

A Standard Reference Model for Data Archives

John S. Hughes¹, Daniel J. Crichton², David Giaretta³, Mike Kearney⁴, Robert R. Downs⁵, Matthias Hemmje⁶, Rosemarie Leone⁷, Terry Longstreth⁸, John Garrett⁹, Sean Hardman², Ronald Joyner², Michael McAuley², Costin Radulescu², Bruce Ambacher¹⁰

¹Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA,
John.S.Hughes@jpl.nasa.gov

²Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA

³PTAB Ltd, Dorset, UK

⁴Sponsored by Google, Huntsville, AL 35803, USA

⁵Center for International Earth Science Information Network (CIESIN), Columbia University, Palisades,
NY, USA

⁶FernUniversität in Hagen, Hagen, Germany

⁷European Space Agency (ESA), Frascati, Rome, Italy

⁸Data and Information Standards Consultant, Laurel, MD, USA

⁹PTAB Consultant, Columbia, MD, USA

¹⁰Research Associate, Digital Curation Innovation Center, College of Information Studies, University of
Maryland, College Park, MD

Abstract

An implementable Data Archive Architecture is being developed for trusted digital repositories based on the Reference Model for an Open Archival Information System (OAIS) – ISO 14721. A set of interoperable protocols and interface specifications are planned that will offer capabilities for accessing, merging, and re-using data, both within and across the operational boundaries of trustworthy digital repositories. The model will also provide support for the fundamental scientific need to verify the reproducibility of results.

This standards development task is being performed by the Data Archive Interoperability (DAI) working group within the Consultative Committee for Space Data Systems (CCSDS). The architecture integrates concepts from the OAIS Reference Model, the ISO/IEC 11179 Metadata Registry (MDR) standard, the CCSDS Reference Architecture for Space Information Management (RASIM), the proposed draft recommended practice document, Information Preparation to Enable Long Term Use (IPELTU), and three decades of digital repository development for science research.

Keywords

Interoperability, Information Architecture, Information Model, Digital Repository

1.0 Introduction

Long-term digital preservation methods are critical practices to preserve the benefits of space missions for the mission customers and our society as a whole. A new concept is presented, a Data Archive Architecture called the OAIS Interoperability Framework (OAIS-IF) [2]; a software architecture suite of standards that will support the OAIS and add capabilities for interoperability between users and all archives that comply with the OAIS-IF standards. This architecture is derived from and supports the Open Archival Information System (OAIS) Reference Model (RM) [1] approach developed by the Consultative Committee for Space Data Systems (CCSDS) and adopted by archives world-wide.

Establishing the OAIS-IF provides stakeholders with a conceptual understanding of the components and the relationships among those components needed to implement systems that can foster acquisition, stewardship, and continuing access to data products and related information resources and services that have been selected for use by designated communities. Identifying and describing the components and relationships described within the OAIS-IF enables stakeholders, including data producers, archivists, repository managers, and sponsors to plan, design, procure, implement, and operate systems that support the long-term use of data that are currently being collected or stored. Furthermore, understanding the components and the relationships depicted within the OAIS-IF enables repository system developers to create and acquire the services and features needed to provide the functionality that data stewards can utilize to provide data producers and users with the capabilities needed for sharing and using data, respectively.

2.0 The OAIS Interoperability Framework

The first OAIS-IF product that will be produced by the DAI WG is the Data Archive Architecture Description Document (DA-ADD). The ADD will describe the overall framework from top to bottom. It is a prerequisite to insure a compatible end-to-end design of the other components in the OAIS-IF

architecture. A conceptual framework [2] is illustrated in Figure 1. The upper framework provides the user interfaces for the producers and consumers of the OAIS archive data.

The middle framework is an Archive Abstraction Layer (AAL) with the function of obscuring and encompassing the implementation details to facilitate interoperability and archive platform independence. Example interactions that pass through the AAL would include methods to allow producers and consumers to interact with an archive using OAIS standard informational entities including the Submission Information Package (SIP); Archival Information Packages (AIPs), Dissemination Information Package (DIP) and their components.

Finally, below the framework are the specific plug-in components that connect the framework to a domain archive. These components map the standard OAIS information entities and bind OAIS-IP service interfaces to their counterparts in the domain specific archives. These plug-in components also expose domain specific information models and services.

