

**From production to preservation to access to use:  
OAIS, TDR, and the FDLP**  
Federal Depository Library Conference, October 2011  
Presentation Handout  
James A. Jacobs  
Data Services Librarian emeritus, University of California San Diego

**OAIS**

- ORIGINAL (free): Consultative Committee for Space Data Systems. *Reference Model for an Open Archival Information System*. (OAIS) CCSDS 650.0-B-1. "BLUE BOOK" (CCSDS Secretariat, January 2002). [148pp, free]  
<http://public.ccsds.org/publications/archive/650x0b1.pdf>.
- ISO STANDARD (fee): International Organization for Standardization. *Open archival information system -- Reference model* (ISO 14721:2003). [141 pp, 224 Swiss Francs]  
[http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=24683](http://www.iso.org/iso/catalogue_detail.htm?csnumber=24683)
- NEW DRAFT (free): Consultative Committee for Space Data Systems. *Reference Model for an Open Archival Information System*. (OAIS) Draft Recommended Standard CCSDS 650.0-P-1.1 (PINK BOOK) Issue 1.1, August 2009 [131 pp, free]  
<http://public.ccsds.org/sites/cwe/rids/Lists/CCSDS%206500P11/Attachments/650x0p11.pdf>
- NEW DRAFT ISO STANDARD (fee): International Organization for Standardization. *Open archival information system -- Reference model* (ISO/DIS 14721). [141 pp, 98 Swiss Francs]  
[http://www.iso.org/iso/iso\\_catalogue/catalogue\\_ics/catalogue\\_detail\\_ics.htm?csnumber=57284](http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csnumber=57284)

**TRAC / TDR**

- Trusted Digital Repositories: Attributes and Responsibilities*. An RLG-OCLC Report. Mountain View, CA: RLG, May 2002.  
<http://www.oclc.org/programs/ourwork/past/trustedrep/repositories.pdf>
- TRAC: Center for Research Libraries and OCLC. *Trustworthy Repository Audit and Certification: Criteria and Checklist*, Version 1, (2007). [94pp, free]  
[http://www.crl.edu/sites/default/files/attachments/pages/trac\\_0.pdf](http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf)
- TDR: Consultative Committee for Space Data Systems, *Audit And Certification Of Trustworthy Digital Repositories*, "Red Book," Issue 1, DRAFT RECOMMENDED PRACTICE, CCSDS 652.0-R-1 (Washington D.C.: CCSDS, October 2009). [78pp, free]  
<http://public.ccsds.org/sites/cwe/rids/Lists/CCSDS%206520R1/Attachments/652x0r1.pdf>
- DRAFT ISO STANDARD (fee): *Audit and certification of trustworthy digital repositories*. ISO/DIS 16363. [http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=56510](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=56510)

## Other Related Standards

*Data Seal of Approval* (n.d.) Data Seal of Approval: Quality Guidelines for Digital Research Data, online at: <http://www.datasealofapproval.org/>

*DRAMBORA*: The Digital Repository Audit Method Based on Risk Assessment. Digital Curation Centre and DigitalPreservationEurope (DPE). <http://www.repositoryaudit.eu/>

NESTOR. *Catalogue of Criteria for Trusted Digital Repositories*. Network of Expertise in long-term STORage. nestor Working Group on Trusted Repositories Certification:, Version 1 (draft for public comment), June 2006, Frankfurt am Main. [http://files.d-nb.de/nestor/materialien/nestor\\_mat\\_08-eng.pdf](http://files.d-nb.de/nestor/materialien/nestor_mat_08-eng.pdf)

*PAIMAS*: [ISO 20652:2006] *Producer-Archive Interface Methodology Abstract Standard*. CCSDS 651.0-M-1. MAGENTA BOOK. May 2004. <http://public.ccsds.org/publications/archive/651x0m1.pdf>

*PREMIS*: PREMIS Data Dictionary for Preservation Metadata. Version 2.0, PREMIS Editorial Committee, March 2008; <http://www.loc.gov/standards/premis/v2/premis-2-0.pdf>

ISO 9000 family of standards. “An international consensus on good quality management practices.”

ISO 15489-1:2001 and ISO 15489-2:2001. Records management.

ISO 15889:2003 Space data and information transfer systems—Data Description Language—EAST Specification;

ISO 21961:2003 Space data and information transfer systems—Data Entity Dictionary Specification Language (DEDSL)—Abstract syntax;

ISO 23081-1:2006, ISO/TS 23081-2:2009 Metadata for records

ISO/IEC 27002:2005. Code of practice for information security management.

## Other Material

Beedham, Hilary, Julie Missen, Matt Palmer, and Raivo Ruusalepp. *Assessment Of UKDA And TNA Compliance With OAIS And METS Standards*, "Appendix 5: A Set Of Questions For OAIS Compliance Self-Testing." (Wivenhoe Park, Colchester, Essex: UK Data Archive, 2005), p. 111 <http://www.esds.ac.uk/news/publications/oaismets.pdf>.

This is a 3-page list of questions prompted by OAIS. They are generic enough to be relevant to most OAIS archives.

Hernon, Peter 'Information Life Cycle: Its Place in the Management of U.S. Government Information Resources', *Government Information Quarterly*, 11 (1994), 143-170 doi:[10.1016/0740-624X\(94\)90002-7](https://doi.org/10.1016/0740-624X(94)90002-7).

Jacobs. James A., "Privatization of GPO, Defunding of FDsys, and the Future of the FDLP." *FreeGovInfo*. (Aug. 11, 2011) <http://freegovinfo.info/node/3416>

NLM Digital Repository Working Group. *Requirements for an NLM Digital Repository: Report and Recommendations*. (2007). <http://www.nlm.nih.gov/digitalrepository/NLM-DigRep-Report-rev032007.pdf>

Ross, Seamus, and Andrew McHugh, 'The Role of Evidence in Establishing Trust in Repositories', *D-Lib Magazine*, 12 (2006) doi:[10.1045/july2006-ross](https://doi.org/10.1045/july2006-ross).

