

CloudedBats.org

Free software for open data

Arnold Andreasson. Arnold Andreasson Konsult AB, Sweden. info@cloudedbats.org
Poster at EBRS 2017, the 14:th European Bat Research Symposium, 1-5 August 2017 in Donostia.



Main problem to solve

We need more data about bats from all regions where they occur, and it needs to be collected soon if we want to document the current situation before it starts to change more rapidly.



One solution to start with

One possible solution, among others, is to record and store sound in full spectrum or time expanded formats for future analysis when better tools and reference libraries of bat sounds are available.



Problems that must be handled

- Recording units for high quality recordings are expensive.
- It is time consuming to analyse recorded sound from many detectors.
- It is hard to find reference recordings to be used for automatic and computer assisted analysis.
- Species identification algorithms must be better.
- The process of sharing and publishing data and results is complicated.

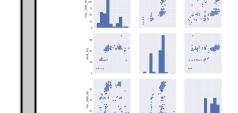


The CloudedBats project

CloudedBats.org is a spare time/hobby project where I am trying to use my skills in software development to address these problems. My motivation is that I want to learn more about bats and to get more practice when using modern software techniques.

The programming language Python is used for all parts and the code is published at GitHub.com.

The CloudedBats software is completely open and free and you can use it as you want, even as a part in a commercial product. More info can be found here: <http://cloudedbats.org/>



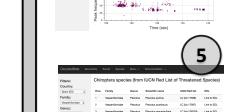
The dataflow for CloudedBats in 10 steps

My goal for the CloudedBats project is to develop free and open software for the whole chain from the ultrasonic microphone to quality approved results and sound files published and stored for future use. Some parts of the code can be used directly and some parts are more on an experimental level. Hopefully, the CloudedBats software can be used as a template/starting point for other biological monitoring projects, and also as an inspiration to produce and publish more open and free data.



1. Microphone

Support for high quality ultrasonic microphones is essential. Pettersson microphones have been used for many years by professionals. **Status:** Support for Pettersson M500 (500 kHz) and M500-384 (384 kHz). Other brands of microphones with USB for Linux may also work.



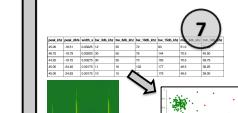
2. Detector for passive monitoring

Standard hardware can be used, except for the microphone. **Status:** Working software for passive monitoring. **Hardware:** Microphone (350€+VAT), Raspberry Pi 3B computer (40€), case (10€), Micro-SD for software (15€), USB memory for sound files (15€), GPS (20€). Any power supply for smartphones can be used, for example a mobile phone charger or a powerbank (10-30€).



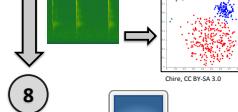
3. DSP, Digital Sound Processing

DSP is needed at different steps in the data flow. **Status:** Basic DSP is working. Experimental work in progress on extracting different sets of metrics from pulses and sequences.



4. Presentation and visualisation

Web applications makes it more easy to work together compared to software running on personal computers. **Status:** Web application for testing is up and running with early development prototype for visualisation: http://test.cloudedbats.org/bat_activity



5. Species lists

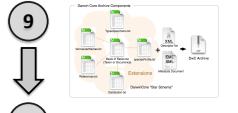
A species list to be used as a taxonomic backbone is needed. **Status:** CloudedBats uses the IUCN Red List of Threatened Species for this: <http://test.cloudedbats.org/species>



6. Reference libraries for bat sound

It is not possible to train computers to identify bats without reference libraries of sound. They are difficult and expensive to develop and must probably be based on caught and marked bats.

Status: So far I have only a limited set of reference recordings. Nothing from outside Europe.



7. Species identification

This seems to be nearly impossible for some species, even for the limited amount of species found in Europe. But the development of new machine learning algorithms takes place quickly.

Status: Have started to look into Python software libraries for machine/deep learning.



8. Storage and processing in the cloud

Large inventory activities of bat sound will produce a huge amount of data. **Status:** Have identified the need for storage and processing capacity beyond standard desktop computers.



9. Survey results, reporting formats

Inventory reports, scientific articles and data papers contains a lot of important data that has better impact if it is spread. **Status:** Working on support for the format DarwinCore-Archive.



10. Storage for future generations

Natural history museums have digital collections and uses DarwinCore-Archive for biological data exchange.

The map shows what they have stored for bats (Chiroptera).

Status: The map does not accurately reflect the bat species abundance distribution around our world.

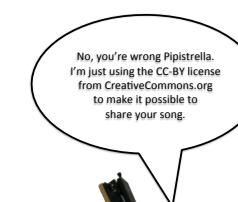


gbif.org/species/734



I'm Pipistrella, the queen of darkness.
But the stupid man below think he's the owner of my song. *)

Facts:
*) Pipistrella's song is a derived work, originally created by her ancestors more than 30 million years ago. This means that the creator has been dead for more than 70 years and the work should be assigned as "public domain".



No, you're wrong Pipistrella.
I'm just using the CC-BY license from CreativeCommons.org to make it possible to share your song.

