Contribution ID: **39**   Type: **Short Talks**

PROVENANCE IN FUSION: APPLYING THE W3C-PROV MODEL TO FUSION DATA

Friday, 3 December 2021 15:15 (10 minutes)

With a growing emphasis on open data, and a drive to improve the interoperability of fusion data, a holistic and standardised approach to data provenance is required. Modern and emerging research methods that handle large data sets will require that metadata, including provenance, can be handled similarly and programmatically, regardless of their source. The FAIR data principles outline guidance for improving data quality and machine-actionability, featuring provenance as a key measure of how reusable data are. Well described provenance also increases the level of trust that users can have in the data, encouraging the wider use of it in studies. The W3C-PROV model provides a way of describing provenance that is both human and machine-readable. As part of the work carried out by the FAIR4fusion project, an approach to the curation of provenance for fusion data is presented. The challenges and areas for development are discussed, with examples including data from the Mega-Amp Spherical Tokamak (MAST) in the UK, and WEST in France. WEST data is provided as Interface Data Structures (IDS) –the format developed for ITER data –demonstrating a potential for fusion data to be united and distributed under a common ontology, with provenance described according a globally-recognised standard.

Country or International Organisation

United Kingdom

Affiliation

UKAEA

Primary authors: CUMMINGS, Nathan (UKAEA); IMBEAUX, Frédéric (CEA-IRFM); MORALES, Jorge (CEA); Mr ARTAUD, Jean-François (CEA, IRFM)

Presenter: CUMMINGS, Nathan (UKAEA)

Session Classification: Friday 3 Dec

Track Classification: Big data