

The Global Biodata Coalition: Towards a sustainable biodata infrastructure

Chuck Cook‡, Guy Cochrane‡

‡ Global Biodata Coalition, Strasbourg, France

Corresponding author: Chuck Cook (ccook@globalbiodata.org)

Abstract

Progress in life and biomedical sciences depends absolutely on biodata resources—databases comprising biological data and services around those databases. Supporting scientists in data operations and spanning management, analysis and publication of newly generated data and access to pre-existing reference data, these biodata resources together comprise a critical infrastructure for life science and biomedical research. Familiar scientific infrastructures—for example the Conseil Européen pour la Recherche Nucléaire ([CERN](#)) or the [Square Kilometer Array](#), are distinct, constructed, physical entities that are centrally funded and managed at one or more identifiable locations. By contrast, the primary infrastructure of the life sciences—comprised of databases and other biological data resources—is globally distributed, virtually connected, funded from multiple sources, and is not managed as a coordinated entity. While this configuration supports innovation, it lends itself poorly to the long-term sustainability of individual biodata resources and of the infrastructure as a whole. The [Global Biodata Coalition \(GBC\)](#) brings together life science research funding organisations that recognise these challenges and acknowledge the threat that the lack of sustainability poses. They agree to work together to find ways to improve sustainability.

In the presentation, we will provide an overview of the global biodata resource infrastructure, focusing in particular on challenges to providing sustained long-term funding to the resources that comprise the infrastructure. This will provide a global context to other presentations in the session, which focus on biodata resources in Australia.

Covering some of the work that GBC has carried out to understand and classify biodata resources and the entire biodata resource infrastructure, we will outline the Global Core Biodata Resource programme and Inventory project and also introduce the stakeholder consultation processes around approaches to sustainability and open data. Finally, we will lay out the path GBC is taking to engage researchers, informaticians, funding organisations and other stakeholders in moving towards greater sustainability for these critical resources

Keywords

data resources, biodiversity, open data, databases, life sciences, funding

Presenting author

Guy Cochrane

Presented at

TDWG 2023

Conflicts of interest

The authors have declared that no competing interests exist.