Contribution ID: 19   Type: **Session**

Units, Symbols and Terminology for Data Driven Science: Philosophy to Pragmatism

Tuesday 14 October 2025 11:30 (1h 30m)

The rapid changes in the way we undertake scientific research driven by digitalisation and artificial intelligence means we need to look again at the basis of scientific methods, requiring the input from the philosophy of science and one end while being pragmatic about the uses of technology at the other. This approach will provide a way to ensure that our research and teaching is relevant, effective and efficient.

More specifically, this means revisiting the topics of quality and provenance frameworks as developed by certain disciplines; exploring the role of metadata and semantics in such frameworks and in the context of use of data by AI systems; exploring how the digital representation of units is best managed in such frameworks.

Consequently, the session will address the following questions:

1. How should we best interact with AI systems, how can they contribute to scientific discovery?

2. How can we ensure that sensible provenance is provided with data that is consumed and produced by AI systems? Why is it so hard to get quality metadata?

3. How do we ensure that the metadata and semantic framework is useful and used by researchers and not seen as yet another barrier to work?

4. The role of metadata standards and ontologies. How do we evolve the publication and dissemination framework to provide these details?

5. What software tools are needed to support researchers to use digital terminology and units? What support can be built into programming languages?

This session builds on the agenda set by a successful small conference held at the Royal Society of Chemistry in London in March 2025, and convened by the UK Physical Science Data Infrastructure (PSDI, www.psdi.ac.uk), the International Union for Pure and Applied Chemistry (IUPAC, Green and Gold Book projects) and CODATA (DRUM Task Group).

Proposed speakers:
We appreciate that this list would be too many for a 90 min session but we wanted to show that we had a wider group of speakers to call upon to support the session as we will need to raise funding for some of them to be able to attend. We will seek funding to bring several speakers to Brisbane and take advantage of the hybrid format to enable those who can not travel to participate online or via recorded talks if the time zones are an issue.

1. Philosophy of science in the AI age –Vanessa Seifert (Greece), Will McNeil (UK)

2. IUPAC –Digital Transformation of IUPAC –Stuart Chalk. (USA) , Jeremy Frey (UK)

3. Digital SI –Max Gruber (Germany) Pragmatic Semantics –Samantha Pearman-Kanza (UK) 4.Usable Metadata –Cerys Willoughby (UK)

4. Definition of units, symbols and terminology in PSDI - Aileen Day (UK) (presenting/participating remotely)

5. M-Layer –Blair Hall, (New Zealand)

6. CODATA DRUM Task Group members

.

Primary authors: Dr DAY, Aileen (University of Southampton); Dr HALL, Blair; FREY, Jeremy (University of Southampton); Dr GRUBER, Max (PTB); Dr PEARMAN-KANZA, Samantha (University of Southampton); Prof. CHALK, Stuart (UNF); Dr SEIFERT, Vanessa (University of Athens)

Presenters: Dr DAY, Aileen (University of Southampton); Dr HALL, Blair; Dr GRUBER, Max (PTB); Dr PEARMAN-KANZA, Samantha (University of Southampton); Prof. CHALK, Stuart (UNF); Dr SEIFERT, Vanessa (University of Athens)

Track Classification: SciDataCon2025 Specific Themes: Rigorous, responsible and reproducible science in the era of FAIR data and AI