

Creating a Framework for Applying OAIS to Distributed Digital Preservation

Eld Zierau

The Royal Library of Denmark
Søren Kierkegaards Plads 1
1219 København K
ph. +45 33 47 46 90
elzi@kb.dk

Matt Schultz

Educopia Institute
1230 Peachtree Street, Suite 1900
Atlanta, GA 30309
ph. +1 616 566 3204
matt.schultz@metaarchive.org

ABSTRACT

This paper describes work being done towards a *Framework for Applying the Reference Model for an Open Archival Information System (OAIS) to Distributed Digital Preservation (DDP)*. Such a Framework will be helpful for future analyses and/or audits of repositories that are performing digital preservation in highly distributed ways. A great deal of work has already been accomplished toward the Framework itself, including selective community reviews of a white paper, case study interviews with DDP practitioners, and an analysis of OAIS as it relates to DDP. The paper will provide background information on this set of work, describe the research carried out to-date, and explain the proposed Framework components, including concepts and terminology, placement of OAIS functional entities, and roles and responsibilities for carrying out DDP.

Categories and Subject Descriptors

A.1 INTRODUCTORY AND SURVEY

A.2 REFERENCE

E.2 Data Storage Representations: Linked representations, Object representation

E.5 Files: Backup/recovery, Optimization, Organization/structure

H.3.7 Digital Libraries: Collection, Standards, Systems issues

General Terms

Management, Documentation, Design, Reliability, Security, Standardization, Theory.

Keywords

Distributed Digital Preservation, OAIS, vocabulary, functional entities, roles and responsibilities, framework

1. INTRODUCTION

This paper presents work being done towards creating a *Framework for Applying the Reference Model for an Open Archival Information System (OAIS) to Distributed Digital Preservation (DDP)*. Such a Framework will be helpful for future analyses and/or audits of repositories that are performing digital preservation in highly distributed ways.

The use of distribution is actually quite commonplace within the preservation field, but there has not been a commonly accepted definition for “*distributed digital preservation*”. As the preservation field has matured, the term “*distributed digital preservation*” has been applied to myriad preservation approaches, spanning initiatives that intentionally build distributed infrastructures as essential components of their preservation models to other initiatives that merely engage in

back-up measures for their digital objects or performing all necessary preservation actions within one organization. The reasons for adopting distributed approaches can also be varied in focus and motivation (e.g. achieving cost effectiveness through shared resources [7], expanded support for large amounts of data [6], and general sustainability of the enterprise in the face of contingencies and threats [7]).

In this Framework project, we use the term Distributed Digital Preservation (DDP) to emphasize the practice of applying distribution in intentional ways, both organizationally and technically, for accomplishing digital preservation, for example through geographic distribution, infrastructure heterogeneity, and organizational diversity. A more concise definition, for DDP is: the use of *replication*, *independence*, and *coordination* to address the known threats to digital content through time to ensure their accessibility.

Worldwide, numerous digital preservation initiatives are already engaging highly distributed methodologies, infrastructures, and organizational apparatuses in order to achieve the reliable persistence of digital content. Examples of such organizational and/or technical initiatives include Archivemata (https://www.archivemata.org/), the Danish Bit Repository [5], Chronopolis (http://chronopolis.sdsc.edu/), LuKII (www.lukii.huberlin.de/), LOCKSS (http://www.lockss.org/), UC3 Merritt (https://merritt.cdlib.org/), Data-PASS (http://www.data-pass.org/), DuraCloud (http://www.duracloud.org/), MetaArchive Cooperative (http://www.metaarchive.org/), DPN (http://d-p-n.org/), Internet Archive (http://archive.org/), iRODS (http://www.irods.org/), and many others.

These existing DDP approaches, as well as those that have yet to come into being, currently lack common vocabulary and conceptual frameworks for building effective, reliable, and auditable distributed preservation environments. Such agreed-upon terminology and theoretical models would help initiatives to describe and compare their infrastructures and operations. They would also help to increase understanding and awareness of the DDP process both by practitioners and by organizations seeking preservation solutions. Finally, they would provide auditors (including self-auditors) with a crucial foundation for assessing the reliability of a variety of distributed approaches.

