



Managing the explosion of high resolution topography in the geosciences

Christopher Crosby (1), Viswanath Nandigam (2), Ramon Arrowsmith (3), Minh Phan (2), and Benjamin Gross (1)

(1) UNAVCO, Boulder, CO, United States (crosby@unavco.org), (2) San Diego Computer Center, University of California San Diego, La Jolla, CA, United States, (3) School of Earth and Space Exploration, Arizona State University, AZ, United States

Centimeter to decimeter-scale 2.5 to 3D sampling of the Earth surface topography coupled with the potential for photorealistic coloring of point clouds and texture mapping of meshes enables a wide range of science applications. Not only is the configuration and state of the surface as imaged valuable, but repeat surveys enable quantification of topographic change (erosion, deposition, and displacement) caused by various geologic processes. We are in an era of ubiquitous point clouds that come from both active sources such as laser scanners and radar as well as passive scene reconstruction via structure from motion (SfM) photogrammetry.

With the decreasing costs of high-resolution topography (HRT) data collection, via methods such as SfM and UAS-based laser scanning, the number of researchers collecting these data is increasing. These “long-tail” topographic data are of modest size but great value, and challenges exist to making them widely discoverable, shared, annotated, cited, managed and archived. Presently, there are no central repositories or services to support storage and curation of these datasets.

The U.S. National Science Foundation funded OpenTopography (OT) Facility employs cyberinfrastructure including large-scale data management, high-performance computing, and service-oriented architectures, to provide efficient online access to large HRT (mostly lidar) datasets, metadata, and processing tools. With over 225 datasets and 15,000 registered users, OT is well positioned to provide curation for community collected high-resolution topographic data.

OT has developed a “Community DataSpace”, a service built on a low cost storage cloud (e.g. AWS S3) to make it easy for researchers to upload, curate, annotate and distribute their datasets. The system’s ingestion workflow will extract metadata from data uploaded; validate it; assign a digital object identifier (DOI); and create a searchable catalog entry, before publishing via the OT portal.

The OT Community DataSpace enables wider discovery and utilization of these HRT datasets via the OT portal and sources that federate the OT data catalog, promote citations, and most importantly increase the impact of investments in data to catalyzes scientific discovery.