As you think about what documents and other kinds of evidence you might use to demonstrate OAIS/TRAC compliance, this piece should be of help. The authors "explore the role of evidence within the certification process, and to identify examples of the types of evidence (e.g., documentary, observational, and testimonial) that might be desirable during the course of a repository audit."

Steinhart, Gail, Dianne Dietrich, and Ann Green, 'Establishing Trust in a Chain of Preservation: The TRAC Checklist Applied to a Data Staging Repository (DataStaR)', *D-Lib Magazine*, 15 (2009) <http://www.dlib.org/dlib/september09/steinhart/09steinhart.html>.  
An example of a "staging repository" model.

Strodl, Stephan, Christoph Becker, Robert Neumayer, and Andreas Rauber, "How to Choose a Digital Preservation Strategy," in *Proceedings of the 2007 Conference on Digital Libraries - JCDL '07* (presented at the the 2007 conference, Vancouver, BC, Canada, 2007), p. 29 doi:[10.1145/1255175.1255181](https://doi.org/10.1145/1255175.1255181)

Although this paper focuses on one specific strategy for preservation (the PLANETS Preservation Planning approach), it nevertheless demonstrates how a one organization made decisions on how to implement OAIS to preserve digital objects for a given purpose.

U.S. Government Printing Office. Library Services and Content Management Business Unit and Federal Digital System (FDsys) Program Management Office. *Information Technology Glossary*. ("GPO Glossary v1.0"). (Last updated: 07-03-2008).  
[http://www.fdlp.gov/home/repository/doc\\_download/530-information-technology-glossary](http://www.fdlp.gov/home/repository/doc_download/530-information-technology-glossary)

U.S. Superintendent of Documents. *SOD 301, Dissemination/Distribution Policy for the Federal Depository Library Program*. Sept. 28, 2006.  
[http://www.fdlp.gov/home/repository/doc\\_download/567-disseminationdistribution-policy-for-the-federal-depository-library-program-sod-301](http://www.fdlp.gov/home/repository/doc_download/567-disseminationdistribution-policy-for-the-federal-depository-library-program-sod-301)

## OAIS Conformance

### 1.4 CONFORMANCE

A conforming OAIS archive implementation shall support the model of information described in 2.2. The OAIS Reference Model does not define or require any particular method of implementation of these concepts.

A conforming OAIS archive shall fulfill the responsibilities listed in 3.1.

### 3.1 MANDATORY RESPONSIBILITIES

This subsection establishes mandatory responsibilities that an organization must discharge in order to operate an OAIS archive.

The OAIS must:

- Negotiate for and accept appropriate information from information Producers.
- Obtain sufficient control of the information provided to the level needed to ensure Long-Term Preservation.
- Determine, either by itself or in conjunction with other parties, which communities should become the Designated Community and, therefore, should be able to understand the information provided.
- Ensure that the information to be preserved is **Independently Understandable** to the Designated Community. In other words, the community should be able to understand the information without needing the assistance of the experts who produced the information.
- Follow documented policies and procedures which ensure that the information is preserved against all reasonable contingencies, and which enable the information to be disseminated as authenticated copies of the original, or as traceable to the original.
- Make the preserved information available to the Designated Community.

## Sample TDR Metrics

### 3.4 FINANCIAL SUSTAINABILITY

**3.4.1 The repository shall have short- and long-term business planning processes in place to sustain the repository over time.**

#### **Supporting Text**

This is necessary in order to ensure the viability of the repository over the period of time it has promised to provide access to its contents for its Designated Community.

#### **Examples of Ways the Repository Can Demonstrate It Is Meeting This Requirement**

Up-to-date, multi-year strategic, operating and/or business plans; audited annual financial statements; financial forecasts with multiple budget scenarios; contingency plans; market analysis.

#### **Discussion**

An annual business planning process is commonly accepted as the standard for most organizations.

**3.1.2.1 The repository shall have an appropriate, formal succession plan, contingency plans, and/or escrow arrangements in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope.**

#### **Supporting Text**

This is necessary in order to preserve the information content entrusted to the repository by handing it on to another custodian in the case that the repository ceases to operate.

#### **Examples of Ways the Repository Can Demonstrate It Is Meeting This Requirement**

Written and credible succession and contingency plan(s); explicit and specific statement documenting the intent to ensure continuity of the repository, and the steps taken and to be taken to ensure continuity; escrow of critical code, software, and metadata sufficient to enable reconstitution of the repository and its content in the event of repository failure; escrow and/or reserve funds set aside for contingencies; explicit agreements with successor organizations documenting the measures to be taken to ensure the complete and formal transfer of responsibility for the repository's digital content and related assets, and granting the requisite rights necessary to ensure continuity of the content and repository services.

#### **Discussion**

A repository's failure threatens the long-term sustainability of a repository's information content. It is not sufficient for the repository to have an informal plan or policy regarding where its data goes should a failure occur. A formal plan with identified procedures needs to be in place.

From production to preservation  
to access to use:  
OAIS, TDR, and the FDLP

Depository Library Council Meeting  
Federal Depository Library Conference  
October 2011

James A. Jacobs  
Data Services Librarian Emeritus  
University of California San Diego

---

---

---

---

---

---

---

---

Slides, handouts, links, bibliography:  
[freegovinfo.info/oais\\_tdr](http://freegovinfo.info/oais_tdr)

---

---

---

---

---

---

---

---

- <goals>
- A non-technical introduction
  - An overview, key points, some detail, a few examples
  - Share my experiences and impressions
  - Relevance: to you, to your library, to FDLP
  - A (new/old) context for addressing everyday questions

The concepts I'll present today should be useful to you regardless of your job title or responsibilities. You don't have to be a technician or programmer or digital preservation specialist to make use of these concepts.