The preservation field today relies heavily upon the Reference Model for an Open Archival Information System (OAIS) to provide theoretical and abstract models and vocabulary for digital preservation [3]. This OAIS standard provides a significant portion of the foundation for the Framework this initiative intends to develop. The Framework will elaborate upon the OAIS model to define the methodology and structure of the actions undertaken in organizationally and/or technically distributed preservation

repositories. The Framework does *not* intend to redefine existing standards. Instead it merely seeks to elaborate with additional models and vocabulary.

2. FRAMEWORK BACKGROUND

The awareness of the need for this *Framework for Applying OAIS to DDP* has emerged independently both in North America and in Europe. In North America, discussions first began in early 2010 between MetaArchive Cooperative, Chronopolis and the Library of Congress. Later that year in October 2010, representatives from the Library of Congress and several LOCKSS-based groups met at the 1st Annual Private LOCKSS Network (PLN) conference. It was here that a “constellation group” was formed to discuss the issues and lay the foundation for work related to the Framework.

A formal Working Group was convened in early 2011 that could begin to document the full range of elaborations that may be needed to help apply OAIS to the DDP environment. In this period, the Working Group prepared a Statement of Purpose for the initiative, established a workspace to document DDP use cases and gap analyses¹, outlined a white paper, and solicited participation from a number of DDP practitioners.

Focused conversations began in early 2012 between the Educopia Institute (<http://www.educopia.org/>) and the Royal Library of Denmark (<http://www.kb.dk/>) to review the Library’s evolved model, known as the IR-BR model [9]. The Royal Library of Denmark is a pioneer in proposing this model, which is an approach to achieving reliable, auditable distributed preservation. The IR-BR model has a great deal of valuable concepts and terminology that present themselves as valuable to the Framework. There is more on the IR-BR model later in this paper.

The Working Group has now grown to include numerous well-known organizations that embody a wide variety of use cases for DDP. These organizations include Archivemata, Chronopolis, Data-PASS, the Danish Bit Repository, DuraCloud, Internet Archive, LOCKSS, MetaArchive Cooperative, and UC3 Merritt

2.1 The Need for a Framework

It is important first and foremost to acknowledge that the Reference Model for an Open Archival Information System (OAIS) does not assume any specific technical or organizational infrastructure, but rather seeks to abstract out the functional and information package requirements that should be achieved in any implementation. It also describes the roles and responsibilities that an archive must undertake, but does not describe where those responsibilities reside, either at an organizational or physical/geographical level.

In practice, some institutions have taken centralized approaches to building an archive that conforms with OAIS - which is to say that their digital objects are ingested and registered into storage resources residing at one geographic and organizational location and stewarded by one organizational center. In this paradigm, the locus of concern, responsibility, and implementation is highly centralized.

Other initiatives have taken distributed approaches to building a repository or network - meaning their digital objects may be ingested and registered into storage resources that reside in multiple geographic locations and that may be stewarded via the use of various distributed services by multiple organizations, all in

order to accomplish effective bit preservation. Likewise, a strong case can also be made for the importance and relevance of distribution for the proper hosting, maintenance, and application of services like format identification, validation and migration/normalization guidance (e.g., format registries such as PRONOM, UDFR, etc.). As discussed further in Section 3.3.3. Terminology below, in this paradigm responsibility may be decentralized and span multiple physical and/or institutional locations. Although OAIS is in no way antithetical to such DDP approaches, it does not explicitly define or describe how OAIS principles and models map onto these distributed infrastructures--be they organizational, geographical, or systems-based - each of which protect against different risks and can even be used in combination. A common framework can help to define and make sense of the proper coordinations.

Although OAIS begins to consider issues of interoperability between separate archives (Section 6: Archive Interoperability), it does not explicitly address the range of interactions that may occur between separate organizational entities as part of the work of one distributed digital preservation repository (e.g., collaboratives that share archival management across multiple, distinct institutions (e.g., through the use of embedded peer-to-peer software, distributed micro-services, geographically dispersed Cloud storage services, or other configurations) or even a single organization that manages an archive comprised of distributed infrastructure components—again such as CDL & Archivemata’s use of distributed micro-services.

For this reason, early DDP practitioners have encountered the need for additional documentation to describe in greater detail the different functions, roles, and responsibilities that emerge in this distributed landscape. Such a Framework will be helpful for future analyses and/or audits of repositories that are performing digital preservation in highly distributed ways.

