Geophysical Research Abstracts
Vol. 16, EGU2014-11611, 2014
EGU General Assembly 2014

Geosamples.org: Shared Cyberinfrastructure for Geoscience Samples

Kerstin Lehnert (1), Lee Allison (2), David Arctur (3), Jens Klump (4,5), and Christopher Lenhardt (6)

(1) Columbia University, Lamont-Doherty Earth Observatory, Palisades, United States (lehnert@ldeo.columbia.edu), (2) Arizona Geological Survey, Tucson, Arizona 85701, United States, (3) University of Texas at Austin, Austin, TX 78758-4445, United States, (4) Helmholtz Centre Potsdam GFZ German Research Centre For Geosciences, D-14473 Potsdam, Germany, (5) CSIRO, Australian Resources Research Centre, Kensington WA 6151, Australia, (6) Renaissance Computing Institute, Chapel Hill, North Carolina 27517, United States

Many scientific domains, specifically in the geosciences, rely on physical samples as basic elements for study and experimentation. Samples are collected to analyze properties of natural materials and features that are key to our knowledge of Earth's dynamical systems and evolution, and to preserve a record of our environment over time. Huge volumes of samples have been acquired over decades or even centuries and stored in a large number and variety of institutions including museums, universities and colleges, state geological surveys, federal agencies, and industry. All of these collections represent highly valuable, often irreplaceable records of nature that need to be accessible so that they can be re-used in future research and for educational purposes.

Many sample repositories are keen to use cyberinfrastructure capabilities to enhance access to their collections on the internet and to support and streamline collection management (accessioning of new samples, labeling, handling sample requests, etc.), but encounter substantial challenges and barriers to integrate digital sample management into their daily routine. They lack the resources (staff, funding) and infrastructure (hardware, software, IT support) to develop and operate web-enabled databases, to migrate analog sample records into digital data management systems, and to transfer paper- or spreadsheet-based workflows to electronic systems. Use of commercial software is often not an option as it incurs high costs for licenses, requires IT expertise for installation and maintenance, and often does not match the needs of the smaller repositories, being designed for large museums or different types of collections (art, archeological, biological).

Geosamples.org is an alliance of sample repositories (academic, US federal and state surveys, industry) and data facilities that aims to develop a cyberinfrastructure that will dramatically advance access to physical samples for the research community, government agencies, students, educators, and the general public, while supporting, simplifying, and standardizing the work of curators in repositories, museums, and universities, and even for individual investigators who manage personal or project-based sample collections in their lab. Geosamples.org builds upon best practices and cyberinfrastructure for sample identification, registration, and documentation developed by the IGSN e.V., an international organization that governs the International Geosample Number, a persistent unique identifier for physical samples. Geosamples.org will develop a Digital Environment for Sample Curation (DESC) that will facilitate the creation, identification, and registration of 'virtual samples' and network them into an 'Internet of Samples' that will allow to discover, access, and track online physical samples, the data derived by their study, and the publications that contain these data. DESC will provide easy-to-use software tools for curators to maintain digital catalogs of their collections, to provide online access to the catalog to search for and request samples, manage sample requests and users, track collection usage and impact. Geosamples.org will also work toward joint practices for the recognition of intellectual property, build mechanisms to create sustainable business models for continuing maintenance and evolution of managing sample resources, and integrate the sample management life-cycle into professional and cultural practice of science.