---

---

---

---

---

---

---

---



# Part 1

OAIS, TDR, Information Lifecycle:  
What are they?

---

---

---

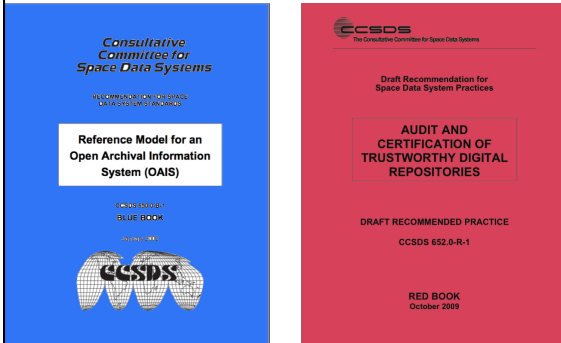
---

---

---

---

<background>



---

---

---

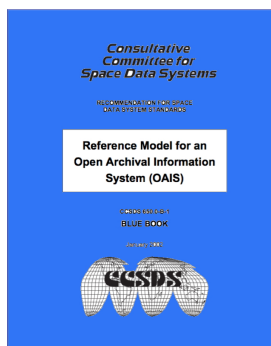
---

---

---

---

# OAIS



---

---

---

---

---

---

---

### OAIS

1. It defines the **functional concepts** of a long-term archive and gives consistent, unambiguous **terminology**.
2. It gives us a functional framework for **designing** archives.
3. It gives us a standard for "**conformance**."

i find it best to think about oais as having three different uses.

---



---



---



---



---



---

### OAIS

(1) OAIS **terminology** helps us **discuss** and examine the issues across domains.

allows talk lib - archv - programmers - db admins, producers and consumers

---



---



---



---



---

### OAIS

(2) OAIS focus on Functionality helps us **design** archives that work.

OAIS is a "**Reference Model**" not an "implementation"

- It **does not prescribe**.
- It does give you the terminology to **describe** what you do and how you do it.
- It describes **functions**.

, it does not tell you to build a data center, or what OS to use, or what file formats to use, or what metadata standards to use, etc. We'll see more about what a "reference model" is in a minute.

---



---



## OAIS

(2) OAIS focus on Functionality helps us **design** archives that work.

OAIS prompts us to ask **questions**.

- **Who** are we doing this for?
- **What** information will we preserve?
- **How** will we ensure preservation and understandability?

---

---

---

---

---

---

---

---

## OAIS

(2) The OAIS focus on Functionality helps us **design** archives that work.

OAIS is **generalizable**.

Since it is not prescriptive, OAIS can be used by all kinds of libraries and archives.

It is equally applicable to libraries and archives, to large and small institutions. It doesn't matter if you are in a public, academic, special, or school library. it is not even about digital-only archives!

---

---

## OAIS

(3) OAIS gives us a Standard against which we can measure **Conformance**

This is where TDR comes in

---

---

---

---

---

---

---

---



TDR is based on OAIS.  
 It translates the general principles and functions and general requirements to a measurable checklist that can be used for certification.

---

---

---

**TDR**

Provides 109 “**metrics**” for measuring conformance to OAIS.

Metrics cover 3 broad areas:

- Organizational Infrastructure
- Digital Object Management
- Technical Infrastructure And Security Risk Management

The 3 broad areas correspond roughly to OAIS categories

---

---

---

---

---

**TDR**

The Metrics are **not prescriptive**.

Asks the archive to document:

- How it answered questions prompted by OAIS.
- Is it doing what it planned to do?
- Is it working?

ideally, many repositories of different kinds can all be measured with same metric

---

---

---

---

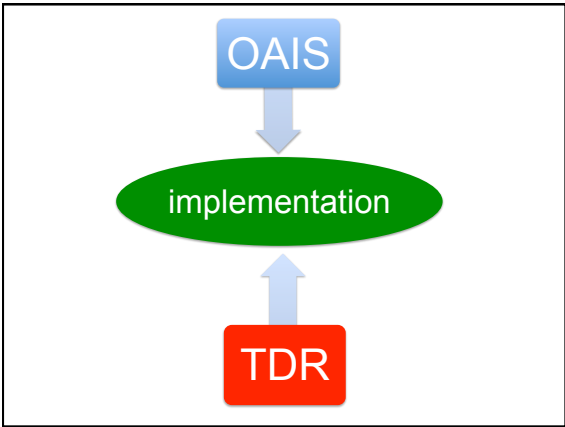
---

### TDR

Gives an archive a shorthand way of communicating its **trustworthiness** to Producers and Users

Again we see this is about communication. A stamp of approval or a “certification” based on common standards provides producers and consumers a way of understanding what an archive is doing and that it is doing it well. Since OAIS is applicable to different kinds and sizes of archives, TDR provides a way of consistently documenting what each archive is doing as well as how well it is doing what it does.

---



OAIS gives you the macro, “high-level” “top-down” contextual overview and TDR provides a micro, “bottom-up” way to evaluate the specifics of your implementation. Together they give a context for developing, running, and evaluating an archive, but they do NOT prescribe how you will implement your archive.

---

---

---

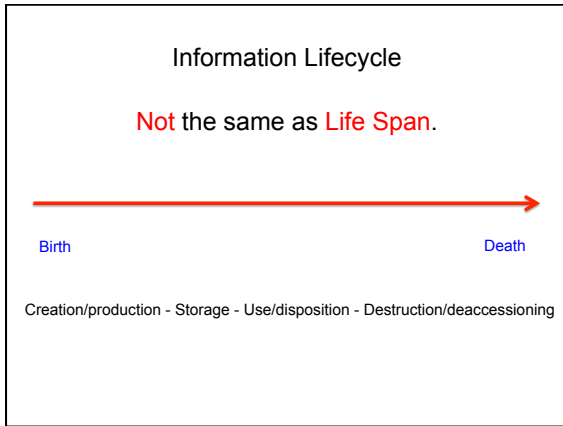
---

---

---

---

Information Lifecycle



Since consistent terminology is a theme today and, since the term “information Lifecycle” is used very inconsistently, I want to define how I use the term.

