

EGU21-13177

<https://doi.org/10.5194/egusphere-egu21-13177>

EGU General Assembly 2021

© Author(s) 2021. This work is distributed under the Creative Commons Attribution 4.0 License.



Data Stewardship Practices for Earth Observation Transient and Optimized Analysis Platforms

Kaylin Bugbee¹, Rahul Ramachandran¹, Ge Peng², and Aaron Kaulfus¹

¹NASA MSFC, Huntsville, United States of America (kaylin.m.bugbee@nasa.gov)

²University of Alabama in Huntsville, Huntsville, United States of America

Access to valuable scientific research data is becoming increasingly more open, attracting a growing user community of scientists, decision makers and innovators. While these data are more openly available, accessibility continues to remain an issue due to the large volumes of complex, heterogeneous data that are available for analysis. This emerging accessibility issue is driving the development of specialized software stacks to instantiate new analysis platforms that enable users to quickly and efficiently work with large volumes of data. These platforms, typically found on the cloud or in a high performance computing environment, are optimized for large-scale data analysis. These platforms can be transient in nature, with a defined life span and a focus on improved capabilities as opposed to serving as an archive of record.

While these transient, optimized platforms are not held to the same stewardship standards as a traditional archive, data must still be managed in a standardized and uniform manner throughout the platform. Valuable scientific research is conducted in these platforms, making these platforms subject to open science principles such as reproducibility and accessibility. In this presentation, we examine the differences between various data stewardship models and describe where transient optimized platforms fit within those models. We then describe in more detail a data and information governance framework for Earth Observation transient optimized analysis platforms. We will end our presentation by sharing our experiences of developing such a framework for the Multi-Mission Algorithm and Analysis Platform (MAAP).