Credit for Data Reuse Is Driven by Making Data FAIR: The PARSEC Project Approach

Shelley Stall[1], Alison Specht[2], Margaret O'Brien[3], Anne Cambon-Thomsen[4], Laurence Mabile[4], Yasuhiro Murayama[5], Pedro Luiz Pizzigatti Corrêa[6], Helen Glaves[7], Helena Cousijn[8], and Eric Olson[9]

[1]American Geophysical Union
[2]University of Queensland
[3]University of California Santa Barbara
[4]University Paul Sabatier Toulouse III
[5]National Institute of Information and Communications Technology
[6]University of São Paulo
[7]British Geological Survey
[8]DataCite
[9]Center for Open Science

November 24, 2022

Abstract

Research data are a vital component of the scientific record. Discovering and assessing data for possible reuse in future research is challenging. The Belmont Forum has recently awarded funds to three international teams as part of a four-year Collaborative Research Action (CRA) on Science-driven e-Infrastructure Innovation (SEI) for the Enhancement of Transnational, Interdisciplinary and Transdisciplinary Data Use to improve data management practices that will increase data reuse. One of these awardees, PARSEC, comprises two interwoven strands, one focused on improving data practices for reuse and credit, and one for synthesis science. The data specialists work alongside synthesis science researchers as they determine the influence of natural protected areas on socioeconomic outcomes for local communities. They collaborate with the researchers to better understand their motivations and work practices, and to aid them in the data-related steps that need to be taken during the research lifecycle. This will ensure their data and code are FAIR-compliant and thus enhance the likelihood of their data being reused and their analyses reproducible. The PARSEC team is working with Research Data Alliance (RDA), Earth Science Information Partners (ESIP), DataCite and ORCID to build awareness of the elements required for data creators to receive credit and automated attribution for their data contributions, and the tools that will make it easier to observe usage. Credit for data is an important incentive for researchers to make their data reusable. When data are FAIR and cited, their related publications have higher visibility. We shall discuss various ways in which we are working across the science-data interface in our multi-country and multi-disciplinary working environment to improve data (and code) reuse through better management and crediting. Make your Data FAIR, Cite your Data, Get Credit, Increase Reuse and reap the rewards!

Posted on 24 Nov 2022 — CC-BY 4.0 — https://doi.org/10.1002/essoar.10502388.1 — This a preprint and has not been peer reviewed. Data may be preliminary.

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

Shelley Stall, American Geophysical Union; Alison Specht, University of Queensland; Margaret O'Brien, University of California Santa Barbara; Anne Cambon-Thomsen, University Paul Sabatier Toulouse III; Laurence Mabile, University Paul Sabatier Toulouse III; Yasuhiro Murayama, Integrated Science Data System Research Lab.; Pedro Luiz Pizzigatti Corrêa, University of Tennessee; Helen Glaves, British Geological Survey; Helena Cousijn, DataCite; Eric Olson, ORCID, Inc

Objectives

1) Predict the socioeconomic outcomes of natural protected areas (PAs) on rural communities using a novel combination of satellite imagery and artificial intelligence
2) Determine the influence of PAs on consumption expenditure and asset health of rural communities
3) Improve future environmental decision-making
4) Improve digitial connections between researchers, their funding, publications and data
5) Improve recommendations for the research data workflow and skills for research teams
6) Increase the number of citations to data sets and better attribute them to the data creator
7) Promote credit for open and FAIR data management and preservation for data reuse
8) Provide tools for researchers to view how the data they have deposited is used and cited.

Improve data workflow for research teams

Synthesis-science strand (David Mouillot)

WP1: Stratified sampling of 200 rural communities close to and far from natural protected areas (PAs) using matching algorithms.

WP2: Estimate socioeconomic conditions in the selected rural communities using remote sensing and artificial intelligence.

WP3: Using paired comparison tests determine whether proximity to a PA can improve socioeconomic outcomes. Identify contributing factors.

Data-science strand (Shelley Stall)

WP5: Develop leading practices, toolkits and workshops to support data sharing.

WP6: Improve capability for researchers to view how deposited data has been used, cited and reused (widget, web-accessible researcher profile).

WP4: Dissemination (website, data sharing, scientific publications, newsletters, conferences).

Research data are a vital component of the scientific record. Discovering and assessing data for possible reuse in future research is challenging. The Belmont Forum has recently awarded funds to three international teams as part of a four-year Collaborative Research Action (CRA) on Science-driven e-Infrastructure Innovation (SEI) for the Enhancement of Transnational, Interdisciplinary and Transdisciplinary Data Use to improve data management practices that will increase data reuse. One of these awardees, PARSEC, comprises two interwoven strands, one focused on improving data practices for reuse and credit, and one for synthesis science. The data specialists work alongside synthesis science researchers as they determine the influence of natural protected areas on socioeconomic outcomes for local communities. They collaborate with the researchers to better understand their motivations and work practices, and to aid them in the data-related steps that need to be taken during the research lifecycle. This will ensure their data and code are FAIR-compliant and thus enhance the likelihood of their data being reused and their analyses reproducible. The PARSEC team is working with Research Data Alliance (RDA), Earth Science Information Partners (ESIP), DataCite and ORCID to build awareness of the elements required for data creators to receive credit and automated attribution for their data contributions, and the tools that will make it easier to observe usage. Credit for data is an important incentive for researchers to make their data reusable. When data are FAIR and cited, their related publications have higher visibility. We shall discuss various ways in which we are working across the science-data interface in our multi-country and multi-disciplinary working environment to improve data (and code) reuse through better management and crediting.

Cite your Data, Get Credit, Increase Reuse and reap the rewards!

Consortium Leaders: Nicolas Mouquet, David Mouillot, Alison Specht and Shelley Stall.

Participating countries

BRAZIL: University of São Paulo–FAPESP (P. Pizzigatti Corrêa) plus postdoc and technical support (FAPESP)
FRANCE: Foundation for Research on Biodiversity, University of Toulouse III–ANR (N. Mouquet)
JAPAN: National Institute of Information & Communications Technology, Research Institute for Humanity and Nature–JST (Y. Murayama)
USA: American Geophysical Union–NSF (S. Stall)
Cooperating partners: National Computational Infrastructure, Australia (L. Wyborn), British Geological Survey, UK (H. Glaves)
Associated organizations: DataCite, ORCID, ESIP, RDA, EDI, WDS, AST, JWP, TNC

parsecproject.org