DAI WG Report to the CMC

Response to CMC resolution 2016-10-03

# The CMC Resolution:

CMC acknowledges that in today’s  Data Economy, space systems  generating and transporting large amounts of relevant data;  be it science (GAIA is generating 50 GB/day) earth observation (Sentinel-2A generating 1 TB/ h), navigation or telecommunication satellites,  constitute key assets for the Data industry.

In this context, standards supporting proper storage and preservation of data, open access to the space data, merging of data from different sources, analysis and processing big data and all related processes and mechanisms are of paramount importance.

Scientific communities are already for quite some time working nationally and internationally to define adequate standards for their data, be it astronomy, planetary exploration or earth observation.

The question is what should the CCSDS organisation do and how should it position itself in the wake of such space data revolution and  how could the work done over the years to standardise data storage and archiving be leveraged at the service of the new paradigms.

In this context, and following the request of the DAI WG to the CMC for support, CMC would like to encourage the DAI WG to address the following questions within a short paper to be delivered to CMC *by 1 February 2017*:

1.        What is the current DAI roadmap and how does it make a valuable contribution to the data issue, taking into account what has already been produced including the overall worldwide efforts in this area?

2.        Taking into account that there are dedicated international organisations working on archiving and merging scientific data, astronomy data, as well as Earth data, what should be the role of CCSDS in general and that of the DAI WG in particular?

3.        Is the long term data preservation of the housekeeping data properly addressed in these initiatives? Is there a need for developing this part of the data preservation?

4.        In this context, what would be the next project for the DAI WG in the future, should resources be available, and how would this project fit with the current CCSDS reference architecture and work plan?

# Introduction

The Data Archive Ingest (DAI) WG thanks the CMC for the opportunity to respond to their insightful context and questions that were posed in their resolution (a copy of which may be found at the beginning of this document). We have been hoping for an opportunity to bring these important topics to the CMC for some time.

We hope this brief introduction will provide some useful context for the broad scope of the DAI WG. Additional details are provided in the answers to the CMC questions, although, because we have not started the work, these are necessarily still at a high level and may change as we consult space agencies and others.

Space Agencies produce data, either directly, for example Earth Observation data, or indirectly, for example housekeeping data to know how the condition of the spacecraft affects its working. Information from space increases in value when merged with data from other sources and other times, for example longitudinal studies over time, merging space data with grounf truth data, for Climate Change studies.

Merging, analyzing and processing data from multiple sources requires more than simply transferring bits to one place. The challenge is to be able to use data with which one is not familiar, for example if one does not have the appropriate software, while making sure one combines data in the correct way, for example ensuring one does not add distances in feet to distances in meters.

DAI works on Information Management Services for CCSDS, although it was realized that the DAI concepts and outputs are applicable more widely. The DAI concepts include “Long-term”, implying access even after the mission systems have been terminated and the data in the archive must be “self describing” to different communities of users and new systems seeking to use it. Long Term is long enough to be concerned with the impacts of changing technologies, including support for new media and data formats, or with a changing user community.

DAI has not specifically addressed “near-term” telemetry archive access (for example) by MOC systems while the mission is still underway, although our processes and approaches are important practices for both near-term and long-term digital archives. Specifically, it is important to note that long-term preservation must be taken into account at the very beginning of the project because it will have an impact on the total cost of ownership, the selection of the data formats, the sizing of the archiving infrastructures and requires functions to be performed during the mission, capturing data descriptions, etc. So the DAI WG products, while focused on long-term preservation, apply to the entire life cycle of a space mission, even the early conceptual phases (Phase A). Few agency missions consider the long-term preservation implications while developing their systems. Those that do have discovered advantages both in resources and ease of use for the associated information. The DAI WG has produced both general best practice standards as well as technical standards that would be useful to agency missions wishing to gain these advantages. We want to promote awareness of these standards among the CCSDS Space Agencies. We also want to continue development of more detailed technical interoperability standards to continue to move the preservation field forward.

To respond to the CMC, first we will provide some explanations that respond to the context in the resolution’s introductory material. Then we will explicitly answer the questions one-by-one.

We interpret the introductory context material in the CMC resolution to ask the question about what CCSDS and DAI should do to position the CCSDS agencies to function in the “big data” world.

It is well known that the CCSDS DAI WG formulated OAIS and related standards for the purposes of digital preservation of the space agencies’ valuable mission data, but the solutions and practices were so valuable that OAIS has been adopted by archival repositories around the world.

OAIS defines preservation in terms of ensuring the data is understandable and usable by user communities, which is achieved by ensuring the data is adequately described by metadata specific to that data. OAIS refers to such metadata as “Representation Information”. DAI provided a number of reasonably generic standards for creating Representation Information but most will be domain specific standards created by others. Adequate Representation Information for multiple types of data enables the datasets to be merged and used interoperably.

The DAI WG has a roadmap (to be described in response to question 1) and plans to refocus on interoperable protocols for digital archive access. This includes a general framework on the user side, an abstraction layer (similar to that developed by the SM&C WG), and archive interfaces that are specific to each discipline or “designated community”. A very successful effort would result in the user interface and the abstraction layer being another general solution that works well for the global problem of digital preservation, resulting in adoption by the same global community that adopted OAIS. But the discipline-specific archive interfaces developed by DAI would be dedicated to the space agencies’ needs for space science, spacecraft data (housekeeping) and perhaps other data types unique to space programs (such as Enterprise Data for developing spacecraft).