2.2 Research Methods

To approach the development of the *Framework for Applying OAIS to DDP*, a number of research activities have and are being undertaken. These include 1) the development of a white paper that was disseminated for peer stakeholder review; 2) a thorough set of case study interviews with several diverse DDP practitioners; 3) a detailed analysis of the Reference Model for an Open Archival Information System with an eye toward bridging gaps in concepts and terminology, proper positioning of functional entities, and roles and responsibilities for DDP; and 4) a review of literature related to DDP and OAIS.

2.2.1 White Paper

The development of the white paper was intended to make a clear case for the need for a Framework for Applying OAIS to DDP. It covered much of what was addressed in the previous Section 2.1 of this paper. It proved to be an extremely useful resource for focusing the proposed work for building the Framework itself and for disseminating information about the Framework amongst numerous peer stakeholder DDP groups and digital preservation experts.² Feedback garnered from the white paper review has already played an instrumental role in refining the Framework contents and outline (see Section 3 below).

¹ See http://www.loc.gov/extranet/wiki/osi/ndiip/ndsa/index.php?title=DDP_OAIS_Frameworks (please contact matt.schultz@metaarchive.org to request access to the NDSA wiki).

² To request a copy of the Framework White Paper, please contact Matt Schultz or Eld Zierau (mails at top).

2.2.2 Case Study Interviews

Case study interviews were also carried out with each of the Working Group partners, each of whom were asked a series of consistent questions, namely:

- What elements of your organization are distributed? (e.g., management, storage infrastructure, preservation services)
- What has been most challenging about working in these distributed ways?
- If you have used any audit tools (OAIS, TRAC [1], DRAMBORA [4], etc.) have there been any gaps between the concepts and terminology in these tools and the ways you perform your distribution?
- What shortcomings, if any, do you see with these audit tools as you have applied them to your distributed environment?

The responses to these questions were recorded and used to identify distinctive technical and organizational qualities and characteristics of DDP that could be highlighted for the Framework. These elements are also discussed in Section 3 below.

2.2.3 OAIS Analysis

Effectively documenting, within the Framework, the relevant portions of OAIS that have bearing on DDP (and vice versa), is contingent, not only upon our case studies, but upon a thorough analysis of OAIS itself. At the core of OAIS are a set of digital preservation concepts and terminology, functional entities, and roles & responsibilities. The Working Group is reviewing each of these core elements of OAIS and searching for both associations and gaps with respect to DDP as it is best defined both theoretically and through existing case studies.

2.2.4 Review of Literature

Much documentation has been undertaken to describe the proper application of OAIS to digital preservation workflows and repositories. Similarly, much has been written about various distributed implementations for digital preservation. It will be vital that we incorporate relevant information from all such existing publications to ensure proper context, continuity and intelligibility of the Framework for its intended audiences. Users of the Framework will benefit from its association to this broader corpus of information.

3. FRAMEWORK – NOW AND LATER

So far there have been identified a number of topics that are needed in order to make a Framework that can be helpful for future analyses and/or audits of repositories performing distributed digital preservation. This includes *terminology* for DDP and set forth a series of *higher-level concepts, principles* and *guidelines*. Below is a description of the intended audience for the Framework, its scope and publication possibilities, as well as an outline of the Framework's primary proposed components, and the existing and intended contents.

3.1 Intended Audience

There are multiple audiences for this Framework that will need to be kept in view. They span a range of institutional stakeholders responsible for and concerned with the persistence of digital information as well as their designated communities, including governmental agencies, digital libraries and archives, and research data curators, among others.

The primary audience - spanning each of these stakeholders - is those organizations that are seeking to jointly develop or enhance distributed digital preservation (DDP) systems and are in need of

guidance on responsible ways of doing so. A second, and equally important group consists of auditing bodies that are seeking to evaluate such DDP systems, and could benefit from the elaborations and interpretations provided by the Framework. Finally, there are the organizations that are seeking to deposit their digital objects in such systems and seeking to understand their operating principles. There may also be other audiences, including those seeking to access and use the digital objects.

