# CS486C – Senior Capstone Design in Computer Science

## Project Description

| Automatic data verification for community science contribution to conservation monitoring | |
|---|---|
| NAU NORTHERN ARIZONA UNIVERSITY | **Dr. Camille Gaillard,** Post-Doctoral Researcher in Ecological Informatics<br>School of Informatics, Computing, and Cyber Systems<br>Northern Arizona University<br>Camille.Gaillard@nau.edu<br>**Dr. Chris Doughty, Associate Professor** in Ecological Informatics<br>School of Informatics, Computing, and Cyber Systems<br>Northern Arizona University<br>Chris.Doughty@nau.edu<br>**Dr. Duan Biggs**, Associate Professor & Olajos Goslow Chair<br>School of Earth and Sustainability<br>Northern Arizona University<br>Duan.biggs@nau.edu |
| | |

# Project Overview:

1 – Context

In December 2022, 188 governments of the world agreed to conserve 30% of the land and sea globally by 2030. This agreement, the Kunming-Montreal Global Biodiversity Framework, or 30x30, came as a new addition to the Convention on Biological Diversity (CBD). This is an ambitious step for the international community and local populations across the globe. While being a major progress towards the preservation of biodiversity, and the ecosystem services it provides, it comes with multiple challenges. Equity and inclusivity are major concerns for populations directly impacted and engaged in conservation areas, as conservation has a challenging history in this regard[1]. Ensuring the achievement of conservation targets requires intensive monitoring of essential biodiversity variables.

Citizen scientists contribute valuable data for academic research and conservation. Enhancing equity and inclusivity can expand its reach and value.

2 – Our team work

In a recent paper[2], we proposed to answer both concerns by turning them into an opportunity for populations directly impacted by conservation. Specifically, we advocate for advancing the monitoring of biodiversity for conservation targets by relying on equitable and inclusive citizen science (also known as community science). Citizen science is the method by which non-academic persons can contribute to the advancement of science and knowledge by their observations. By pairing remote sensing with citizen science, we can ground-truth satellite data more accurately and

1 Blanc, Guillaume. *The invention of green colonialism*. John Wiley & Sons, 2022.
2 https://doi.org/10.1038/s41893-024-01309-7
3 https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0226534
4 https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0295298
5 https://www.mdpi.com/1424-2818/16/1/42

expand the use of these products. Citizen science observations can also provide information about parameters not observable with remote sensing.

3 – Key concerns

Several citizen science apps have been deployed, such as eBird, iNaturalist and Globe. While they are formidable tools in their own regard, we identify two components which could enhance their contribution to achieving the 30x30 conservation targets. First, their reliability is still debated and could be improved[3,4,5]. For example, they might be mismatched with other sources of data, e.g. satellites, which is a challenge for validation. Second, citizen scientists are not incentivized to contribute to standardized biodiversity monitoring through these apps. Though experiments are being conducted on this question, it will raise specific concerns.

4 – Broader Solution

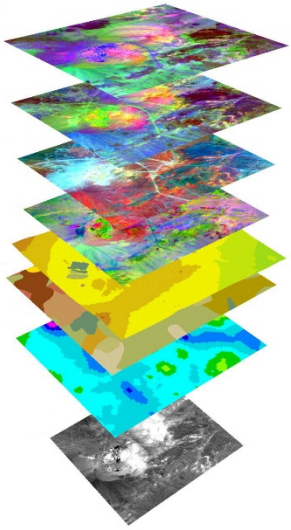
This first item is critical to ensure the usefulness of the collected data. From a conservation and ecology perspective we need to verify data and metadata content, integrity, quality and authenticity. This will allow to accurately monitor conservation target through time and contribute to reporting to the 30x30 agreement of the CBD. Further, considering financial incentivization of citizen scientists, ensuring the accuracy of the contributed data is critical to a fair payment system.

Example of camera trap acquired in Flagstaff's forest, to be analyzed for species identification

5 – Key features

Several open-source image identification tools are available to detect and count species and individuals in camera trap images. However, they might perform differently depending on the species and area. The first step will be to automatically identify which tool is best suited for any camera trap image. The second step will be to use metadata to improve the accuracy and informativity of the analysis. Using GPS location of the camera trap, we can cross check with distribution range map, such as the ones provided by the IUCN (International Union for the Conservation of Nature). This requires verifying metadata quality, for example by comparing displayed weather with recorded weather for the area. The last step will be to store and organize the output of these image analyses.

While this tool will be developed based on camera trap data, it should be ready to plug-in into existing citizen science apps.



Layers of data allow to give an accuracy ranking to each contributed data point. For instance, a species recorded outside of its bioclimatic range is less likely to be correct.

6 – Overall Impact

With this plug-in ready module, we will be able to implement a pilot citizen science project to evaluate the feasibility of verifying contributed data and metadata, based on existing or newly developed apps. Successful outcomes will enable the mainstreaming of payment and verification in citizen science apps. By deploying, at global scale, our proposed solution for monitoring essential biodiversity variables, and engaging local populations as citizen scientists for conservation, we will foster equity and inclusivity towards global conservation targets.

**Some of the key functions supported by this product should include:**

- Minimum Viable Product: An offline tool for automatic animal species identification and verification, able to be plugged in into existing citizen science apps.
    - It will be able to ensure that data points are congruent with existing data relatively to the specific parameters of the contributed data, *e.g.* species identified on a photo match with known distribution range of said species, or time of day matches with known behavior of said species.
    - It should be able to flag image quality and avoid duplicates.

- Complete Product: In addition to the above functionality the final product should integrate:
    - A data storage and management feature. Processing thousands of hours of camera trapping is a challenge in terms of image analysis, but also in terms of processed data management. This should be handled considering storage space efficiency and user friendliness. Ideally, processed data should be made available as excel/csv files and as maps, including species names, abundance biomass, and activity (time).
- Stretch Goals: Additional features would take this software/module to the next step.
    - A point system: Data are attributed value or points based on relevant criteria, such as rarity of the data point (place, time, species) and quality of the data point.
    - Identification of animal metrics, such as size (height, weight) or behavior (rest, movement, feeding)
    - Duplicate flagging expanded to pseudo-duplicates; *i.e.* two camera trap next to each other capturing the same thing with a different angle.
    - Porting the app from computer based to app based (iOS, Android).

***A note on not reinventing the wheel…*** To our knowledge there is no available tool providing all the outlined. What is unique about this project is none of its sub-components, but their integration and their function and goal. Thus, we encourage you to use existing code and can provide support from developers of existing citizen science apps.

# Knowledge, skills, and expertise required for this project:

- This project is to be developed in coordination with another project "Data integrity and abuse prevention to enable equity in community science for conservation", as they both contribute to a common end-goal.
- Skills/knowledge in image recognition, data management and automated processes (i.e. to automatically compare submitted data to existing databases)
- Coding language can be flexible
- Skills/knowledge in mobile app development (iOS or Android)

# Equipment Requirements:

- No specialized software or equipment should be required for this project, beyond a standard software development stations and free IDEs, frameworks, and other tools.
- Camera trap data will be provided.