This architecture would have the unique feature of enabling cross-discipline research in trusted digital archives. The Space Science researcher would use the consumer interface and abstraction layer and the space science specific “plug-in” to access space science results and spacecraft telemetry. But (s)he would also be able to install a plug-in for geological data archives or weather data archives by installing those “plug-ins”, developed by those communities. As noted above adequate Representation Information enables interoperability and merging data from different sources. Meanwhile the geology or weather researcher would similarly be able to plug in the space science archive plug-in to correlate their discipline data with spacecraft mission results.

The importance of this cannot be overstated. In the future, the greatest technological advances in the world are expected to come from cross-discipline research.

From our extensive contacts with the digital preservation community, we can verify that no other organizations are working towards generic standards for preservation or for interoperable protocols to provide these capabilities to the space agencies or any archival organization, although there is some complementary domain specific work.

So as you see, the answer to your general question about the relevance and role of CCSDS in the “big data” world is easily apparent and profound: First enable broad, open access to space mission “big data” archives to both space program insiders and outsiders, while preserving the data for generations to come. And then to share that capability with the world at large to enable digitally-driven cross-discipline technological advancement of mankind, and contribute to solving the “digital dark ages” problem.

Some other benefits of the DAI WG approach are:

* Development with the ISO community spreads the cost of development to other organizations outside the space agencies such as members of DAI from the US National Archives (NARA), National Libraries such as the Dutch (KB) and French (BnF), museums, Universities and private companies such as GAEL.
* It expands the global relevance of our spaceflight missions to the broader data industry, and contributes to the best kind of technological advances (cross-discipline).
* It will raise the stature of CCSDS as a great leader in interoperability standards.

Granted that the DAI WG members may not be objective with this next statement, but our assessment is that the work described here is the most important work in CCSDS because over the long term (for future generations) it is the most beneficial to the space agencies and the world at large. And we are perplexed about why space agency support has dropped off so greatly in this area, compared to the broad support that was provided for the Panel Two work early in the CCSDS era.

# Question 1: The DAI Roadmap

What is the current DAI roadmap and how does it make a valuable contribution to the data issue, taking into account what has already been produced including the overall worldwide efforts in this area?

First, some history. DAI has previously placed the greatest emphasis on the core underlying processes for digital preservation (Reference Model for an Open Archival Information System (OAIS), Information Preparation to Enable Long Term Use ([IPELTU](http://cwe.ccsds.org/fm/Lists/Projects/DispForm.aspx?ID=368&Source=http%3A%2F%2Fcwe%2Eccsds%2Eorg%2Ffm%2FLists%2FProjects%2FAllItems%2Easpx&ContentTypeId=0x0100B63160D64FE81342BE42A874DE7E703D))) before addressing interoperable protocols or interface specifications extensively. Even so, some interface specifications were produced (Producer-Archive Interface Specification ([PAIS](https://public.ccsds.org/Pubs/651x1b1.pdf))). The relationship between the DAI Standards and Recommended Practices completed or under development is shown below, supporting the flow of data from creation to transfer to the archive to preservation.