3.2 Scope

The Framework currently seeks to address first and foremost the various areas outlined below with respect to both DDP and OAIS. As noted below in Section 4. Discussion and Further Work, focused attention will be given during the later drafting and review stages as to how best to address the relationship of distribution to functional preservation services beyond those of a more generalized repository implementation, which remains the primary focus of the Framework. Such functional preservation services are integral to an overall digital preservation endeavor and encompass things like format registry services and how these may or not be managed and hosted by multiple organizations (e.g. PRONOM, UDFR, etc.). What is likely to be somewhat out of scope (at least initially) is the incorporation of lessons from the general open source software community where collaboration, sustainability, and extensibility (as opposed to hosting and maintaining distributed technologies and administrative resources) are more at issue. The open source community remains important and such lessons are likely to be full of useful insights. Though they are outside of the immediate focus and somewhat broad for the purposes of the immediate Framework, efforts will be made to study what such communities can contribute to the final Framework.

3.3 Publication

This Framework could potentially take numerous forms. One exemplar that already exists is the *Producer-Archive Interface Model Abstract Standard (PAIMAS)* [2], which is a supplemental standard to OAIS itself. Taking this document approach would require review and approval by the Consultative Committee for Space Data Systems (CCSDS) and the International Organization for Standards (ISO). This may not be the most appropriate status for the work as it is currently being proposed. The Framework could perhaps more appropriately exist as a simple community-reviewed document or publication hosted and made available by a respected organization. In addition to such traditional document forms, the Framework could also exist as a modular web resource. There may be other forms. The Working Group will continue to discuss the proper publication forms as the Framework drafting proceeds. Drafting of the Framework is scheduled to be undertaken by the Educopia Institute and the Royal Library of Denmark (financed by the Danish Ministry of Culture) in concert with the Working Group throughout the remainder of 2013, with community reviews scheduled for Fall/Winter 2013.

3.4 Outline

The Framework outline has been developed on the basis of findings from the case studies and reviews, suggestions from the Working Group, stakeholder reviews of the white paper, as well as preliminary analyses of those components of OAIS that are most relevant to DDP (and vice versa). Based on this outline the Framework proposes to cover the following elements.

3.4.1 Introduction

This section will concisely state the purpose and rationale for the Framework and provide an overview of the components.

3.4.2 Background & Overview of DDP

The Framework proposes, first and foremost, to put forth a thorough set of background and overview information on distributed digital preservation (DDP) on its own terms as it has evolved to distinguish itself from more centralized approaches to digital preservation where digital objects are ingested and registered into storage resources residing at one geographic and organizational location and stewarded by one organizational center. This will be helpful for orienting readers to this unique environment, and set the boundaries for effectively applying OAIS for the purposes of the Framework.

3.4.3 Terminology

Many of the case studies have underscored the importance of having improved terminology to describe the parts of a distributed system. Some reviewers voiced support for missing terminology that presents itself as very DDP specific, like some of the proposed supporting terms described below. As mentioned in the introduction, this is more of a task of extending the terminology of OAIS, rather than prescriptively redefining any of the existing OAIS terminology. Nevertheless, where it seems appropriate and helpful, urging new or improved standardization of terminology for the digital preservation community will not be avoided.

So far, the need has emerged for two types of terms to be defined:

- *Major Terms* like the definition of Distributed Digital Preservation, which in its simplest form is the one given in the introduction based on use of *replication*, *independence*, and *coordination* (expanded definitions will also be included in the final Framework)³.
- *Supporting Terms* which include what is meant by *independence*, and *coordination*, as well as terms like *storage unit*, *storage node*, *storage environment*, *cache/pillar*, but also broader definitions of OAIS terms like *replication* and *disaster recovery*.

DDP's definition derives in part from the matured experience of performing proper bit preservation, where bit safety is best ensured by the proper coordination of independent replications of data. The independence of replications requires that replicas be distributed geographically and organizationally and coordinated through timely and effective integrity checking. DDP's definition is also derived in part from the maturing experience of performing functional preservation where sustainable preservation is dependent on shared knowledge and solutions for how and when to do format migration. Here the responsibilities of format registries and the development of trustworthy services (e.g. a migration micro-service) may best rely on their hosting at different organizations. See Section 4. Discussion and Further Work for how this area of distributed functional preservation will be explored throughout the development of the Framework.

A typical example of a supporting term in DDP are any one of those used for *Storage Unit*, *Storage Node*, *Storage Environment*, *Cache*, and/or *Pillar*. This is the unit that forms the basic storage for a copy of data in a bit repository. This unit is based on specific

technology with an organisation around it which can be responsible for basic operations, technology watch, etc. The difficulty here is to arrive at more standardized usage, since most of the terms have other meanings as well, and therefore can be misinterpreted in other contexts.

