



Contribution ID: 599 Contribution code: TUPS70

Type: **Poster Presentation**

## The Data Platform: an independent system for management of heterogeneous, time-series data to enable data science applications

*Tuesday, 21 May 2024 16:00 (2 hours)*

The Data Platform is a fully independent system for management and retrieval of heterogeneous, time-series data required for machine learning and general data science applications deployed at large particle accelerator facilities. It is an independent subsystem within the larger Machine Learning Data Platform (MLDP) which provides full-stack support for such facilities and applications [1]. The Data Platform maintains the heterogeneous data archive along with all associated metadata and post-acquisition user annotations. It also facilitates all interactions between data scientists and the data archive, thus it directly supports all back-end data science use cases. Accelerator facilities include thousands of data sources sampled at high frequencies, so ingestion performance is a key requirement and the current challenge. We describe the operation, architecture, performance, and development status of the Data Platform.

### Footnotes

[1] C.K. Allen, C. McChesney, et. al., "The Machine Learning Data Platform: Full Stack Support for Data Science Based Modeling, Control, and Optimization in Particle Accelerator and Large Experimental Physics Systems", IPAC24.

### Funding Agency

Work performed under the auspices of the U.S. Department of Energy with funding by the Office of High Energy Physics SBIR Grant DE-SC0022583.

### Paper preparation format

Word

### Region represented

North America

**Primary author:** MCCHESENEY, Craig (Osprey DCS LLC)

**Co-authors:** ALLEN, Christopher (Osprey DCS LLC); DALESIO, Leo (EPIC Consulting); DAVIDSAVER, Michael (Brookhaven National Laboratory)

**Presenter:** MCCHESENEY, Craig (Osprey DCS LLC)

**Session Classification:** Tuesday Poster Session

**Track Classification:** MC6: Beam Instrumentation, Controls, Feedback, and Operational Aspects:  
MC6.D13 Machine Learning