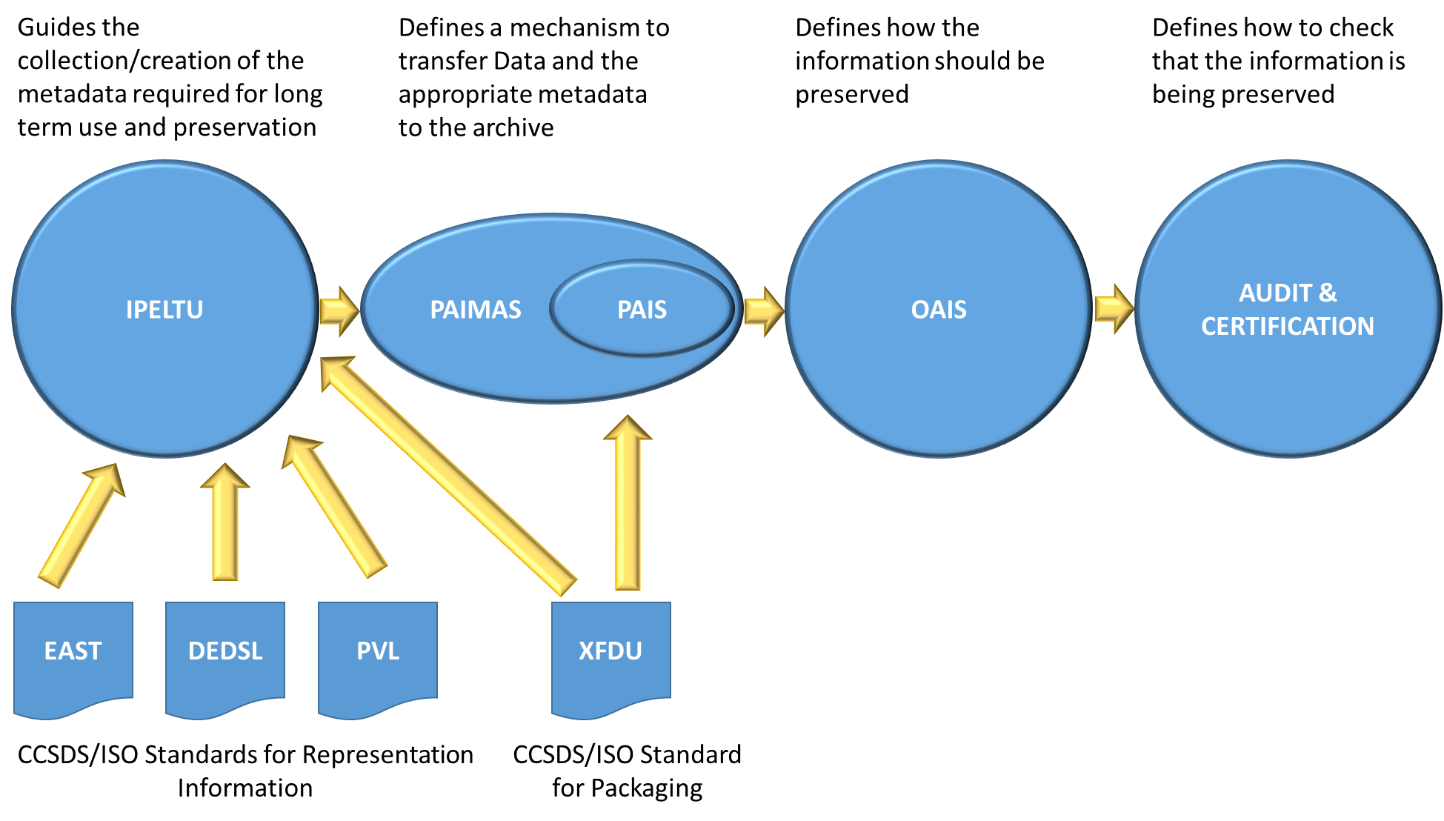


Figure Relationship between DAI standards and Recommended Practices

The roadmap for the DAI WG planned work has two components: Data architecture and schedule.

## The DAI vision for architecture and protocols

Notable characteristics of the planned protocol architecture (Figure 2) are:

* User interfaces (Producer and Consumer) at the top.
* An Archive Abstraction Layer in the middle (similar to the MAL from the SM&C WG). Our initial thoughts are that this would include methods to
  + Take in data, Representation Information and other metadata from the Producer in order to create Archival Information Packages (AIPs), the key component for long term preservation.
  + Query archive about AIPs available.
  + Access an AIP.
  + Send metadata from the AIP to the user interface.
  + Send user selection of data to be retried from the user interface to the archive
* Further illustrations are provided in the associated PPT.
* Bindings on the bottom to interface to specific types of archives.
* Legacy interfaces to current archives are shown on the right side. The new architecture can be an add-on to legacy systems, and operate in parallel.
* The overall framework would apply to all OAIS repositories/archives around the world. And they will help develop it. Under the CCSDS umbrella, we can be assured it will meet the needs of the Space Agencies.
* Unique “bindings” may actually be APIs or protocols. Only the bindings specific to spaceflight mission archives need be developed with the participation of the CCSDS space agencies. Other bindings can be developed elsewhere in the industry. For example, the Web/HTML binding may be developed by the IIPC (International Internet Preservation Consortium) but will be available to space agencies (and their national archives) for their web preservation purposes.
* Four example bindings are illustrated, but we anticipate that many other organizations would develop bindings for their specific data types or communities.
* So the users (producers and consumers) have a standardized interface that is stable, and they have the ability to “plug in” bindings for many different data types to access data in other communities besides their own.

In a few words, all-inclusive, all-serving, CCSDS-compatible and modular. And supportive of cross-discipline research.