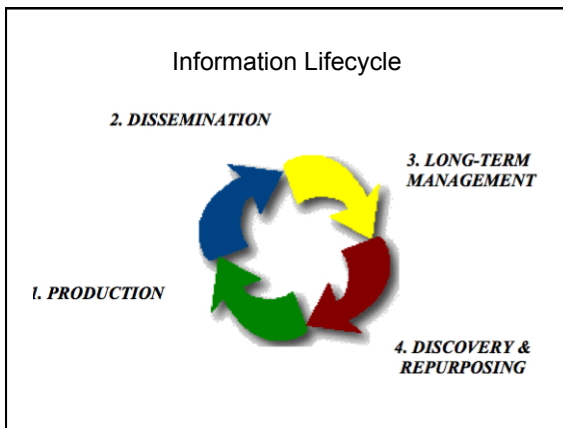
First, it is often used to mean the same thing as the Life Span of information.

Life Span has two characteristics:

- 1) it is linear and
- 2) as a “birth to death” timeline it explicitly assumes “death” or withdrawal or destruction etc.

Different actors in the life cycle think of the “final step” in the lifespan (death) differently.

Death may = distribution, dissemination, destruction, etc.



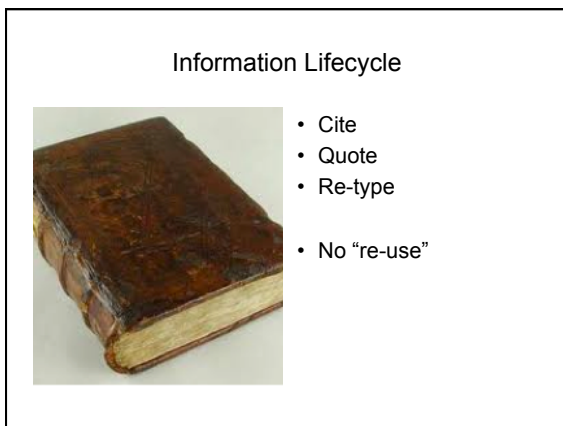
The Information Lifecycle is not linear.

It is a model that assumes information use, re-use, and repurposing.

Where the lifespan model provides each actor with a limited, linear perspective of its activities, the life-cycle model provides an overview from the perspective of the information itself.

Different actors may participate in different parts of the life cycle,

but the model gives us a way of thinking about the permanent use of information.




---

---

---

---

---

---

---

---

Information Lifecycle

```

100110011100100000
0110010101110010
000011000010111000
010011000010111000
011011100110011100
110010 0111011101
0100000011100000110
000001101001011011
0110000101110000011
100110011100100000
0110010101110010

```

- Copy
- Subset
- Excerpt
- Re-mix
- Mashup
- Reformat
- Literally re-use

---

---

---

---

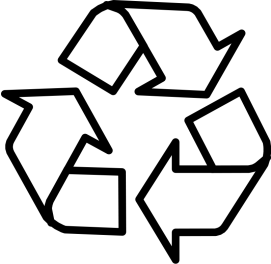
---

---

---

---

Information Lifecycle



So when we think about the information lifecycle of digital information, we can think of it as recycling  
 And, we need to track not just citations, but objects, provenance, etc.

In fact, 'authenticity' takes on new meaning in this environment. It is no longer the just the object that conveys authenticity, but the "digital object" plus its provenance and metadata and the trust we have in its source of the digital object.

Part 2

OAIS, TDR:  
Some details

---

---

---

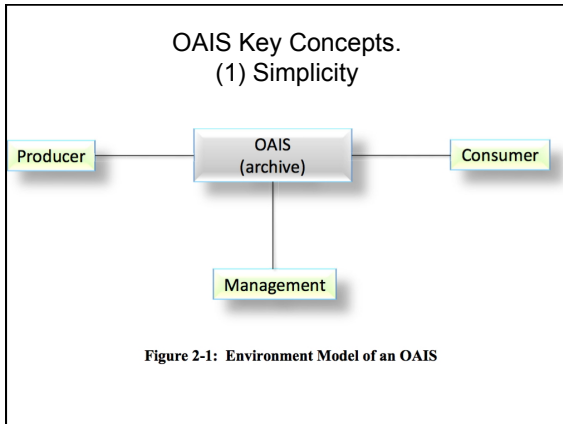
---

---

---

---

---



It highlights ROLES in the information lifecycle and describes the roles or functions of the archive in detail.

It notes the need for management and preservation planning.

OAIS does have a lot of detail and lots of terminology and diagrams and “richness.” but, at heart, it describes very basic, simple, easily understood concepts.

This what makes it applicable to many different archives.

**OAIS Key Concepts.**  
(2) The Designated Community

An **identified group** of potential Consumers who should be able to **understand** a particular set of information.

The Designated Community may be composed of multiple user communities. [1.7]

The Concept of the “Designated Community” is fundamental to OAIS. It is mentioned over 75 times and is in every section of the book.

OAIS is all about your choosing your designated community.

One of the main questions you have to answer when you design an archive is to ask “for whom am I doing this?” And that is critical so that you will know what to collect and how to make it understandable to your community.

A Designated Community can be very small and specialized or very large and general: it is up to you.

And an archive can have more than one Designated Community.

