Geophysical Research Abstracts
Vol. 19, EGU2017-11645-1, 2017
EGU General Assembly 2017

Case study: Curation and publication of physical samples using persistent identifiers

Pavel Golodoniuc, Anusuriya Devaraju, and Jens Klump
CSIRO Minerals, Perth, WA, Australia

Physical samples are important observational units in Earth and Space sciences. Samples and their derived data play vital role in scientific validation and reproducibility. Systematic practices and technical solutions are required to curate and publish samples and their data on the Web. Persistent identifiers ensure unambiguous identification, and enable linkage and citation of samples and associated data sets. The International Geo Sample Number (IGSN) is a persistent and globally unique identifier for physical samples and sample collections. IGSNs are allocated to clients (e.g., laboratories, projects, individual users) through agents. Agents are institutions that represent the IGSN e.V., the IGSN Implementation Organization. For example, CSIRO, Geoscience Australia and Curtin University are IGSN allocating agents in Australia. Clients register IGSNs for samples through the respective agent's registration services.

The Discovery research program of CSIRO is conducting research in the prolific metallogenic provinces of the Capricorn Orogen – a regional study aimed to define distal footprints of covered ore systems. In CSIRO, the Capricorn Distal Footprints (CAPDF) project is one of the early adopters of the IGSN in its sample curation. The project involves collection of various samples, including water, rock, sediment, vegetation and regolith, which are collected by different researchers. The application of IGSN requires not only the supporting infrastructure and tools (e.g., user interface, service, metadata model) but also systematic workflows to cater different users, i.e. practices of laboratories or individual researchers. We present the application of the IGSN in the context of the CAPDF project as a sample inventory management system, samples identification in publications and sample discovery through a web portal. We describe workflows that demonstrate IGSN integration into existing sample curation practices and highlight challenges and benefits of such a system. The integration process takes a non-invasive approach and does not require a significant shift in existing operating procedures. We also specify how samples identified with IGSNs are linked to their datasets stored in our repository. This contributes towards enabling synoptic integration of heterogeneous geochemical datasets and their samples for regional-scale exploration.