Replication and *Disaster Recovery* are examples of supporting terms that are extension of OAIS terms. In OAIS the term *Replication* primarily has a migration context and meaning, whereas in the DDP environment it is more likely to have a proactive integrity measure context and meaning that is closer to OAIS in its treatment of *Disaster Recovery*. However, disaster recovery as a primary context and meaning for replication also does not do full justice to its usage in the DDP environment. The Framework will make this clearer.

3.4.4 DDP and OAIS Relationships

The Framework can add value for both DDP and OAIS by focusing on the responsible use of distribution at both technical and organizational levels. DDP and OAIS relationships will be related to one another in the Framework by bridging DDP perspectives with OAIS perspectives, especially through analysis of the proper positioning of functional entities, and roles and responsibilities.

- *Functional Entities* will describe scenarios for the placement of the components of an OAIS functional entity across distributed environments, including those that entail the coordination/collaboration of multiple organizations. An example can for instance be found in the IR-BR model (see *Models* below).
- *Roles & Responsibilities* will describe the roles and responsibilities at institutions that are replicating and preserving digital information in geographically and/or organizationally distributed ways.

3.4.5 Case Studies

The purpose of this section of the Framework will be to provide some case study examples of how various efforts are modeled and implemented in distributed ways both organizationally and technically. By no means are these case studies intended to be fully representative of the DDP field at large, nor are they meant to convey mutual exclusivity (i.e., some DDP groups may share case study qualities with others). Nonetheless they will document an impressive array of configurations that can lend insight into the various qualities and characteristics of DDP.

The case studies are based on interviews that were focused towards discovering the aspects of DDP and OAIS as utilized by the various organizations and technical initiatives summarized below.

Archivematica Case Study

Community-Driven Support for Distributed Digital Preservation

Technical collaboration is often integral to supporting a DDP implementation. Archivematica's open source and OAIS-based micro-services infrastructure is highly dependent upon a distributed and collaborative community of both users and developers. This case study will highlight best practices for carrying out coordinated technical approaches to accomplishing digital preservation via such modularized and flexible platforms.

Chronopolis Case Study

Balancing Partnerships for Distributed Digital Preservation

The institutions that comprise the Chronopolis program have structured the administrative workflows necessary to streamline the deployment of resources across three very heterogeneous

³ We will follow an approach similar to ALCTS (i.e., short, medium, and long definitions). See here: <http://www.ala.org/alcts/resources/preserv/defdigpres0408>.

organizations. This case study will describe the importance of positioning and coordinating administrative responsibilities effectively across multiple independent organizations that are collaborating toward a shared and distributed repository infrastructure.

Danish Bit Repository Case Study

Shared Flexible Bit Preservation among Institutions for DDP

This case study discusses a platform for shared bit preservation, which allows distribution of copies of data (regardless of whether media is online or offline) as well as services upon them. Shared bit preservation especially challenges roles and responsibilities in connection with how independence among copies of data is maintained, as well as how to ensure that bit preservation solutions for different requirements of bit integrity, confidentiality and availability can be offered.

Data-PASS Case Study

Coordinating Stakeholders for Distributed Digital Preservation

The primary stakeholders in the Data-PASS partnership have balanced institutional independence with the sharing and coordination of resources among institutions participating in collaborative preservation. This case study will explain the importance of cultivating stakeholder buy-in, managing partnership relations, coordination of operations, development of shared practices, and participation in common infrastructure, in order to effectively coordinate actions and ensure sustainability of the overall DDP organizational endeavor.

DuraCloud Case Study

Leveraging Cloud Infrastructure for DDP

This case study will discuss DuraCloud’s innovative approach to managing ingest and storage workflows across multiple commercial and public Cloud storage providers. The case study will also discuss the role and importance of external content auditing and making effective use of service agreements in the Cloud service landscape.

Internet Archive Case Study

Fit-to-Purpose Roles & Responsibilities for DDP

This case study will discuss the Internet Archive’s impressive efforts to seek international partners that can fill focused roles in the overall set of organizational and technical responsibilities for accomplishing DDP for Internet Archive. Much of this key positioning has been for the sake of optimizing the processes necessary for managing and replicating Internet Archive’s large amounts of data.

