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 scientists 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 professionals represents data communities (RDA, ESIP), societies (AGU), and infrastructures for data attribution (e.g., DataCite and ORCID) are working to develop leading practices on data citation, attribution, credit, and reuse. Using an integrated approach 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 management usage and reuse.

Dealing with international complexity

PARSEC comprises over 30 team members from six countries, speaking 8 languages. To date there are 4 PIs, and a lead researcher representing each country. PARSEC team is diverse, with ecologists, statisticians and modellers, 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 step. This provides the context to understand characterise and engage the technological challenges. 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 management usage and reuse.

PARSEC roadmap proposal to meet data strand objectives

Step 1: Inventory success stories / landscape / good practices / tools / RDA or other recommendations

Step 2: Collect the needs of researchers / how to support them

Step 3: Synthesize practices to researchers / validation by the whole community / material development

Step 4: Train and support research teams, adapt best tools to researchers needs

Step 5: Assess efficiency to resolve issues and ensure sustainability of applied practices

Data Strand in PARSEC

In the PARSEC project, data science experts are partnering with a multi-country synthesis science research team to build relevant tools and processes for better data and software management that integrate into the research lifecycle and are to be shared with the wider research community. In recent months, we have seen important papers and books published that support this larger goal of helping researchers incorporate data and software management during their whole lifecycle, but there is still much work to be done to develop the methods and foster the cultural change needed to implement them in practice across scientific communities. The main challenges identified so far can be grouped into: technological ones and cultural ones. We would like to share these considerations and work collaboratively on incremental improvements that would be beneficial for all.

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 seeds 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 seeds were planted where, didn’t fertilise or water regularly, and allowed the weeds to take over.

Like watering and weeding, a little, consistent work throughout the research life cycle will result in data that are of value to others and able to be shared as part of any research output. In PARSEC, we value scientific data as an important contribution to scholarly work. We are taking careful and measured steps towards ensuring our data are well managed, FAIR and capable of having an enduring role in the scholarly ecosystem beyond the lifetime of our project.

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 contribute to the promotion of FAIR, CARE and TRUST through authentic examples from an active research project.