Unique Sample Identification in the Solid Earth Sciences: SESAR (Solid Earth Sample Registry) and the International Geo Sample Number IGSN

K. A. Lehnert (1), S. L. Goldstein (1), W. C. Lenhardt (2), S. Vinayagamoorthy (2)

1. Lamont-Doherty Earth Observatory of Columbia University, USA
   (lehnert@ldeo.columbia.edu, steveg@ldeo.columbia.edu)
2. Center for International Earth Science Information Network, Columbia University
   (clenhardt@ciesin.columbia.edu, sri@ciesin.columbia.edu)

Data generated by the study of solid earth samples provides the basis for models and hypotheses in all disciplines of the Geosciences from tectonics to mantle dynamics to surface processes, and is fundamental to our understanding of the history and evolution of the earth's composition, climate, and biosphere. Sample-based data are diverse ranging from descriptions of lithology, texture, mineral or fossil content, stratigraphic context, and physical properties to age determinations, to chemical composition including major and trace element abundances, radiogenic and stable isotope ratios of rocks, minerals, fluid or melt inclusions. The usefulness of these data is critically dependent on their integration as a coherent data set for each sample. If different data sets for the same sample cannot be combined because the sample cannot be unambiguously recognized, valuable information is lost.

Unfortunately, such data integration has until now been seriously impeded by the lack of a consistent naming protocol that allows unambiguous identification of samples on a global scale. Different samples are often given identical names, and the same sample is often renamed as it is passed among investigators, making it difficult to link disparate data for a sample. SESAR, the Solid Earth Sample Registry [SESAR], is now addressing this problem by building an on-line registry that will for the first time provide globally unique identifiers for geological samples. The implementation of the IGSN in data publication and digital data management will facilitate the ability of investigators to build on previously collected data on samples as new measurements

are made or new techniques are developed, and will provide the solution to a basic requirement for links among and integration of data within the emerging Geoscience information network.

Each sample that is registered at SESAR will obtain a globally unique serial number, *the* **International Geo Sample Number (IGSN)**. Sample registration will require an authentication process (login) and includes submission of information about the sample such as sample type, collection location, collection time and collector/owner. Samples that can be registered with SESAR range from entire sampling events such as dredge hauls or drill cores, to individual specimens such as rocks, minerals, fossils, aerosols, and synthetic materials, to processed samples such as rock powders, mineral or fossil separates or concentrates, grain size fractions, thin sections, or leachates. Links between 'parent samples' and 'sub-samples' (child samples) will be established. Procedures for batch loading of groups of samples or large collections will be available.

SESAR is working closely with existing data management projects such as CHRONOS [CHRO], PaleoStrat [PALEO], the ODP Janus Database [ODP], the Marine Geoscience Data Management System [MGDMS], and EarthChem [EARC] to design and build a system architecture that is compatible and interoperable with these external systems so that samples can be registered with SESAR seamlessly via external data systems when users submit sample-based data to these systems. Implementation of the IGSN in sample-based data management systems will dramatically advance interoperability and data sharing, opening an extensive range of new opportunities for discovery and for interdisciplinary approaches in research.

1 References

[CHRO] CHRONOS: An Interactive Network Of Data and Tools for Earth System History, http://www.chronos.org

[EARC] EarthChem: Advanced Data Management for Solid Earth Geochemistry, http://www.earthchem.org

[MGDMS] Marine Geoscience Data Management System, http://www.marine-geo.org

[ODP] Janus Web Database, http://www-odp.tamu.edu/database/

[PALEO] PaleoStrat: A Paleontological and Stratigraphic Information System, http://www.paleostrat.com

[SESAR] SESAR: Solid Earth Sample Registry, http://www.geosamples.org