science or other research projects to integrate customisable automatic identification tools and help users identify species from images.

### Description

This is a user guide to **AI-Taxonomist**, an online service which enables citizen observatories, citizen science or other research projects to integrate automatic identification tools, adapted to their needs, supporting identification of particular groups of species from images.

**AI-Taxonomist** has been developed by [Inria](#) as part of [Cos4Cloud](#) a European Horizon 2020 funded project. The Cos4Cloud project has developed [thirteen services](#) boosting citizen science technological services to help increase and improve the quantity and quality of observations.

This is one of three Cos4Cloud services developed by [Inria](#), for existing and new citizen science initiatives providing artificial intelligence (AI)-based Species Photo ID and AI-based Biodiversity Prediction as tools and resources for biodiversity. Inria is one of the research organisations behind [Pl@ntNet](#) - a citizen observatory with an extensive database that uses AI to help people identify and better understand plants.

### Service Coordinator



### Cos4Cloud Coordinator

*The following content has been provided to help guide users of this service.*



### WHAT IS AI-TAXONOMIST?

**AI-Taxonomist** can be used as a web-based API (application programming interface) or through a web GUI i.e., a web-based application. It uses biodiversity search engines i.e. the Global Biodiversity Information Facility (GBIF), operating under the [GBIF Data user agreement](#). **AI-Taxonomist** was developed to search a set of species that occur in GBIF, which is the default option, however it can work with any other image database.



### WHO IS THIS SERVICE FOR?

**AI-Taxonomist** is an online service for citizen science platforms, projects and citizen observatories (COs) or other research projects, that supports the integration of automatic identification tools for species identification.

Based on machine learning, it includes the option to search and find similar results searching images of species that potentially resemble an observation for example, those made by citizen scientists.

**AI-Taxonomist** provides a toolkit for developers or data scientists with a good background in data engineering and AI technologies who want to use this web service to provide additional identification tools for their users. For example, a museum developing an app related to a particular group of species could use this service and integrate a photo-identification tool for that species to help their users with identification.



### WHAT ARE THE BENEFITS OF THIS SERVICE?

This service provides innovation for citizen observatories:

- It facilitates the integration of AI-based species identification in citizen observatories.
- It allows users to automatically identify species from images, using AI.
- It delivers search queries presenting the most similar or likely options from the GBIF database.



### HOW DOES IT WORK?

**AI-Taxonomist** integrates automatic identification tools customisable to a citizen observatory or citizen science project or platform needs, helping users identify species from pictures:

- The service provides a toolkit that facilitates customising to particular requirements i.e. particular groups of species. The toolkit is based on the deep-learning-based similarity search engine already used in the *Pl@ntNet* platform extending this to other species groups.
- This service allows users to identify observations submitted with multiple images demonstrating different parts of the same individual organism (i.e. for plants: leaf, flower, fruit, etc.).
- It performs the search then returns the most similar observations in the database for each of the identified species, thanks to a scalable similarity-search engine that relies on high-dimensional data hashing and deep representations.
- Based on machine learning, the identification toolkit includes a search to find similar results, i.e. amongst photographs of species that potentially resemble the observation made by a citizen scientist.
- **AI-Taxonomist** can establish a list of species of interest. For example, if a citizen observatory monitors bees, **AI-Taxonomist** can be integrated to help users automatically identify different bee species.



### HOW TO USE AI-TAXONOMIST – A STEP BY STEP GUIDE

Once applied **AI-Taxonomist** allows users to define a list of species of interest and uses a system that can be queried which operates through:

1. a web-based API (Application Programming Interface) i.e. GBIF and
2. a web GUI (Graphical User Interface) i.e. a web-based component.

Using illustration images licensing, **AI-Taxonomist's** backend and frontend are under the MIT open-source licence.

## 01

### WEB-BASED API: BUILD AN IDENTIFICATION ENGINE FROM A GBIF QUERY :

#### Required operational environment and setup

- Follow the step by step process [here](#).

#### ai-taxonomist

(almost) Automatically generate a Pl@ntNet like identification engine from a GBIF occurrences Darwin Core Archive

By using this tool, the user agree to the GBIF data user agreement (<https://www.gbif.org/terms/data-user>)

#### Required environment

- python 3.9
- docker >=19.03.12
- Cuda capable GPUs
  - cuda >=11.0,
  - nvidia-docker >=1.0.1  
(<https://docs.nvidia.com/datacenter/cloud-native/container-toolkit/install-guide.html#docker>)
- github.com access
- GitHub docker register access

