

# IGSN: International Collaboration Makes Samples and Sample Data FAIR

\*Kerstin Lehnert<sup>1</sup>, Jens Klump<sup>2</sup>, Lesley Wyborn<sup>3</sup>, Sarah Ramdeen<sup>4</sup>

1. Columbia University, 2. CSIRO Kensington, Australia, 3. Australian National University, Canberra, Australia, 4. Ronin Institute, USA

In recent years, the International Geo Sample Number (IGSN) has seen increasing uptake as a Persistent Identifier (PID) for physical samples because there is growing recognition that unambiguous citation and tracking of physical samples allows previously impossible linking of samples to data and publications, linking and integration of sample-based observations across data systems, and paves the road toward advanced data mining of sample-based data. The IGSN leverages the existing Handle system as its technical basis, while following the example of DataCite for its governance structure and technical architecture.

The IGSN was originally developed for the solid Earth Sciences as a centralized system, but has evolved into an international PID system that is now adopted by a growing number and range of stakeholders worldwide, including researchers, collection curators, and data managers, and by other disciplines that need to refer to physical samples. More than 6.7 million samples have been registered so far.

The recent expansion of the IGSN beyond the geosciences confirms the power of its concept and implementation, but imposes substantial pressures on the existing capacity and capabilities of the IGSN architecture and its governing organization. Significant updates to the IGSN organizational and technical architecture are necessary at this point to keep pace with the growing demand and expectations. As stated by Wittenburg et al. [2017], *“in order to be useful and reliable, PID registration and resolution systems need to be trustworthy and sustainable”*. Essential criteria for trustworthiness include an organizational foundation that ensures longevity, sustainability, proper governance, and regular quality assessment of registration services, as well as a reliable and secure technical platform based on open standards that is sufficiently scalable and flexible to accommodate the growing diversity of specimen types, use cases, and stakeholder requirements.

A major planning activity is currently underway funded by a recent award from the Sloan Foundation with participation of an international group of experts to re-design and improve the existing organization and technical architecture of the IGSN for it to be able to respond to, and support in a sustainable manner, the rapidly growing demands of a global and increasingly multi-disciplinary user community in a landscape of maturing research data infrastructures. This will include a revision of the current business model of the IGSN e.V. as it is no longer able to sustain the organization through the necessary professionalization of its operations. The goal is to ensure that the IGSN will be a trustworthy, stable, and adaptable persistent identifier system for material samples, both technically and organizationally, that attracts, facilitates, and satisfies participation within and beyond the Geosciences, that will be a reliable component of the evolving research data ecosystem, and that is recognized as a trusted partner by data infrastructure providers and the science community alike.

Keywords: Research data , Samples, Persistent Identifier, FAIR