**OAIS Key Concepts.**  
(3) The Long Term

“Indefinitely”  
“Permanent”

Long Term: A period of time long enough for there to be concern about the impacts of **changing technologies**, including support for new **media** and **data formats**, and of a **changing user community**, on the information being held in a repository.

This period extends into the indefinite future. [1.7]

---

---

---

---

---

---

---

---

OAIS Key Concepts.  
(3) The Long Term

Permanence of the **Information**  
*...not the institution.*

“The information being maintained has been deemed to need **Long Term Preservation**, even if the OAIS itself is not permanent.” [1.1]

One result of this is that an OAIS archive can be designed to be temporary.

There are already examples of “staging repositories” whose designated community is another archive that will take over preservation at some later time.

---

---

---

OAIS Key Concepts.  
(3) The Long Term

Not just ‘*bit* storage’ -- but long-term **information preservation**, **and access**, **and understandability**. [2]

---

---

---

---

---

---

---

OAIS Key Concepts.  
(3) The Long Term

For years, preservation simply meant collecting. The sheer act of pulling a collection of manuscripts from a barn, a basement, or a parking garage and placing it intact in a dry building with locks on the door fulfilled the fundamental preservation mandate of the institution.

In this regard, preservation and access have been mutually exclusive activities often in constant tension.

*In the digital world, the concept of access is transformed from a convenient byproduct of the preservation process to its central motif.*

The content, structure, and integrity of the information object assume center stage; the ability of a machine to transport and display this information object becomes an assumed end result of preservation action rather than its primary goal.

•Paul Conway, "Preservation in the Digital World" (1996)  
<http://www.clir.org/pubs/reports/conway2/>

---

---

---

---

---

---

---

### OAIS Key Concepts. (3) The Long Term

- Information must be:
  - Not just preserved, but *discoverable*. [2.2.2]
  - Not just discoverable, but *deliverable*. [2.3.3]
  - Not just deliverable as bits, but *readable*. [2.2.1]
  - Not just readable, but *understandable*. [2.2.1]
  - Not just understandable, but *usable*. [4.1.1.5]

---

---

---

---

---

---

---

---

### OAIS Key Concepts. (3) The Long Term

How do you know if your information is usable?

- The archive must ask **questions** and make **decisions**.
- The archive must design an **implementation** to address the needs of its **Designated Community**.

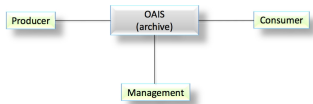


Figure 2-1: Environment Model of an OAIS

The key here is that the archive-function is different from the producer or consumer role.

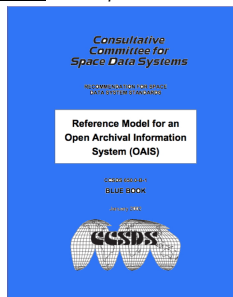
The archive-role is one of addressing the functional needs of the information and the community.

It is NOT a passive role.

It is not a role of accepting what the producer gives without question.

### OAIS Key Concepts. (4) Reference Model

*Reference Model for an Open Archival Information System*




---

---

---

---

---

---

---

---



OAIS Key Concepts.  
(4) Reference Model

What a Reference Model is *Not*:

- Not an implementation.
- Not a system or software or hardware.
- Not about file formats.
- Not about particular metadata standards.

It does not tell you that you must use METS or MODS or MARC

It does not tell you that you have to use XML or that you cannot use proprietary formats.

It does not tell you to use Oracle and avoid MySQL

It is not about How? but What? What functionality.

OAIS Key Concepts.  
(4) Reference Model

Ref. Model for a "Land Vehicle"

- Propulsion
- Steering
- Breaking

a box of parts (steering wheel, engine, breaks) would not conform.

The questions you ask, when confronting the reference model, are things like  
are we building a vehicle for freight or passengers?  
What Capacity do we need?  
What about Mileage, Total Cost of Ownership, Maximum Speed, Acceleration, etc. ?

The answers you come up with determine your implementation.

OAIS Key Concepts.  
(4) Reference Model

Ref. Model for a "Land Vehicle"

Implementations:

- Propulsion
- Steering
- Breaking



All of these, different as they are from each other, can conform to the reference model.

Conforming to OAIS

1. Conform to the **OAIS Information Model**
2. Fulfill **6 OAIS Responsibilities**

So, if the reference model helps us to \*design\* an implementation, how do we know if the implementation \*conforms\* to the model?

OAIS specifies two requirements. I'll go over these quickly.

---



---

Conforming to OAIS

**OAIS Information Model**

Content + Metadata

---



---



---



---



---



---

Conforming to OAIS

OAIS Information Model [2.2]

**Information Packages** [2.2.2]

- Content, plus
  - Representation Information (**understandability**)
  - Descriptive Information (**discovery, identification**)
  - Preservation Description Information (**preservation**)
  - Packaging Information (**management**)

It describes in some detail these concepts (again, w/o specifying any file formats or database schemas or how you will or should record information.)

And it does not specify WHAT information to preserve.

It does specify the FUNCTIONAL needs of the information you preserve.

---



---

Conforming to OAIS

The Three Information Package *Variants* [2.2.3]

- Submission Information Package (SIP)
- Archival Information Package (AIP)
- Dissemination Information Package (DIP)

---

---

---

---

---

---

---

Conforming to OAIS

SIP → AIP → DIP

First, note that OAIS specifies FUNCTION of these different packages.

This is not about format or media or file-type.

---

---

---

---

Conforming to OAIS

- Information must be:
  - Not just preserved, but *discoverable*. [2.2.2]
  - Not just discoverable, but *deliverable*. [2.3.3]
  - Not just deliverable as bits, but *readable*. [2.2.1]
  - Not just readable, but *understandable*. [2.2.1]
  - Not just understandable, but *usable*. [4.1.1.5]

The OAIS Information Model describes and gives us consistent terminology for how we “package” information so that it meets all these functional requirements.