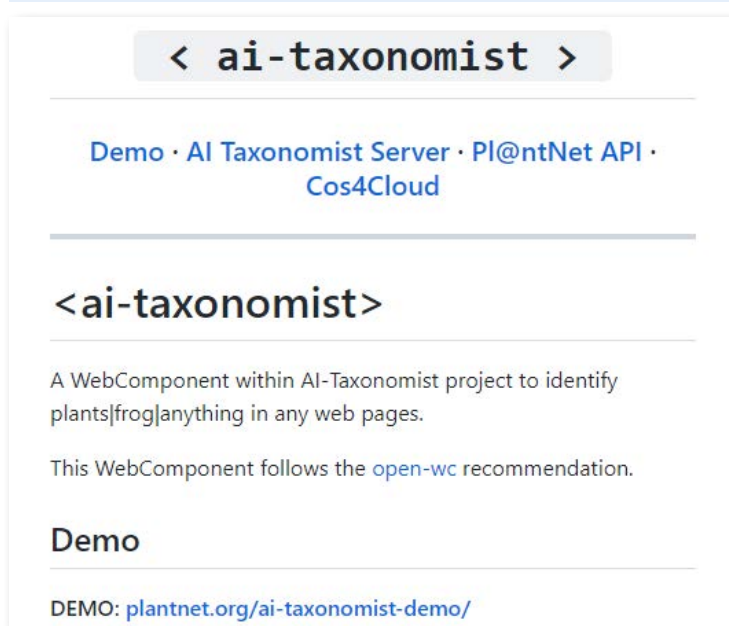
- This allows a set up which automatically generates a *Pl@ntNet-like* identification engine from a GBIF dataset for different types of species.
- By using this tool, users agree to the [GBIF data user agreement](#).

## 02

### WEB-BASED API: USE AI-TAXONOMIST ONLINE TO IDENTIFY DIFFERENT SPECIES:

Use the web-based version of AI-Taxomist powered by the **Pl@ntNet-API**

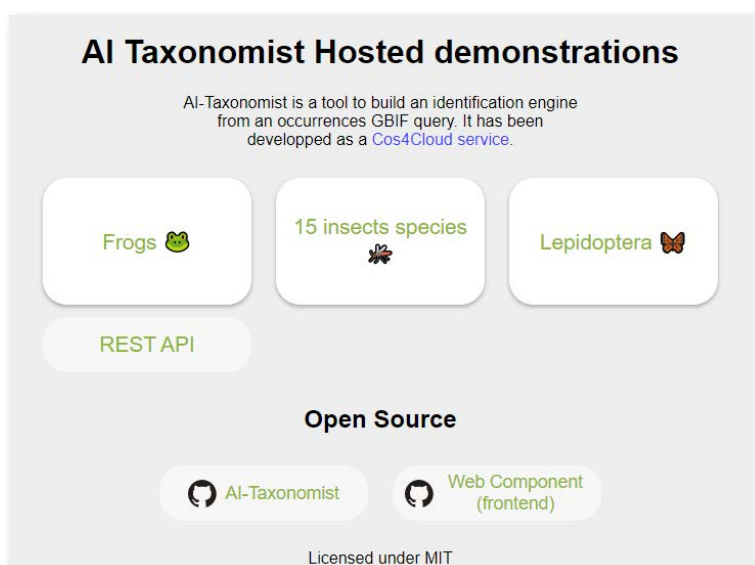
- Follow the step by step process [here](#).



- This presents a web-based API version of **AI-Taxonomist** service which facilitates identification of different species built on the underlying Pl@ntNet-API service.
- an API Key is required for integration: get the [Pl@ntNet-API key here](#).

#### Check out AI-Taxonomist

- Access [hosted demonstrations](#) and test the functionality.





### INFORMATION AND FURTHER SUPPORT

**Introduction to this Cos4Cloud service:** <https://cos4cloud-eosc.eu/services/ai-taxonomist/>

**AI-Taxonomist online – try the demos:** <https://ai-taxonomist.plantnet.org/>

**Build an identification engine based on GBIF queries:** <https://github.com/plantnet/ai-taxonomist#readme>

**Back end and code repository:** <https://github.com/plantnet/ai-taxonomist#readme>

**Web component:** <https://github.com/plantnet/ai-taxonomist-webcomponent>

**Pl@ntNet-API documentation can be found here:** <https://my.plantnet.org/doc/openapi>

**About Pl@ntNet:** <https://plantnet.org/en/>

**Get the Cos4Cloud infographic for this service:** [AI-Taxonomist infographic](#)

**Additional resources:** [About the Cos4Cloud Toolbox and Evidence Hub](#)

This system and user guide is a training resource for AI-Taxonomist, one of the technological services co-designed by the Cos4Cloud project (<https://cos4cloud-eosc.eu/>). **Title: AI-Taxonomist System and User Guide.** Service Coordinator: Inria. This guide is one of the Training and Capacity Building resources in the Cos4Cloud Toolbox & Evidence Hub developed by the Open University in collaboration with project partners. Contact: [cos4cloud-toolbox@open.ac.uk](mailto:cos4cloud-toolbox@open.ac.uk).



This project is part of:

**EUROPEAN OPEN  
SCIENCE CLOUD**



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 863463.