Challenges in data and software management and sharing at the global level: the case of the PARSEC project.

Romain David, Corrêa Pedro Luiz Pizzigatti, Rorie Edmunds, Danton Ferreira Vellenich, Ana Heredia, Jeaneth Machicao, Laurence Mabile, Yasuhiro Murayama, Margaret O’Brien, Solange Santos, et al.

To cite this version:


HAL Id: hal-03014752
https://hal.science/hal-03014752
Submitted on 19 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
PARSEC objectives:

Building New Tools for Data Sharing and Reuse through a Transnational Investigation of the Socioeconomic Impacts of Protected Areas

The PARSEC project is designed to provide a unique opportunity for data and synthesis science to collaborate and exchange in real-time toward the goal of improving research outcomes, data sharing, and data reuse. It will also help pioneer new scientific and data science technologies aimed at improving both the management and conservation of global biodiversity.

We have two interactive teams in this project: a Synthesis Science team and Data Science team.

- The Synthesis Science team is employing artificial intelligence techniques to analyse satellite images and socio-economic information to better predict and mitigate the effect(s) of the creation of Protected Areas that may potentially threaten the livelihoods and health of local (indigenous) communities. Like most researchers who investigate complex environmental problems, the team depends significantly on the availability of high quality, spatially dispersed, multidisciplinary, and time-series data.

- The Data Science team of leading environmental data management and conservation of global biodiversity experts is partnering with a multi-country, spatially dispersed, multidisciplinary, and time-series data communities (RDA, ESIP), societies (AGU), and infrastructures for data attribution (e.g., DataCite and ORCID) working to develop leading practices on data citation, attribution, credit, and reuse. Using an integrated work with the synthesis science team as a case study of the wider scientific community, the data-science team will provide a review of best practices for data management and stewardship to optimise data access and reuse. The team will also develop and implement a new tool to better track data usage and reuse.

Dealing with international complexity

PARSEC comprises over 30 team members from six countries, speaking 4 languages. There are 4 Pls, and a lead researcher representing each country. PARSEC team is disciplined diverse, with ecologists, statisticians and modelers, economists, spatial scientists, data scientists and experts in artificial intelligence. An overarching challenge for the project is sharing what we learn and develop with the broader research community. Engaging with working groups within the RDA is one way to get additional feedback.

Opportunities to work directly with research teams, such as those funded by our grantor, the Belmont Forum, will be an important first step. This project articulates the technological challenges prompted by the many aspects of team diversity and physical location as well as the cultural challenges. We are seeking better understanding and we embrace the value of different team dynamics, methods of communication, and preference to research approach.

Planned outcomes

- To be able to determine the influence of natural protected areas on consumption expenditure and asset health of rural communities in a range of countries;
- To have refined recommendations for the research data workflow and skills for research teams;
- To have developed tools for researchers to view how the data they have deposited are used and cited. This should improve the number of citations to data sets and better attribute them to the data creator;
- To have contributed to the promotion of FAIR, CARE and TRUST through authentic examples from an active research project.

You reap what you sow!

A garden that is planned, cultivated and managed well will yield more produce than one where you chose the wrong plants, forgot to mark what were planted where, didn’t fertilise or water regularly, and allowed the weeds to take over. A garden that is planned, cultivated and managed well will yield more produce than one where you chose the wrong plants, forgot to mark what were planted where, didn’t fertilise or water regularly, and allowed the weeds to take over.