



The Lab of CLEF dedicated to  
biodiversity data

## LifeCLEF 2020

**Alexis Joly** (INRIA, LIRMM), **Henning Müller** (HES-SO), **Hervé Goëau** (CIRAD, AMAP), **Stefan Kahl** (Chemnitz University of Technology), **Pierre Bonnet** (CIRAD, AMAP), **Hervé Glotin** (University of Toulon, LSIS CNRS), **Willem-Pier Vellinga** (Xeno-Canto), **Fabian Robert Stoeter** (Inria, LIRMM), **Andrew Durso** (University of Geneva), **Maximilien Servajean** (University of Montpellier), **Benjamin Deneu** (Inria, LIRMM), **Christophe Botella** (INRA, Inria, AMAP)



# Four tasks



Task 1 - PlantCLEF: cross-domain plant identification



Task 2 - BirdCLEF: bird species detection and separation in audio soundscapes



Task 3 - GeoLifeCLEF: location-based prediction of species based on environmental and occurrence data



Task 4 - SnakeCLEF: image-based snake identification

**NEW**



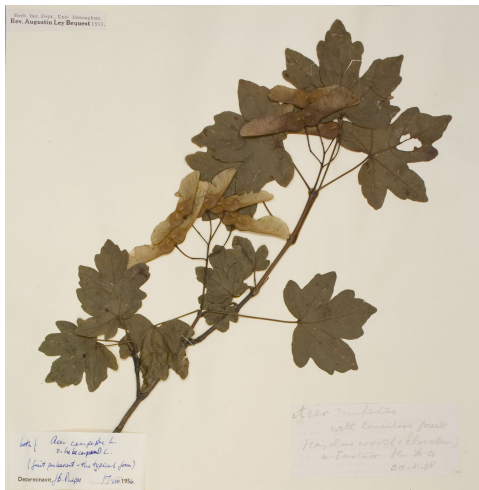


# PlantCLEF 2020

## Cross-domain plant identification

**Scenario:** Predict plant species in pictures based on a training set of herbarium sheets

- Herbarium sheets are the only available training data for many species
- A difficult cross-domain classification task (drying, pressing, ageing, etc.)





# PlantCLEF 2020

## Cross-domain plant identification

**Data:** 100K herbarium sheets & 10K plant pictures

- Herbarium: eRecolNat, iDigBio
- Pictures: eRecolNat, Pl@ntNet, EoL

TRAINING SET



TEST SET



VALIDATION SET





# BirdCLEF 2020

## Bird detection in soundscapes

**Scenario:** Predict the list of species that are audible in a 5-second segment of a soundscape recording.

### Training data (~75,000 audio files):

- Mono-species recordings + metadata from Xeno-canto
- ~800 Classes (South & North America, Central Europe)

### Test data (~20 days of audio):

- Colombia and USA soundscapes from 2019
- Previously unreleased test data from the USA
- New soundscapes from Germany with expert labels





# BirdCLEF 2020

## Bird detection in soundscapes

### Rules:

- Train on mono-species recordings only
- Test on soundscapes only
- Validation data must not be used for training
- No model ensembles

### Metrics:

- rMap and cMap as in 2018 & 2019
- F-measures (F1, F0.5)
- We are open for input from participants





# GeoLifeCLEF 2020

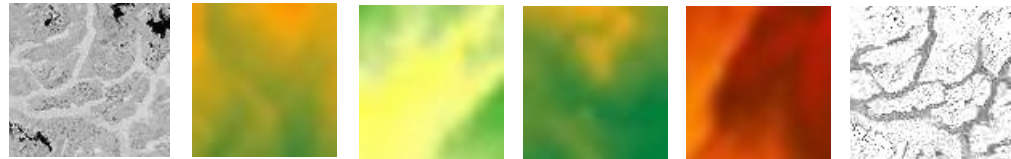
## Location-based species recommendation

**Scenario:** Predict the list of species that are the most likely to be observed at a given location

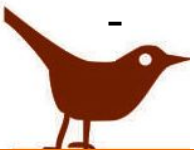
**Data:** Biodiversity occurrence data (e.g. 1M) associated to multi-modal environmental images



**1 occurrence of *Malva Silvestris***



Channels: climatic data, elevation, soil occupation, satellite, etc.





# SnakeCLEF

## Image-based snake identification

### Scenario:

- Predict snake species in photos taken in the wild
- over half a million victims of death & disability from venomous snakebite annually

**Data:** 187K images of 85 species, with geographic information at the continent and country level

- Pictures: iNaturalist, HerpMapper, Flickr, IndianSnakes.org
- Can be divided into training & testing as desired
- Other, more private testing data are available for later validation



**UNIVERSITÉ  
DE GENÈVE**