MetaArchive Case Study

Building Community for Distributed Digital Preservation

Bringing together multiple organizations to accomplish distributed digital preservation is an opportunity to share lessons and develop mutually beneficial technologies. Creating such a community of praxis across multiple organizations can be both challenging and rewarding. This case study of the MetaArchive Cooperative will document the structures and approaches that have made community building possible for this DDP group.

UC3 Merritt Case Study

Dedicated Services for Distributed Digital Preservation:

This case study will explain how the California Digital Library has distributed its UC3 Merritt repository services to optimize the performance of its micro-service oriented architecture and workflows.

3.4.6 Models

The Framework will also investigate different models that can be of help in understanding DDP.

One model being explored can be found in Figure 1, which illustrates the example of distribution for bit preservation [8].

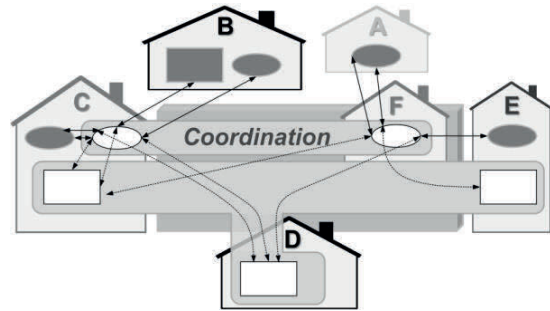


Figure 1. Distribution for Digital Preservation.

The squares could represent a “storage unit” that each hold one copy of data, the houses could represent an organization, and the circles represent services and processes. The dark circles/squares are fully internal, while the white are part of a shared distributed solution. In this case “house” A is only a consumer of a distributed solution, where it does not have any copies of data itself. On the other hand “house” D only has a role as a provider of a “storage unit” to hold a copy of data. Similarly “house F” is only a provider of a coordinated processing service within a distributed solution. This is just to illustrate that DDP can exist in many forms and with varying complexity.

Another example is the IR-BR model shown in Figure 2 [9]. This model was made in connection with a feasibility study for the Danish Bit repository platform.

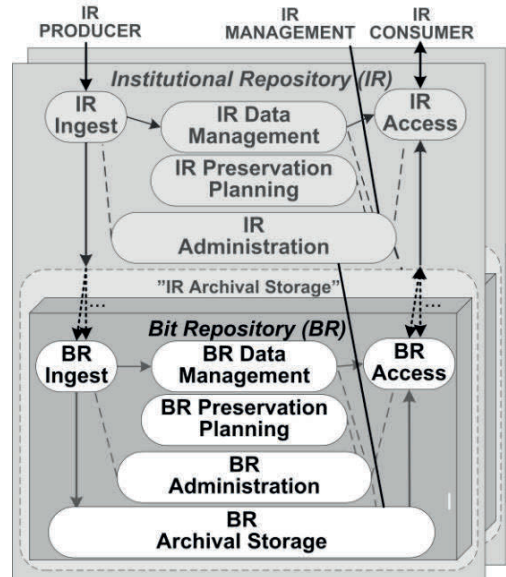


Figure 2. The IR-BR model.

The IR-BR model illustrates a shared Bit Repository (BR) as an OAIS that is shared among a number of Institutional Repositories (IRs) that are also OAIS repositories. Thus the BR is an OAIS within many OAISes.

The idea is that each IR will have their own Technology Watch as part of their *IR Preservation Planning*, which covers ingested file formats. However, the total *Preservation Planning* for the IR and its use of the BR will also cover Technology Watch as part of the *BR Preservation Planning*, thereby covering things like storage media. Furthermore, the BR relies on different “storage units” as depicted in Figure 1. That means the *Preservation Planning* for an

IR covers its own internal *IR Preservation Planning*, as well as the *BR Preservation Planning*, which can be distributed between many organizations with their own technology watches for the media used in their “storage units”.

Similarly, additional models may be added to the Framework, in case they can add value in analysis and/or audit of DDP.