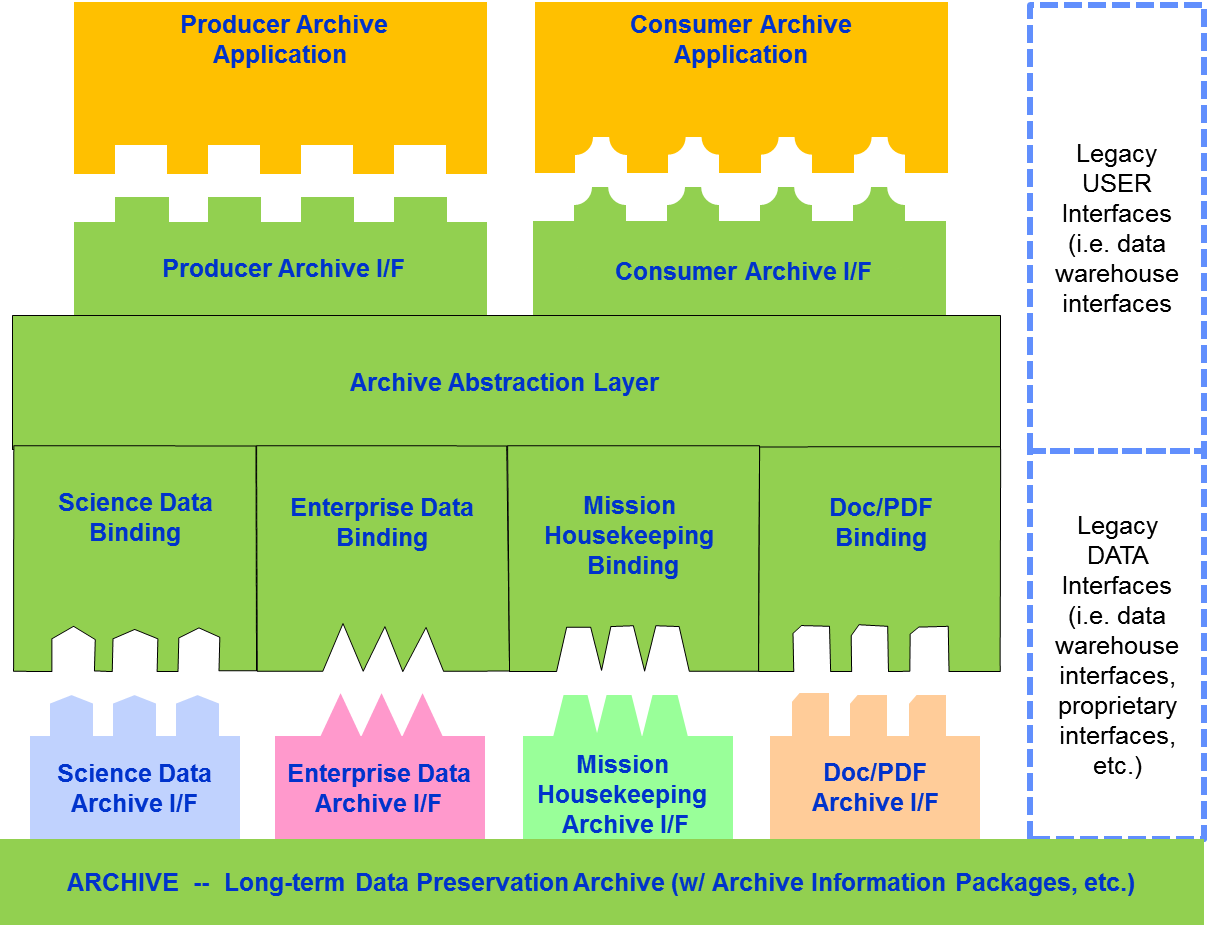


Figure Abstract Protocol Architecture

We are not locked in to this architecture. We are not investing significant resources to refine this current architectural concept until it has been vetted with both the CCSDS and ISO communities and the review with ISO communities will take some time.

In particular, current pressures on DAI to address distributed archives and cloud-based archives may change this architecture concept and the associated document tree.

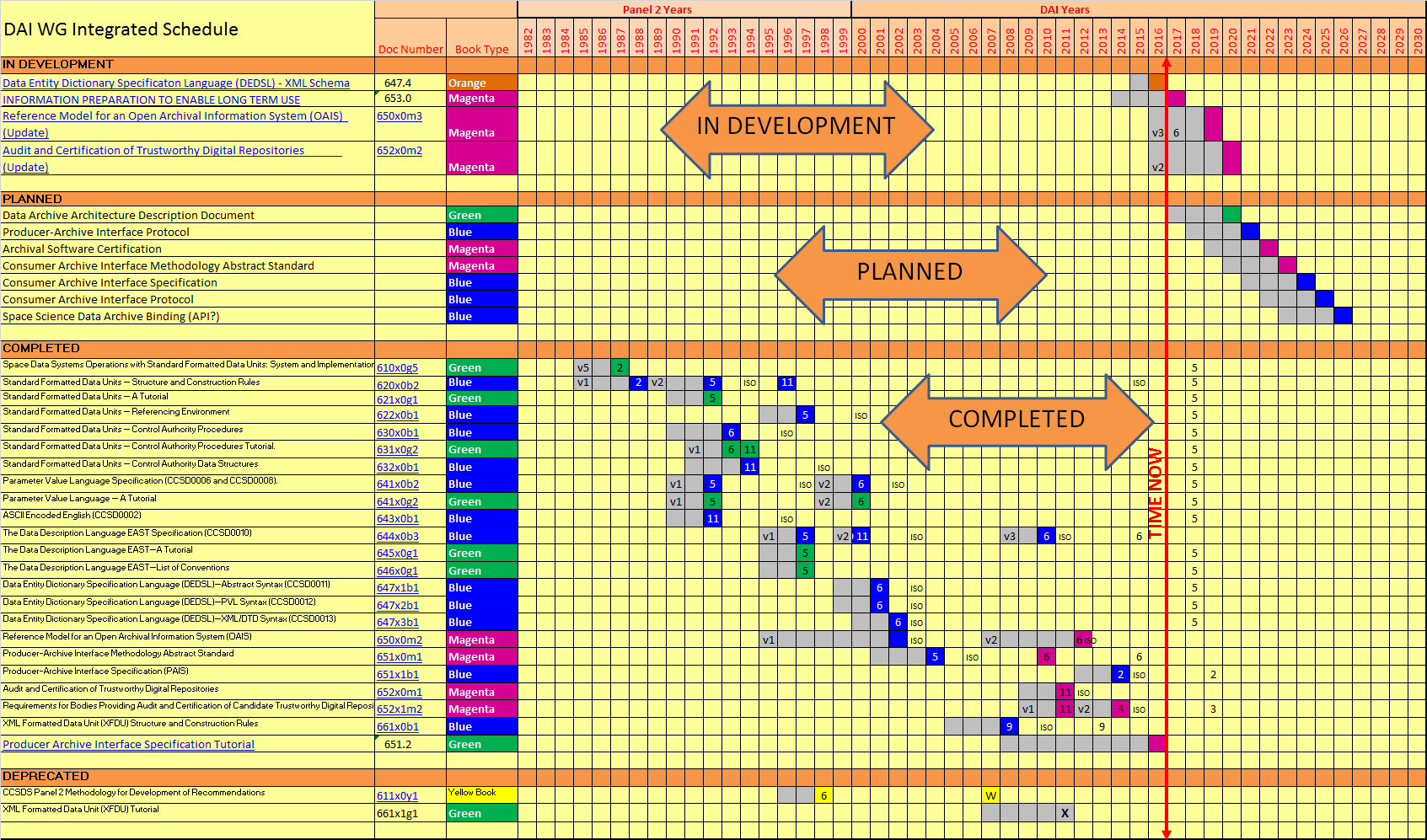
If this brief introduction to the long-range vision of Digital Preservation work in CCSDS is not self-explanatory, we would be happy to again attempt a special topic presentation to the CMC at an upcoming telecon or meeting.

