Developing an OAIS based Interoperability Framework

David Giaretta, … ???

ABSTRACT

The Open Archival Information Reference Model (OAIS) has been very successful in guiding preservation of digitally encoded information. From OAIS the ISO 16363 standard, and the certification process for Trustworthy Digital Repositories, were developed.

OAIS (2012) included a roadmap for a number of follows on standards including interfaces between OAIS based type Archives, submission of information to an archive, and delivery of information from an archive, and metadata information retrieval about digital and physical data sources.

The OAIS Interoperability Framework (OAIS-IF) aims to provide the set of standards to fulfil this roadmap.

The work began with a reminder by Vint Cerf that there was a need for a communications protocol which would allow an OAIS Archive to send AIPs to another OAIS Archive, which is one of the fundamental preservation options identified by OAIS. While considering this, the CCSDS DAI Working Group realized that, when bringing in the need to look at communications protocols, one could leverage the OAIS concepts to improve cross-disciplinary usability of digitally encoded information by building on the ideas of usability, the authenticity, the context and Provenance of such information.

An OAIS Archive guarantees usability for a Designated Community. Clearly most other current and future users are not in that Designated Community, and therefore are not guaranteed to be able to understand that data. What OAIS identifies is that a user needs Representation Information in order to understand and use the digitally encoded information. Therefore one needs ways of incrementing the amount of valid and valuable Representation Information that may be made available to a user – beyond that required for the Designated Community. By supporting the gathering of the Data Objects as well as the appropriate Representation Information from many disparate sources, there will be an increase in the ability to combine those pieces of digitally encoded information to produce new insights through interdisciplinary research.

Users may also wish to understand the information, such as Provenance – how and where the information was collected - and other “metadata”, associated with the digital objects. All of these are most likely to be digitally encoded in the modern World. OAIS tells us that we can apply the same concepts here as for the “primary” information such as scientific data. Provenance, gathered in one way can then be combined and compared with Provenance gathered in a completely different way, in order to be sure that one is combining primary data which is compatible e.g. gathered at the same time and in a consistent way.

These realizations are allowing us to develop standards which are widely and consistently usable across all types of information.

The CCSDS formal or official process for implementable standards requires prototyping to show that two or more sets of independent developers can use the standards to produce implementations which can interoperate. In this process the DAI Working Group has been impressed by the way in which modern software tools allow one to use the standards to create prototypes which are remarkably scalable, robust and cloud-ready. Those implementations are technology dependent of course but, as a fundamental requirement, the standards are technology agnostic.

This paper will describe the way in which the OAIS-IF related standards, which should be extremely widely applicable, are being developed. We have taken into account the need to simplify what changes are needed by archives and users. In addition, we can extend the applicability beyond OAIS conformant Archives, so they can be used by any source of information, for any purpose.

If widely adopted these standards will increase the cross-disciplinary usability of all kinds of digitally encoded information, making it easier to combine and interoperate with them, thereby increasing their value.