pyOpenSci: Open and reproducible research, powered by Python

Michael Trizna[‡], Leah A Wasser[§], David Nicholson[|]

‡ OCIO Data Science Lab, Smithsonian Institution, Washington, United States of America
§ Earth Lab, University of Colorado Boulder, Boulder, United States of America
| Biology Department, Emory University, Atlanta, United States of America

Corresponding author: Michael Trizna (triznam@si.edu)

Abstract

pyOpenSci (short for Python Open Science), funded by the Alfred P. Sloan Foundation, is building a diverse community that supports well documented, open source Python software that enables open reproducible science. pyOpenSci will work with the community to openly develop best practice guidelines and open standards for scientific Python software, which will be reinforced through a community-led peer review process and training. Packages that complete the peer review process become a part of the pyOpenSci ecosystem, where maintenance can be shared to ensure longevity and stability in code. pyOpenSci packages are also eligible for a “fast tracked” acceptance to JOSS (Journal of Open Source Software). In addition, we provide review for open science tools that would be of interest to TDWG members but are not within scope for JOSS, such as API (Application Programming Interface) wrappers. pyOpenSci is built on top of the successful model of rOpenSci, founded in 2011, which has fostered the development of several useful biodiversity informatics R packages. The pyOpenSci team looks to following the lessons learned by rOpenSci, to create a similarly successful community. We invite TDWG members developing open source software tools in Python to become part of the pyOpenSci community.

Keywords

APIs, software development, open source, data retrieval, data extraction, reproducibility

Presenting author

Michael Trizna

©
.

Presented at

TDWG 2021

Conflicts of interest