A package suitable for submission, may not be suitable for preservation,

and a package suitable for preservation may not be suitable for delivery to users.

---

---

Conforming to OAIS



While this may seem obvious, it is actually a little bit revolutionary, or at least evolutionary.

For, it was not too long ago that we thought in terms of acquiring, preserving, and delivering information in the same package.


Note that this is entirely acceptable to the OAIS model.

---



---

Conforming to OAIS



These Packages must be Functional:

- Is this SIP preservable, or can I make it into a preservable AIP?
- Is this AIP deliverable and usable, or can I make it into a deliverable, usable DIP?

OAIS doesn't specify how we accept, or store, or deliver the packages. It only specifies that they work: that they fulfill their functions.

---

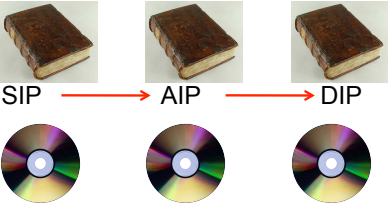


---



---

Conforming to OAIS

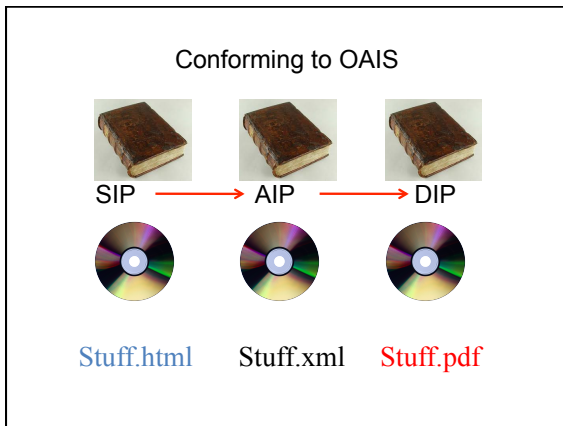


As another example, think of portable digital media like CD-ROMs and DVDs.

OAIS does not tell us these are good or bad, but only prompts us to ask if they work.

OAIS implies that it is not acceptable for us to ASSUME that the portability of a Medium makes the Content "portable" or preservable or deliverable.

We have to be sure when we accept a SIP that we can preserve it or convert to a preservable format; and, when we design a delivery mechanism, that the DIP will be usable and understandable by the Designated Community.



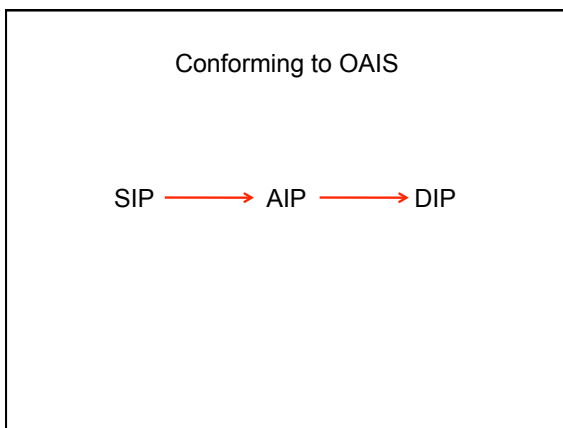
And....  
 OAIS does not prescribe any methods or tools or formats for these packages.

We might choose in the case of book for all 3 packages to be the same

We might also choose to convert html to xml for preservation and convert it to pdf for delivery.

We might choose any number of possible combinations of strategies.

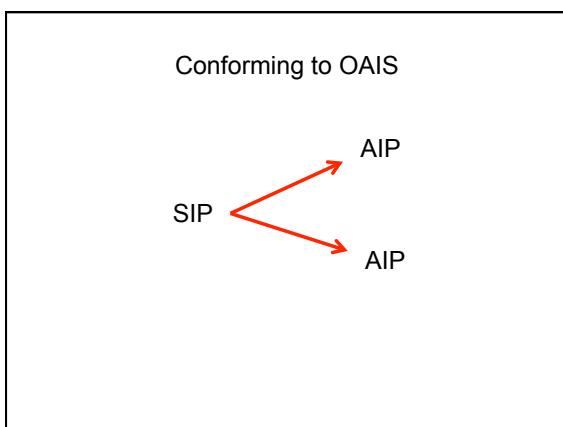
OAIS only asks us to ensure that they work.



Let's quickly visit a big advantage that these 3 separate conceptual packages produce.

Here is one way we could deal with content:

We get content in a package of information as a SIP and we store that content as an AIP and we deliver that content as a DIP.

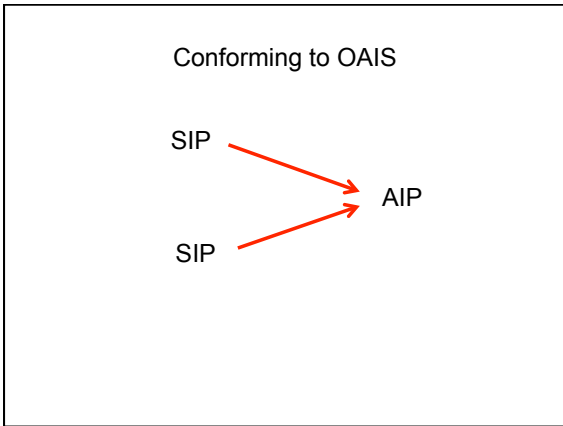


But that is not our only option.  
 By decoupling submission, archiving, and delivery, we have other options.

For example, one sip might be divided into multiple AIPs.

You might, for instance, get a SIP that has an issue of a journal w/ 10 articles and decide that for archival purposes you want each AIP to be one article. So you split the SIP into 10 AIPs.