The DAI WG intends to “socialize” this concept with other Digital Preservation organizations around the world, and to hopefully get their support to this as a long-range industry-wide vision. A CMC resolution supporting this long range direction (while not yet approving specific projects) would be a welcome support to this effort.

## The DAI Notional Schedule

In the DAI summary notional schedule below we cover historic, current and future work items. For the “road map” you can see where we must finish current reviews/projects before we can start on the new architecture/protocol work. The “future” tasks can be mapped to specific features of the architecture in Fig 2. We realize that the mapping is not included in this paper for the sake of brevity, but we can provide that if desired by the CMC. The major points on this schedule are:

* Each of these documents must of course follow the normal workflow approval process. If the CWE project of the first of the future documents, “Data Archive Architecture Description Document”, is approved, we intend to produce a draft of that book after one year so that this can help to justify the approval of the second project
* We anticipate a pace of about a document per year, with 3 or 4 documents in work at any given time.
* Some may be able to be combined, which could shorten the delivery cycle.



# Question 2: Other Organisations and DAI WG Role

Taking into account that there are dedicated international organisations working on archiving and merging scientific data, astronomy data, as well as Earth data, what should be the role of CCSDS in general and that of the DAI WG in particular?

It is quite true that scientific communities have for quite some time been working nationally and internationally to define adequate standards for their data, be it astronomy, planetary exploration or earth observation. However, all of the ones that we know about are defining formats and ontologies that are discipline-specific. They are working only in their “designated communities” (in OAIS terms) and they are not addressing the proper storage and preservation of data in trustworthy digital repositories. These domain specific standards complement the CCSDS standards for digital preservation. In particular OAIS defines “Representation Information” as the information needed to understand and use the data being preserved. The domain specific formats and ontologies are specific examples of Representation Information.

A survey was conducted using the MOIMS-DAI extensive contacts in these many areas.

Some specific results of our survey are:

* The SM&C “Archive” service refers to short-term storage capabilities rather than to a Long Term OAIS
* Open Geospatial Consortium (OGC); The list of OGC standards is available at <http://www.opengeospatial.org/standards>. While there are OGC standards unique to earth science topics and that touch on digital preservation for the vertical applications of that domain, there is not a general standard that supports digital preservation (the main focus of DAI). For example, their APIs are specific to geospatial data. We have some common membership in both the DAI WG and the OGC, so there is no risk of losing sight of both efforts.
* ISO TC211/WG7 also produces Geographic Information standards. They have incorporated OAIS into much of their work, and our assessment is that our work is complimentary. In particular in the workplan of TC211 the only reference to preservation is to ISO/CD 19165 “Geographic information – Preservation of digital data and metadata” in which they say: ‘This standard has a close relation to the standards developed by the ISO/TC 211 “Geographic information / Geomatics” and to archival standards of ISO such as the ISO 14721 “Open archival information system (OAIS) – Reference model”’. Moreover OAIS is referenced as a normative standard and much of the glossary is taken from OAIS.
* International Internet Preservation Consortium (IIPC); The IIPC conference and the Preservation Working Group in April 2016 was attended by DAI representatives. During the conference the community expressed interest in developing tools, but when the question was raised about developing standards, the response from the community was negative. No activity has begun.
* Object Management Group (OMG); While the OMG has listed “Archives” on their list of future efforts, they have not begun any work, and there is no schedule for starting. The topic for them is simply spacecraft telemetry archives (not other space mission related archives) and they are not addressing long-term digital preservation methods. This info is from a phone call to the co-chair of the OMG Space Domain Task Force.
* The Research Data Alliance (RDA) has many groups, two of the most relevant are the [Active Data Management Plans Interest Group](https://www.rd-alliance.org/groups/active-data-management-plans.html) and the [Preservation e-Infrastructure Interest Group](https://www.rd-alliance.org/groups/preservation-e-infrastructure-ig.html). Both these groups have common members between each other and with MOIMS-DAI. This ensures the consistency of the work of the groups. The roadmap of complementary work here is that RDA groups will help in the review and implementation of the CCSDS standards. Other RDA groups are developing domain specific Representation Information.
* ISO TC46 and the BnF (National Library of France) draft for ISO/CD 20614 “Data Exchange Protocol for Interoperability and Preservation” which aims to provide a standardized framework for the various information exchanges (data and its metadata) between the repositories and their partners. This document has OAIS, as well as PAIS and PAIMAS, as normative references. It is worth noting that it states that “The standard is limited to information exchanges between the Repository and its partners”.
* Long Term Data Preservation framework (LTDP); The goals (see <https://earth.esa.int/web/gscb/ltdp>) are;
  + To preserve the European, and Canadian, EO space data sets for an unlimited time-span.
  + Ensure and facilitate the accessibility and usability of the preserved data sets respecting the individual entities applicable data policies.
  + Through the adoption of a cooperative and harmonized collective approach among the data owners (LTDP Framework) based on the application of European LTDP Common Guidelines and sustained through cooperative (multi-source) long term funding schemes.
  + Ensure, to the maximum extent, the coherency with the preservation of other non-space based environmental data and international policies..

