

# Facilitating Open science for Discovery and Scientific Progress Seismology

Lorraine J. Hwang and Louise H. Kellogg

Computational Infrastructure for Geodynamics, University of California, Davis

Seismology has a rich tradition of sharing data dating back to the earliest recorded paper records and exemplified today by organizations such as IRIS which provides access to time series data from thousands of stations worldwide. As research grows increasingly complex and data- and computationally-driven, traditional scholarly publication has expanded to recognize that data, data products, and scientific software should also be considered scholarly contributions. Recognition of research products is especially important in seismology, where data and data products such as earthquake catalogs or seismic imaging results are often reused by other researchers. In addition, research often uses specialized software developed by the scientists themselves requiring significant effort over years or decades by teams of researchers. Properly curating scientific data and software facilitates scientific discovery and promotes reproducibility, replication, discoverability, and accessibility of research, while providing credit to scientists who publish data and develop software.

As part of its mission to promote good scientific software practices in solid earth geophysics, the Computational Infrastructure for Geodynamics (CIG) maintains a community repository of research software for geodynamics. CIG's repository includes seismology codes used by the IRIS community. CIG (<https://geodynamics.org/>) originated in part from recognition of the tremendous effort required to develop and sustain high quality software developed by the scientific community for geophysics. To enable software users to give credit for software, and to provide examples of how to cite and attribute software, we developed *abc* – the *attribution builder for citation* (<https://geodynamics.org/cig/abc>). The *abc* tool uses Zenodo (<https://zenodo.org/>) as an archive and to assign a unique identifier (DOI) for scientific software and follows guidelines established by groups such as FORCE11 (see figure), an international group of scholars, librarians, publishers, and other stakeholders in the publication of data and software (<https://www.force11.org/>). *abc* also follows FAIR Principles (<https://www.go-fair.org/fair-principles>). CIG aspires to have all actively developed packages use Zenodo and Git integration to automate the archiving process and provide metadata integration with each new software release. We encourage other projects to join the Computational Infrastructure for Geodynamics Community in Zenodo (<https://zenodo.org/communities/geodynamics/>) to promote the discoverability of their data and software products.

Remaining open issues include how to handle legacy software and multi-authored libraries, and how to assign different authorship roles. While CIG does not actively curate data, data curation and model reuse face similar challenges with the additional complication of the size of model datasets produced by large and/or complex simulations.

