

## The Path Toward Open Collections and Samples

\*Kerstin Lehnert<sup>1</sup>, Jens F Klump<sup>2</sup>, Lesley Wyborn<sup>3</sup>, Sarah Ramdeen

1. Columbia University, 2. CSIRO, Perth, Australia, 3. Australian National University, Canberra, Australia

While the paradigm of Open Science has substantially advanced open access to scientific publications, data, and software, with pertinent changes to policies, infrastructure, and funding structures that support open access, similar improvements to broad and consistent accessibility of physical samples and collections in the Earth and environmental sciences are lagging behind. Physical samples in the Earth, environmental, and planetary sciences are a basic element for reference, study, and experimentation, collected and studied because they represent a wider population or a larger context. Collections of samples represent highly valuable, often irreplaceable records of nature that carry enormous potential for future discoveries. We state that in the era of open, transparent, and reproducible science, preservation and persistent access to samples is no longer an option, but a mandate. Samples and collections in the geosciences need to be findable and accessible for reuse by a global research community, and to ensure transparency and reproducibility of research in the modern research ecosystem.

Essential requirements for Open Samples pertain both to their digital and to their physical management, making the path toward open samples challenging. For example, physical infrastructure (boxes, shelves, cabinets, sheds and warehouses) is needed to properly store samples. Humans are usually needed to physically retrieve specimens from their storage and prepare them for transport to the lab. Digital infrastructure is needed to publish information about samples and collections in online catalogs to make them discoverable, to ensure their unambiguous identification, and to link them to data and publications. New digital infrastructures have been emerging to make samples and collections more open and accessible. The IGSN Global Sample Number, an international persistent identifier system for samples, is already cataloguing 7 million samples that have been registered by sample repositories, museums, national geological surveys, labs, and hundreds of individual researchers. IGSNs are increasingly used to reference samples in publications and in digital datasets, allowing tools to emerge that take advantage of a new level of networked scholarly products. We will report on recent developments and initiatives that advance open and FAIR samples, and that promote and support cross-domain convergence on key issues related to the digital representation of physical samples and collections. We will also discuss the need for community agreed policies and procedures that can guide researchers, institutions, and funding agencies to plan and implement solutions for reliably and persistently curating and providing access to both the samples and the sample metadata to achieve truly Open and FAIR samples.

Keywords: Samples , Open Science, Persistent Identifiers