The LTPD provides a number of guidelines in order to harmonize the approach, including a checklist for the types of information which should be preserved. However the work is very much based on OAIS rather than seeking to replace it, for example “Mission Documentation shall include Representation Information, Packaging Information and Preservation Descriptive Information according to OAIS information model [RD-6].

* Pérennisation des Informations Numériques (PIN), **net**workof expertise in long-term **stor**age **(nestor**), and the Digital Preservation Coalition (DPC) are national membership organisations in France, Germany and the UK respectively which provide fora for discussing digital preservation. None are developing standards and in fact all have made significant numbers of suggestions for the update of OAIS.
* The Group on Earth Observations (GEO), has been developing the Group on Earth Observations System of Systems (GEOSS) and relevant documents, such as the Data Sharing Principles and Data Management Principles, as well as Implementation Guidelines for each of these sets of principles. These documents are not standards, but build on existing standards, such as OAIS, and therefore, are complementary to the DAI efforts. The DAI WG also has a member that participates in GEO, which helps maintain coordination and prevent overlap.
* LOng Term Archiving and Retrieval (LOTAR) International which describes itself as “an international collaboration of the Aerospace Industries Association (AIA) and the Aerospace and Defense Industries Association of Europe for Standardization (ASD-STAN) under the supervision of the International Aerospace Quality Group (IAQG). Key enabler of the LOTAR project are the ProSTEP iViP Association and the PDES, Inc.”. It divides its work into the following (our comments are in italics):
  + Basic Parts: provide an overview about the essential information on long time archiving and retrieval of 3D model and product structure data.
    - *This makes extensive use of OAIS concepts, for example the “Metadata for Archival Package” Workgroup”* [*http://www.lotar-international.org/lotar-workgroups/metadata-for-archival-package.html*](http://www.lotar-international.org/lotar-workgroups/metadata-for-archival-package.html) *is developing domain specific OAIS Archival Information Packages (AIPs).*
  + Common Process Parts: define the main process steps of archiving and retrieval of 3D model and product structure data in a detailed and common manner.
    - *This is Representation Information specific to 3D models*.
  + Data Domain Specific Parts: contain requirements and definitions for specific types of data related to long time archiving and retrieval. So far descriptions for CAD 3D model data and Product Data Management (PDM) data exist
    - *This creates specific types of Representation Information*
* International Virtual Observatory Alliance (IVOA) describes itself as follows: “The Virtual Observatory (VO) is the vision that astronomical datasets and other resources should work as a seamless whole. Many projects and data centres worldwide are working towards this goal. The International Virtual Observatory Alliance (IVOA) is an organisation that debates and agrees the technical standards that are needed to make the VO possible.” IVOA has produced and continues to work on these standards – see <http://ivoa.net/documents/index.html>. These standards provide data formats such as VOTable and data models (i.e. Representation Information) specifically useful to astronomers, as well as protocols for accessing and using services for astronomical data, which we will factor into our proposed new work.
* The Metadata Encoding and Transmission Standard (METS) group, hosted by the Library of Congress, (<http://www.loc.gov/standards/mets/>) is, as the name suggests concerned with metadata transmission, and makes extensive use of the OAIS concepts and much of its work defines specific Representation Information.

Those are the principal entities working in the Digital Preservation Arena that might possibly have been embarking on work that could overlap the DAI. However indications are in each case that they are not seeking to replace OAIS, but rather they are all using OAIS as their foundation. Several are developing protocols for exchange of information between repositories (ISO TC46) or between repositories to applications (IVOA, METS). However none are developing interoperable standards that specifically promote interoperability between trusted OAIS repositories between agencies and across disciplines. Therefore, within the data industry, the DAI work supports these unique roles for CCSDS:

* Preserve the space agencies’ digital data assets for future generations. (With a general solution that also meets global needs of many nations and industries.)
* Enable access across disciplines; Space researchers accessing other disciplines and other discipline researchers accessing the science and mission benefits of our space programs. (Again, with a general solution that also meets global needs of many nations and industries.)

That covers the field of all organizations working in digital preservation and archive interoperability that are known to the members of the DAI WG. If the CMC has any other suggestions for organizations that are potentially duplicative of the current or planned work in the DAI WG, we would be very interested in learning about them, not only for the CMC objective of insuring no duplicative work, but also for the DAI WG members’ objectives of keeping abreast of developments in the field of digital preservation.

# Question 3: Preservation of Housekeeping Data

Is the long term data preservation of the housekeeping data properly addressed in these initiatives? Is there a need for developing this part of the data preservation?

The principles of data preservation that have been pioneered by the DAI WG are equally applicable to all data types including housekeeping data. The OAIS processes, the Producer Archive Interface Specification, the future plan for all Digital Preservation protocols all apply equally to spacecraft science data and housekeeping data. They also apply as well to space program enterprise data (CAD, requirements, engineering documentation) and even more broadly to all other data types. The well-developed techniques and anticipated protocols are expected to apply equally to science data, housekeeping data and all other data types that require preservation.

The area where science data and housekeeping data will diverge is in their Representation Information. Some metadata would be common, but some would be specific to the nature of housekeeping data.