3.4.7 Applying to Auditing Methodologies

Many of the case studies address challenges of applying auditing methodologies to their distributed OAIS. Both DRAMBORA and TRAC are inspired by OAIS, thus it is fair to assume that extension and description of the application of DDP to OAIS can be used in addressing some of these challenges. This will include:

- *Trends & Approaches in Auditing* will provide an overview of two dominant approaches to auditing (top-down with TRAC and bottom-up with DRAMBORA) and will highlight some of the trends and directions with the standardizing of the TRAC audit metrics and the rise of audit agencies.
- *TRAC Metrics - How Do They Apply?* will address how areas of TRAC are best interpreted for DDP environments. This description will also address any potential areas unique to DDP that TRAC currently does not address fully.
- *TRAC Auditors - How to Approach DDP?* will address how auditors can prepare to apply TRAC metrics in a DDP environment.
- *DRAMBORA - How to Approach DDP?* will address how DRAMBORA can be best interpreted and used for DDP environments.

3.4.8 Conclusion – Using the Framework

This section will summarize and provide a set of guidelines and suggestions for applying the Framework for both existing DDP practitioners as well as those interested in using such approaches.

4. DISCUSSION AND FURTHER WORK

Though much research has been carried out and the proposed Framework appears well-defined based on contributions from the Working Group and feedback from the broader stakeholder community, there are still relevant questions and topics to be explored as the effort proceeds to the drafting and further review stages. The Framework initiative aims to continue to expand the discussion on this topic and stakeholder participation through subsequent conference papers, poster sessions, and hosted events.

Among the questions and topics that will need further attention, as mentioned above, are the final form(s) the Framework should take and where it should be hosted for broadest uptake and maintenance over time. As also mentioned above, we acknowledge that analyzing distribution for digital preservation in the context of a repository system and environment should be mildly distinguished from collaboration for things like format registries and other open source technology developments. Drawing the proper boundaries is one that likely deserves more attention and discussion.

Further work is also needed to understand the role and needs of consumers of digital objects and trustworthy requirements for access in a distributed digital preservation environment.

Finally, since this Framework initiative does not aim to occupy the place of a standard, the Framework authors are mindful that

care will need to be taken to make such relationships clear throughout the resource.

5. CONCLUSION

This paper has explained the beginnings of a *Framework for Applying the Reference Model for an Open Archival Information System (OAIS) to Distributed Digital Preservation (DDP)*.

This is an international initiative comprised of numerous DDP practitioners and stakeholders that have been working steadily and collaboratively since 2011 to clearly define the area of needed work, which includes researching and documenting proper mappings between OAIS and DDP in the areas of *terminology*, *functional entities*, and *roles & responsibilities*.

This work has resulted in a white paper that has undergone preliminary review by various DDP and digital preservation experts, a series of case study interviews with DDP practitioners, an analysis of OAIS with respect to DDP, and a thorough literature review.

The Framework proposes to include a number of helpful elements, including a Background & Overview of DDP, discussions on Terminology, DDP and OAIS Relationships, a series of Case Studies & Models, as well as a section on Applying the Framework to Auditing Methodologies.

This Framework will be helpful for future analyses and/or audits of repositories that are performing digital preservation in highly distributed ways.

6. REFERENCES

- [1] Center for Research Libraries, Online Computer Library Center. 2007. Trustworthy Repositories Audit Criteria: Checklist and Certification v1.0. (also ISO Standard 16363).
- [2] Consultative Committee for Space Data Systems. 2009. *Producer-Archive Interface Methodology Abstract Standard*, Magenta Book, May 2004, CCSDS Secretariat.
- [3] Consultative Committee for Space Data Systems. 2002. *Reference Model for an Open Archival Information System (OAIS)*: ISO 14721:2003.
- [4] Digital Curation Centre & Digital Preservation Europe. 2007. *Digital Repository Audit Method Based on Risk Assessment*, Version 1.0.
- [5] Jurik, B. A., Nielsen A. B., Zierau, E. 2012. Flexible Bit Preservation on a National Basis, In *Proceedings of the IS&T Archiving 2012*, Copenhagen, Denmark, pp. 2-7.
- [6] Littman, J. 2006. A Technical Approach and Distributed Model for Validation of Digital Objects, In *D-Lib Magazine*, vol. 12 no. 5.
- [7] Rosenthal, D. S. H. 2010. Bit Preservation: A Solved Problem?, In *The International Journal of Digital Curation*, vol. 5, no. 1.
- [8] Zierau, E. M. O. 2011. *A Holistic Approach to Bit Preservation*, Doctoral Thesis, Copenhagen University
- [9] Zierau, E., Kejser, U.B. 2010. Cross Institutional Cooperation on a Shared Bit Repository, In *Journal of the World Digital Libraries*, vol. 3, no. 1, pp. 11-21.