That's OK!



multiple SIPs, from a publisher,  
each containing a volume of a multi-volume book

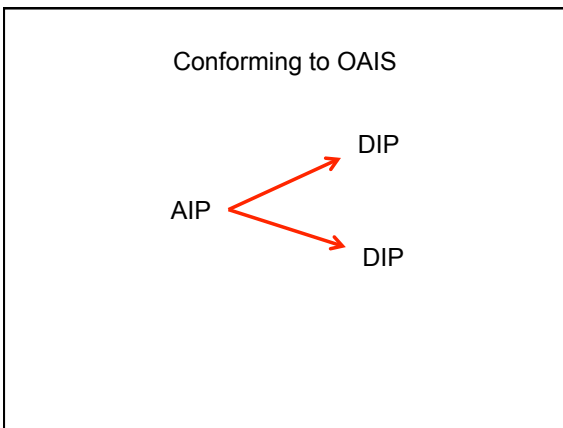
you might decide to combine the volumes into an AIP  
that has the complete book with all of its volumes.

That's OK!

---



---



Similarly, you might want to deliver part of an AIP.

Sn AIP might have text, and a movie, and still images  
and raw data

and you might want to give the user the opportunity  
to get just the text, or just the data, etc.

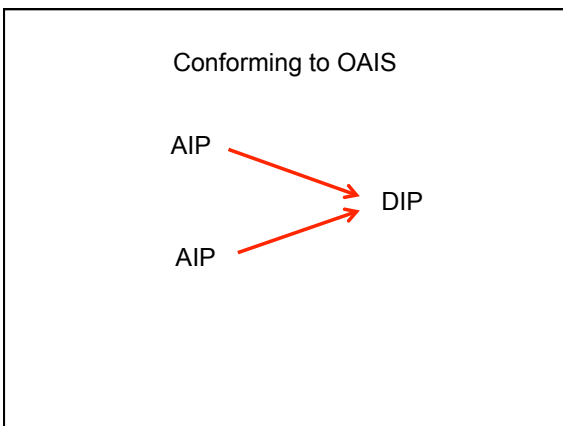
---



---



---



A user might search your archive and find several documents  
and you might give the user to opportunity to get all those  
documents in a single DIP,  
saving the user from repetitively clicking and downloading  
the individual documents.

THE POINT IS THE INFORMATION MODEL ALLOWS US  
to decouple what the producer chooses to give us from  
how we store and preserve that information.

AND we can decouple how we preserve information from  
how we deliver it.

Each package serves its own function,

Conforming to OAIS

1. Conform to the **OAIS Information Model**
2. Fulfill **6 OAIS Responsibilities**

---



---



---



---

Conforming to OAIS

Mandatory responsibilities [3.1]

1. Negotiates For And Accepts Information (**Select**)
2. Obtains Sufficient Control For Preservation (**Acquire**)
3. Determines **Designated Community**
4. Ensures Information Is **Independently Understandable**
5. Follows Established Preservation Policies And Procedures (**Organize and Preserve**)
6. Makes The Information Available (**Provide access**)

I suggest to you that these mandatory responsibilities are the essential, traditional, continuing roles of a library.

These roles are an essential part of the information life cycle. Without these, information will be lost and the life cycle will be broken.

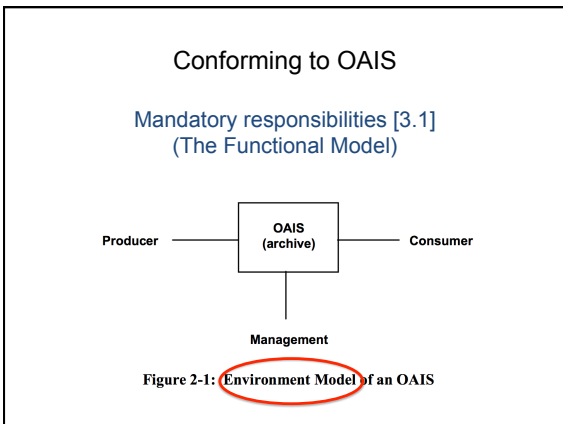
Item 3 deserves a special mention. In the physical world, it was easy for libraries to think of their designated communities as geographically-based. Physical proximity was often THE defining characteristic of the “community”

But, in the digital world, the designated community need not be physically near the library. This gives libraries a new flexibility in addressing the needs of new communities.

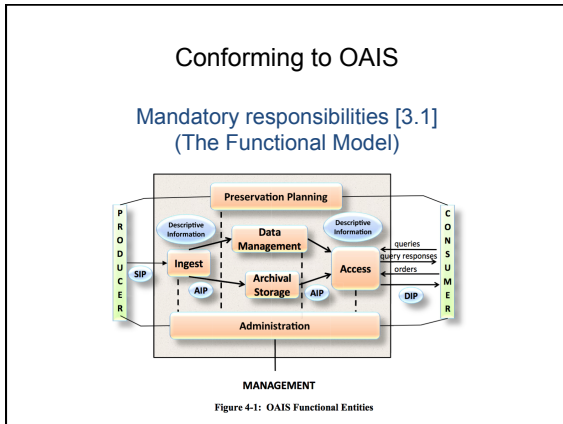
Communities can be based on subjects, or disciplines, or type of information, or type of use of information, or almost any other focus.

Also, item 4 -- "ensuring that the information is Independently Understandable" is an essential one in the digital world, because the “bits” are not understandable. Remember the information must be readable, understandable, and usable.

These two “new” requirements broaden and focus the traditional functions of a library.