For example the Housekeeping data may be well described using the EAST standard (CCSDS 644.0-B-3 The Data Description Language EAST Specification) and CCSDS 647.3-B-1Data Entity Dictionary Specification Language (DEDSL)—XML/DTD Syntax, as its Representation Information, its Provenance encoded in the Open Provenance Model etc, so that an Archival Information Package can be created to ensure the preservability of this data.

When the Housekeeping data is required to be used some time later, when the original systems and even the operators and system designers are no longer available, then the EAST and DESL descriptions and other documentation would allow one to extract the information – numbers and units - from that Housekeeping data. The Housekeeping data may then allow one to correlate, say, instrument temperatures with scientific results in order to improve the calibration of the scientific output from the Agency’s satellite.

It is remotely possible that when DAI seeks to develop the bindings/plug-ins/APIs/protocols as part of the long-range architecture that perhaps a separate protocol will be optimized for housekeeping data. Generally, it is anticipated that all telemetry types would have a common protocol. But that distinctive housekeeping work, if necessary, is also expected to be under the DAI WG umbrella.

# Question 4: Next Steps

In this context, what would be the next project for the DAI WG in the future, should resources be available, and how would this project fit with the current CCSDS reference architecture and work plan?

Following our roadmap as described in the response to question 1, the next steps are:

1. Finish the [OAIS](http://cwe.ccsds.org/fm/Lists/Projects/DispForm.aspx?ID=581&Source=http%3A%2F%2Fcwe%2Eccsds%2Eorg%2Ffm%2FLists%2FProjects%2FAllItems%2Easpx&ContentTypeId=0x0100B63160D64FE81342BE42A874DE7E703D) and associated simultaneous [Audit/Certification](http://cwe.ccsds.org/fm/Lists/Projects/DispForm.aspx?ID=582&Source=http%3A%2F%2Fcwe%2Eccsds%2Eorg%2Ffm%2FLists%2FProjects%2FAllItems%2Easpx&ContentTypeId=0x0100B63160D64FE81342BE42A874DE7E703D) reviews that are underway.
2. Finish the [IPELTU](http://cwe.ccsds.org/fm/Lists/Projects/DispForm.aspx?ID=368&Source=http%3A%2F%2Fcwe%2Eccsds%2Eorg%2Ffm%2FLists%2FProjects%2FAllItems%2Easpx&ContentTypeId=0x0100B63160D64FE81342BE42A874DE7E703D) document that is underway.
3. Begin on the Digital Archive Architecture Description Document (ADD). Development of an ADD is a widely recognized CCSDS practice to support a broad long-range plan of protocol development in other working groups (DTN, etc.) and the DAI WG objective is no different. (Note the title may change in future DAI WG deliberations, but the purpose will remain essentially the same.) DAI will of course take into account other work done in CCSDS including that carried out on System Architecture and SM&C Architecture.
4. Complete the “Producer Specs” with the Producer Archive Interface Protocol either as a standalone document or as an update to the PAIS standard.
5. Develop the Archive Abstraction Layer (the equivalent of the Message Abstraction Layer developed by the SM&C WG). As noted above our intention would be, if the CWE project is approved for the Architectural Design Document, to complete an initial draft in the first year which could then be used to justify the approval of the next CWE project in our roadmap.
6. Begin on the “Consumer Specs” branch of the document tree that is the equivalent to the “Producer Specs” that have already been published.
7. Develop the bindings/plug-ins/APIs/protocols that are specific to the spaceflight industry (telemetry archive protocols and anything else that is needed).

Also, in the timeframe of the development of the Digital Archive ADD (or perhaps earlier) we expect to begin “socializing” the long-range DAI WG plan with other organizations in the Digital Preservation arena. This will be to garner broad support for the overall development plan and to insure that there are not conflicts or overlap with work in other organizations. It should, as well, encourage resources from other organizations to help with the archive framework efforts, and to encourage other disciplines to develop their own bindings/plug-ins/APIs/protocols that are compatible with the framework and which enable cross-discipline research and operations of trustworthy digital archives.

# Concerns from the DAI WG

The DAI WG has some high-level management issues that require CMC attention

## CMC Is Encouraged to Improve Space Agency Support to DAI

As in all CCSDS WGs, there are two aspects of support that is sought from the CCSDS Space Agencies: Adoption and Development. In the case of the DAI WG, we feel that the real work of the DAI WG, ***digital preservation,*** is not well understood by the space agencies’ participants that have a stronger focus on space communications than on mission applications. Hence the space agencies have faltered in their support to the DAI WG.

ADOPTION: First, we encourage the CCSDS space agencies to improve their agencies’ adoption of the published DAI standards.

DEVELOPMENT: The space agencies are encouraged to improve their support to the DAI team in developing the processes and protocols for future interoperable capabilities. The science data and engineering data investment of our agencies’ spaceflight missions are costly to acquire. The final “put to bed” costs of long-term digital preservation are minor in comparison. Besides those obvious benefits of long-range digital preservation within an agency for all data types (from science to enterprise data), the interoperability capabilities that will result will also benefit each agencies’ researchers that need to cross disciplines and access science results across a broad spectrum of science disciplines. (Given access permissions, of course.)

Considering the large storehouse of science data from decades of past space missions, what could be more important?

## CMC Position on Broader Participation for the Broader Good:

CCSDS programs such as DAI and DTN that address the needs of mankind beyond the space agencies is a very good thing for CCSDS. The outside-applicability of the DAI work is a parallel to CCSDS DTN which is now evolving to an IETF standard for the broader Internet for the public at large.

