Contribution ID: **74**   Type: **Presentation**

Accelerating Research with Data Commons and Data Meshes

Monday 13 October 2025 15:36 (11 minutes)

About Data Commons and Data Meshes

A data commons is a cloud-based software platform with a governance framework that enables a research community to manage, analyze and share its data. A data mesh is a collection of two or more data commons, cloud-based computational resources, and other cloud-based resources that interoperate using a common set of core software services and a hybrid governance model. Data commons and data meshes are two important types of data platforms supporting the research community.

Reason for a Session

Today, there are over 50 data commons and similar platforms around the world and a growing number of data meshes. The primary goal of this session is to bring together IDW/RDA participants interested in data commons and data meshes and in sharing lessons learned about building, operating and using data commons and data meshes. Data commons and data meshes are emerging as important platforms for building, distributing and federating AI models. A secondary goal of the session is to bring together IDW participants interested in building AI models over data commons and data meshes.

Relevance to the conference

Data commons support open reproducible research and provide FAIR access to the data they manage, both of which are important themes for the conference. Data meshes provide a mechanism for interoperating and federated data commons and supporting interdisciplinary research, which are also important themes for the conference. The session will include talks on CAREful Indigenous Data Governance and all of the talks will provide regional and international perspectives.

Structure of Session

Four 15 minute talks about data commons and data meshes (see below). One 30 minute panel discussion with audience participation about using data commons to support reproducible scientific research and discovery.

Four 15 Minute Talks

Robert L. Grossman - From Data Commons to Data Meshes

The talk will provide a brief introduction and overview of data commons and data meshes and how they are beginning to support AI. He will discuss some of the lessons learned over the past decade that contribute to a successful data commons and data mesh. Finally, he will provide an introduction to some of the approaches that are emerging for integrating AI models and frameworks with data commons and data meshes.

About Robert L. Grossman. Robert Grossman is a Professor of Medicine and Computer Science at the University of Chicago and the Director of the Center for Translational Data Science at the University of Chicago. He is also the Director of the Open Commons Consortium. He is the lead for the open Gen3 data platform which has been used to build over 20 data commons and data meshes.

Claire Rye - Building a genomic data repository for taonga species in Aotearoa New Zealand

The Aotearoa Genomic Data Repository (AGDR) is an Aotearoa-based resource that enables researchers and Māori communities to fulfil their obligations relating to the guardianship, management, sharing and use of genomic data from biological samples that are taonga (precious or treasured). The AGDR was jointly developed by Genomics Aotearoa and the New Zealand eScience Infrastructure, with funding from the Ministry

of Business Innovation and Employment. Its design is based on the principles of Māori data sovereignty, enabling kaitiaki (Māori guardianship) centric control over who can access genomic data, and for what purposes. AGDR has been developed based on gen3 data commons and Globus, and is in line with the globally-relevant CARE and FAIR Principles, ensuring data is findable, interoperable, and re-usable in cases where there is prior and informed consent, and access is provided by kaitiaki.

About Claire Rye. Claire Rye is a Product Manager at New Zealand eScience Infrastructure (NeSI) based out of the University of Auckland. She works across the Aotearoa Genomics Data Repository and Rakeiora Pathfinder projects and generally looks at research data management and the data lifecycle across NeSI and nationally as part of the Research Data Culture Conversation.

Steven Manos - Some Best Practices for Building Data Commons: The Australian BioCommons Experience

The talk will provide an overview of how the Australian BioCommons develops and operates data commons and discuss some of the the related services supporting research for data managed by data commons.

About Steven Manos. Steven Manos is the Associate Director of Cyberinfrastructure for the Australian Bio-Commons. His interests lie in building digital platforms designed specifically for researchers, which includes developing partnerships and community building.

Bernie Pope - establishing data commons for human omics data in Australia

Australian BioCommons is working with several partner organisations in Australia to build human omics data commons tailored to the particular needs of the respective research communities. This work has received national funding from the National Collaborative Research Infrastructure Strategy (NCRIS), supporting the GUARDIANS program, and from the Medical Research Future Fund (MRFF), supporting the Australian Cardiovascular Disease Data Commons (ACDC) and the OMIX3 program. In this presentation we will describe the steps taken to establish these data commons, including critical aspects of data management and governance.

About Bernie Pope. Bernie is a Professor in the Faculty of Medicine, Dentistry and Health Sciences at The University of Melbourne. He works in the fields of computer science and bioinformatics, where he applies scalable and robust computational methodologies to key challenges in human health, particularly in genomics and cancer. He is Associate Director at Australian BioCommons and leads the Human Genome Informatics division.

Primary author: GROSSMAN, Robert (University of Chicago)

Co-authors: POPE, Bernie (University of Melbourne); RYE, Claire (New Zealand eScience Infrastructure); MANOS, Steven (Australian BioCommons)

Presenters: POPE, Bernie (University of Melbourne); RYE, Claire (New Zealand eScience Infrastructure); GROSSMAN, Robert (University of Chicago); MANOS, Steven (Australian BioCommons)

Session Classification: Presentations Session 2: Data and Research & Data Science and Data Analysis

Track Classification: SciDataCon Persistent Themes: Data and Research