Contribution ID: **134**   Type: **Presentation**

Promoting the Use of Discipline-Specific Metadata for Data FAIRness

Tuesday 14 October 2025 12:47 (11 minutes)

The FAIR guiding principles indicate that scientific datasets should be annotated with "rich" metadata that adhere to relevant community standards. Those standards include metadata reporting guidelines, which enumerate the attributes needed to describe the features of the experiments that led to the corresponding data, and the controlled terms that standardize the values of those metadata attributes. Some scientific communities have been developing their own metadata standards for nearly three decades, while others are only now beginning. The functional genomics community, for example, released the Minimal Information About a Microarray Experiment (MIAME) back in 2001—offering a reporting guideline that has been widely adopted by researchers, by publishers, and by the discipline-specific repositories where microarray data are archived, enhancing data discoverability, interpretability, and reusability. Yet the majority of scientific communities still lack metadata standards, seriously impeding the FAIRness of their datasets.

This panel session will highlight the importance of developing discipline-specific metadata reporting guidelines to support the FAIRness of data both within and across domains.

The panel will involve brief presentations by a group of international experts, with considerable time for discussion among the panelists and with the audience. The panel is being organized by Christine Kirkpatrick, Founder and Head of GO FAIR US and Secretary General of CODATA, and by Mark Musen, Stanford Medicine Professor of Biomedical Informatics Research at Stanford University (USA). The sequence of presentations will be as follows:

Christine Kirkpatrick will introduce the problem and the panel. She will survey the use of discipline-specific metadata standards in science and discuss efforts on the part of GO FAIR US to educate research communities in the development and use of rich metadata standards that are tailored for different classes of experiments within their particular disciplines.

Mingfang Wu, Product Manager at the Australian Research Data Commons, has performed studies of researchers in clinical investigation, social science, and ecology, and she has identified how metadata granularity directly affects their ability to perform dataset search and reuse. She will offer evidence for how discipline-specific metadata can enhance data FAIRness.

Wonsik "Jeff" Shim, Professor of Library and Information Science at Sungkyunkwan University in Seoul, KR, will discuss the use of rich, discipline-specific metadata within the Korean Research Memory—a comprehensive repository of research output from all publicly funded projects supported by the National Research Foundation of Korea. Prof. Shim will describe the repository's rich, hierarchical organization for metadata, particularly in the humanities and social sciences, and how such discipline-specific metadata facilitate data search, data understanding, and data reuse.

Erik Schultes, FAIR Implementation Lead for the GO FAIR Foundation, NL, will discuss the Metadata for Machines workshops that his organization has pioneered over the past eight years. He will present the structure of these 2–3 day events that lead subject matter experts through the process of constructing relevant discipline-specific reporting guidelines and sets of controlled terms to offer new community standards for detailed metadata authoring. He will offer an assessment of the strengths and limitations of intensive workshop environments to help scientists to create and document relevant and meaningful metadata standards and to deploy those standards more routinely.

Mark Musen will provide perspective on the preceding presentations. He then will describe the use of the CEDAR system for managing and disseminating discipline-specific metadata standards. He will discuss how CEDAR encodes a research community's preferences for discipline-specific standards as metadata templates,

and how those templates can become shared representations for use by a variety of data-management technologies—for metadata authoring, for metadata correction, and for metadata harmonization.

Overall, the panel will overview the status of discipline-specific metadata, discussing the feasibility and utility of incorporating such annotation in real-world repositories. The panel will discuss strategies for creating and disseminating new discipline-specific, community-based metadata standards, and for encoding those standards in a way that makes them maximally reusable across a variety of applications that rely on those standards.

The FAIR guiding principles mandate the use of "rich,"community-relevant metadata standards, but they do so without offering guidance for how research groups might actually accommodate such standards. We now have considerable experience within the data-management community that offers practical solutions for the creation, promulgation, and use of discipline-specific metadata—demonstrating that the creation of truly FAIR data is not only desirable, but also manifestly achievable.

Primary authors: KIRKPATRICK, Christine (San Diego Supercomputer Center / CODATA); Prof. MUSEN, Mark (Stanford University)

Co-authors: SCHULTES, Erik (GO FAIR Foundation); WU, Minfang (Austrailian Research Data Commons); SHIM, Wonsik (Sungkyunkwan University)

Presenters: KIRKPATRICK, Christine (San Diego Supercomputer Center / CODATA); SCHULTES, Erik (GO FAIR Foundation); Prof. MUSEN, Mark (Stanford University); WU, Minfang (Austrailian Research Data Commons); SHIM, Wonsik (Sungkyunkwan University)

Session Classification: Presentations Session 3: Rigorous, responsible and reproducible science in the era of FAIR data and AI

Track Classification: SciDataCon Persistent Themes: Open Data, FAIR Data, Innovation, Industry and Development