It is true that many other organizations that are not space-related will benefit from the CCSDS DAI work; National Archives, Libraries, etc. We think this is positive for the global recognition of CCSDS stature, as well as for the pragmatic goal of getting work to benefit the space agencies by using “outside resources”.

We think these standards that become applicable outside of the space community are a great thing for CCSDS, but we would like the CMC to confirm their agreement. If the CMC thinks this is not positive, the DAI WG would like to understand the CMC rationale.

## CMC Expression of Support for the ISO Community

DAI WG needs an expression of commitment from the CMC because the external (ISO) community has expressed concern that the CCSDS space agencies may not fully support their needs. These concerns were expressed as part of the ongoing OAIS review, as well as during conferences and on online blogs. The DAI WG feels that this concern would be eliminated by a resolution from the CMC expressing that CCSDS fully intends to embrace ISO participation in CCSDS reviews, and we intend for the process to be inclusive and transparent.

Compared to other CCSDS WGs, DAI has a large ISO (non-space-agency) participation. If these ISO participants submit their comments only during the ISO review, it would cause extensive delays and rework. DAI really needs ISO participation to happen early, during the normal CCSDS agency review. That’s the reason for this DAI process that is “different” from other CCSDS WGs, and the reason we need this resolution. A proposed draft resolution is provided at the end of this report.

## DAI Working Group Efficiency

DAI WG has been inaccurately accused of inefficiencies in CESG discussions. We would like to put this topic to rest by showing the schedule of past accomplishments (integrated with the current future schedule). We believe that this illustrates a continuous strong pace of productivity when compared to other CCSDS WGs. Also, it should be noted that this is a large output using some external resources. More than any other CCSDS WG, this work uses resources outside space agencies’ budgets, hence it is very efficient for the space agencies.

## DAI Roadmap Details

DAI WG has a new long-range notional plan, and we seek CMC notional approval (understanding that project approvals will come much later). The architecture is illustrated by the block diagram in the response to question 1 (the roadmap). DAI requests that the CMC respond with general agreement that this is a good baseline approach.

## Proposed CMC Resolutions:

The CMC recognizes that for the work of the Digital Archive Ingest Working Group, it is important for the ISO participation to be encouraged during the CCSDS Agency Review phase, in order to avoid extensive delays and rework in later ISO review phases. The CMC also recognizes that it is critically important to support this work for the Digital Preservation community much broader than the nominal participation of the CCSDS space agencies. The CMC therefore resolves to express support for the inclusion of the ISO community, and resolves to commit to an inclusive and transparent review process for the future DAI WG projects in standards for Digital Preservation.

Action to the CMC: (Unless they are already participating in the DAI WG) Please provide to the DAI WG the names and email addresses of:

1. Personnel in your space agency that are working in digital preservation or archives.
2. Personnel in your nation’s archives or library organizations that have responsibilities for digital preservation.

The DAI WG will contact them to ensure they are aware of CCSDS standards for digital preservation, and for them to consider participating in the DAI WG telecons, or to simply to suggest coordination when DAI WG personnel are visiting in the area (Such as ASI attendance at the last CCSDS meeting in Rome).

# General Comments by MOIMS AD/Deputy AD

* Please provide the current membership of the DAI WG and the available resources/sponsors.
  + The membership of DAI is available at <https://mailman.ccsds.org/cgi-bin/mailman/roster/moims-dai> - currently 86 members. A group of 10 people regularly attend the weekly Webex meeting, with others joining from time to time.
* The purpose of the proposed architecture is clear (universal ability to deposit/access any data to archives), however this is a very ambitious goal above all if you want to handle differently the data depending of their types (science, HK, CAD, …) providing specific interactions for the data type. For instance, the SM&C COM does this, but the solution is very simple with basic services. It is not understood how you plan to achieve your goals. Please provide clear and descriptive examples.
  + Brief details are provided in the text above. Since we have completed only the process stage (OAIS, IPELTU, etc.), and we will start the architecture stage in a number of months, and we will start the protocol/API/binding development in many years, it is not reasonable to ask for descriptions of specific designs for interaction patterns. We are at the conceptual stage where SM&C was in 2003-04, and at that time the SM&C team did not have such details. We are confident that even for more complex interactions, a Service Oriented Architecture with an abstraction layer is feasible because there are many examples in the other industries. A major factor in choosing this approach was compatibility with the overall CCSDS Architecture. Changing to some other technical approach would lessen compatibility with MOIMS and CCSDS. Therefore we have embarked on the maximum compatibility path. It is a little confusing to get this question from the CESG when generally the struggle has been to maintain a unified architecture approach with other areas. Yet we seem to be challenged for too much compatibility within MOIMS.
* What is the level of resources that you would need to achieve your plan?
  + The plan is executable with our current support levels (including space agencies, other ISO contributors and volunteers). However, some supplemental skills would probably need to be found once we are past the architecture stage and into the protocol/binding/API development stage. Initial discussions with Google may have possibilities. Also, recent discussion with JPL staff indicates that they may be increasing support (with SMD funding, not SCaN funding) and they could potentially bring such skills. In any case, it is very much too preliminary to expect a resource plan for documents that are not to be started for years to come. When the projects are submitted to the Framework, they will include the appropriate book editor and prototype resource plans at the right time. You mentioned that the DAI could also make use resources from sources external to CCSDS. Please indicate if you have already secure this type of resources. We have listed above organisations which have already contributed. We will be approaching others if these plans are approved.