**Plan Overview**

*A Data Management Plan created using DMPTool*

**DMP ID:** [https://doi.org/10.48321/D1C08D](https://doi.org/10.48321/D1C08D)

**Title:** Projeto ForestEyes - Cidadãos monitorando desmatamento em Florestas

**Creator:** Alvaro Fazenda - **ORCID:** [0000-0002-4052-1113](https://orcid.org/0000-0002-4052-1113)

**Affiliation:** Universidade Federal de São Paulo (unifesp.br)

**Principal Investigator:** Alvaro Luiz Fazenda, Fabio Augusto Faria

**Data Manager:** Hugo Resende, Eduardo Bohid Neto

**Project Administrator:** Hugo Resende

**Contributor:** Hugo Resende, Eduardo Bohid Neto, Amanda Sales, Marília Alves Silva, Fernanda Beatriz Jordán Rojas Dallaqua

**Funder:** São Paulo Research Foundation (fapesp.br)

**Template:** Digital Curation Centre

**Project abstract:**

The conservation of Tropical Forests is a key issue due to their importance in the global ecosystem since they contain great biodiversity, carbon storage, rainfall regulation in adjacent regions, and indigenous peoples sheltering.

Unfortunately, millions of hectares of tropical forests are lost each year through deforestation and degradation for different and complex reasons.

In this research project, a proposal to enhance a novel monitoring system prototype, called ForestEyes, is presented, which aims to combine citizen science and machine learning in detecting deforestation. Citizen Science allows ordinary volunteers to collaborate by inspecting remote sensing image segments over a forest region, seeking to detect deforested or degraded areas, thus participating in a scientific experiment aimed at improving the planet's sustainability, also acquiring more awareness and
involvement with the major issue. Machine Learning techniques will be applied in a forest area by using trained databases with samples labeled by volunteers in the previous phase, expanding the coverage for the detection area. Initial experiments over six official campaigns, demonstrated through solid published scientific works, obtained more than 81,000 contributions from 644 different volunteers, allowing to build of a solid training base for classification models, whose results were compared with the official program for monitoring the Brazilian Legal Amazon (PRODES). The methodology developed proved to be promising, both from the point of view of the volunteer's accuracy and the classification techniques applied. This project proposes improvements to the aforementioned project by expanding the application's database and investigating the use of new remote sensing images from multiple optical and radar sensors (SAR), and different regions and periods of time, which implies scientific and technological challenges in the pre-processing, citizen science and machine learning phase. It is also intended to improve the remote sensing image segmentation process, as well as apply super-resolution techniques to the images.

Start date: 04-02-2019

End date: 12-19-2026

Last modified: 10-19-2023

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal.
Data Collection

What data will you collect or create?

Remote Sensing Images from satellites, normally in GEOTIFF format.

Mask images created applied to original images.

Segmented images to be investigated by citizen science and machine learning.

How will the data be collected or created?

Original data will be collected directly from institutional repository

Other images will be created with a GIS or segmented algorithms by the project team.

Documentation and Metadata

What documentation and metadata will accompany the data?

Geolocation for every image.

ID for each segment of image.

Labeled image pixels and segments for the ground truth.

Ethics and Legal Compliance

How will you manage any ethical issues?

The identity of participants volunteers will not be registered.

All acquired or created images have free access.

How will you manage copyright and Intellectual Property Rights (IP/IPR) issues?

All acquired or created images have free access, but in every case the original authors will be depicted.

Storage and Backup
How will the data be stored and backed up during the research?

Data space in public clouds will be used for backup.

How will you manage access and security?

There are no sensible data, but all of them will be protected by the storage platform used.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

There are no need to destroy the data periodically.

What is the long-term preservation plan for the dataset?

The data/images needs to be stored by a decade, for research purposes.

Data Sharing

How will you share the data?

The data could be shared to other similar projects

Are any restrictions on data sharing required?

Question not answered.

Responsibilities and Resources

Who will be responsible for data management?

The researchers and principal investigators will be responsible to implement the DMP

What resources will you require to deliver your plan?

There are no extra needs to implement the plan
Planned Research Outputs

Image - "Classified segments of images by volunteers in a citizen science platform"

Labeled images by volunteers using zooniverse platform.

Image - "Classified segments of images by machine learning"

---

Planned research output details

<table>
<thead>
<tr>
<th>Title</th>
<th>Type</th>
<th>Anticipated release date</th>
<th>Initial access level</th>
<th>Intended repository(ies)</th>
<th>Anticipated file size</th>
<th>License</th>
<th>Metadata standard(s)</th>
<th>May contain sensitive data?</th>
<th>May contain PII?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classified segments of images by volunteers in a citizen science pla ...</td>
<td>Image</td>
<td>Unspecified</td>
<td>Open</td>
<td>None specified</td>
<td>Creative Commons Attribution 4.0 International</td>
<td>None specified</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Classified segments of images by machine learning</td>
<td>Image</td>
<td>Unspecified</td>
<td>Open</td>
<td>None specified</td>
<td>Creative Commons Attribution 4.0 International</td>
<td>None specified</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>