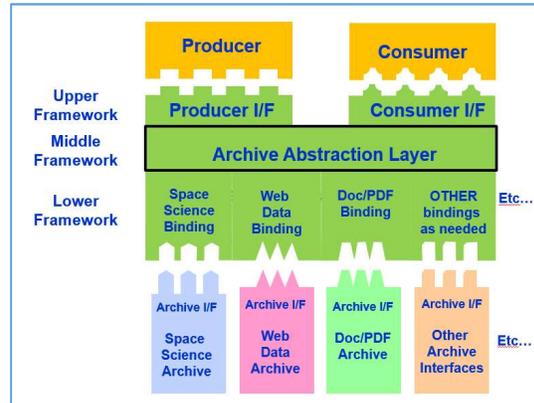


Figure 1 - The overall high-level structure of OAIS-IF

3.0 An Example Interoperable Infrastructure – The Planetary Data System

In 1982 and 1986, the National Research Council (NRC) Committee on Data Management and Computing (CODMAC) issued reports that set guidelines for the development of science data archives. Based on the CODMAC guidelines, the Planetary Data System (PDS) was established in 1989 and since then has aggressively promoted long-term preservation in the Planetary Science community. In Oct 2013 the PDS released PDS4 [3, 4], a complete redesign of its system using lessons learned and principles set forth in the Open Archival Information System (OAIS) and ISO/IEC 11179 [8] reference models. Soon afterwards, the International Planetary Data Alliance (IPDA), a consortium of international space agencies, endorsed the PDS standards and set the course for bringing together the entire science discipline under one interoperable infrastructure.

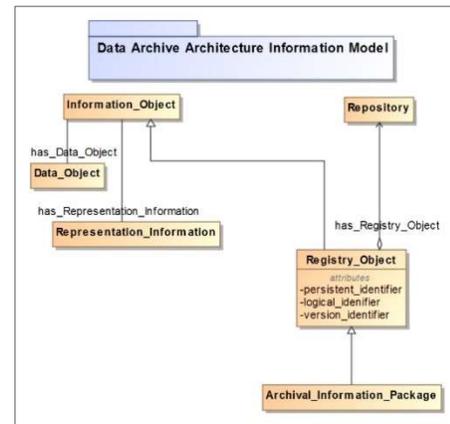


Figure 2- Information Components

4.0 Status of the Architecture Description Document

A draft of the Data Archive Architecture has been created using concepts from sources that include the Reference Architecture for Space Information Management (RASIM) [5], PDS4 specifications [3, 4], the Data Archive Ingest (DAI) WG Report to the CCSDS Management Council [6], SCience Data Infrastructure for Preservation – Earth Science (SCIDIP-ES) [7], and reports and recommendations from the Research Data Alliance (RDA) Data Fabric WG. The draft

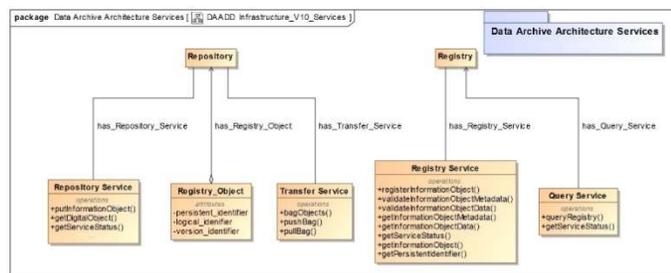


Figure 3 - Service Components

model is being captured and managed in the Cornerstone Framework (NPO-49832), the framework used to in the development of the PDS4 Information Model. A few of the information model and service components are illustrated in Figures 2 and 3 respectively.

5.0 Conclusion

In 2002, the National Virtual Observatory Science Definition Team said, “It is probably safe to say that no other professional community has reached the level of data interchange standards (both syntax and semantics) that we have reached in astronomy.” [9] The international planetary science community has met this challenge with the development and adoption of the PDS4 Information System. The OAIS-IF is introduced as a conceptual model for improving understanding of the components and relationships needed for planning, designing, developing, acquiring, and operating digital repositories that enable the submission, management, preservation, and continuing use of data products and related resources that have been identified as having enduring value for designated communities. The universal acceptance of the OAIS RM provides the opportunity to extend these successes by developing an OAIS Interoperability Framework that will enable interoperability across data archives in general.

6.0 Acknowledgements

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (NASA). © 2018. All rights reserved. Support for Robert Downs was received from NASA under contract NNG13HQ04C for the Socioeconomic Data and Applications Distributed Active Archive Center (DAAC).

References

- [1] ISO 14721:2003: Reference Model for an Open Archival Information System (OAIS), ISO, 2003.
- [2] Kearney, M., et. al., Digital Preservation Archives – A New Future Architecture for Long-term Interoperability, To be published, 15th International Conference on Space Operations, Marseille France, May 2018.
- [3] Planetary Data System - PDS4 Information Model Specification, Version 1.8.0.0, March 2017.
- [4] Planetary Data System - PDS4 System Architecture Specification September 1, 2013, Version 1.3.
- [5] Reference Architecture for Space Information Management (RASIM) CCSDS 311.0-M-1
- [6] Data Archive Ingest (DAI) WG Report to the CCSDS Management Council (CMC), Figure 2: Notional Data Archive Architecture, March 2017
- [7] Science Data Infrastructure for Preservation-Earth Science Data. (2012). Retrieved April 28, 2018, from www.scidip-es.eu.
- [8] ISO/IEC 11179: Information Technology -- Metadata registries (MDR), ISO/IEC, 2008.
- [9] NVO Science Definition Team, “Towards the National Virtual Observatory”, A Report Prepared by the National Virtual Observatory Science Definition Team, 2002.