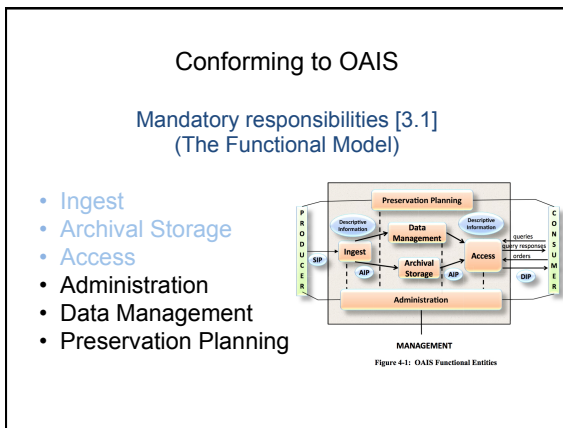
To help archives design system that adequately meet the requirements, OAIS provides a Functional Model to complement its Information Model.



The functional model.  
You can see the producer and consumer to the left and right and the archive with all its functions in the middle.

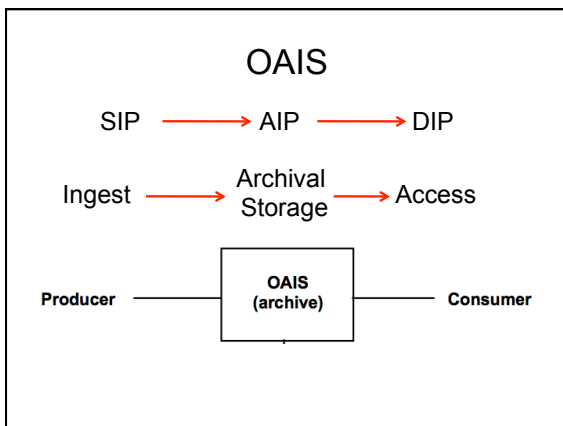
The 6 functions are in orange,  
and the SIP, AIP, and DIP in blue.

This also shows the connections between the modules.



The Functional Model lists  
3 administrative/management/planning functions  
in addition to the functions of ingest, storage, and access  
that we've already seen in the information model.

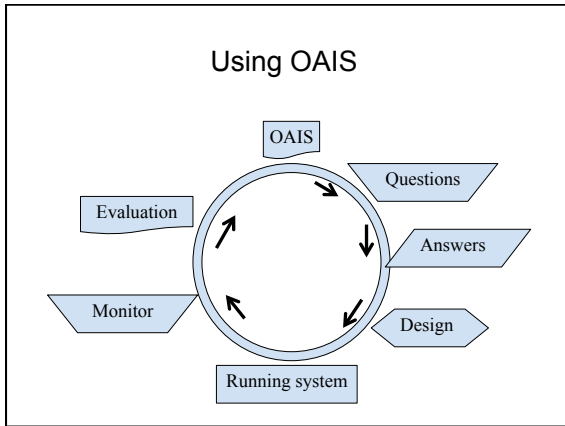
OAIS describes the functional model in some detail,  
but ....



But the key concept here is, again, the simplicity  
and consistency of OAIS.

The information model,  
the functional model,  
and the environment  
are all reflected in these three stages  
or processes  
or functions.



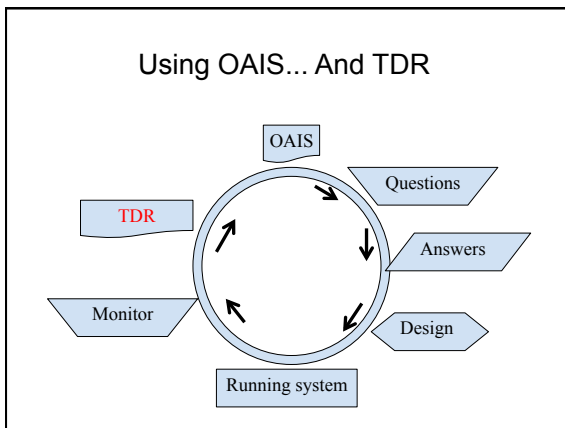


So, Now we can imagine how one might use OAIS to design a repository!

One starts with OAIS functional categories, asks questions to give those concepts specific answers in relation to your particular library;

then you design the systems that will fulfill your own requirements, build the system and monitor it to be sure it works.

Then you can evaluate everything you've done to be sure you are fulfilling your mission.



---

---

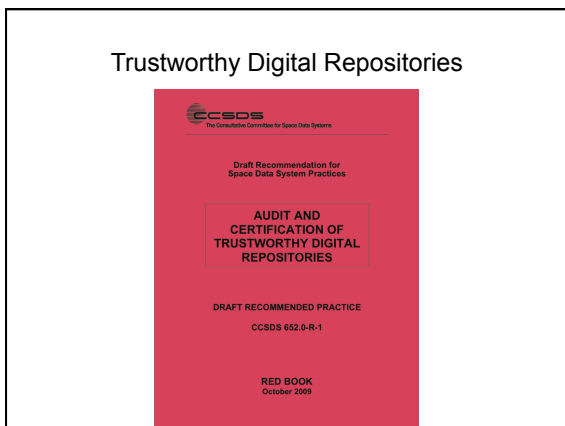
---

---

---

---

---



---

---

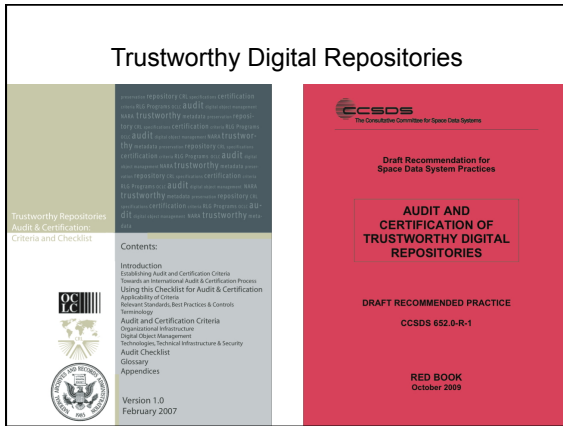
---

---

---

---

---



You may have also heard of “TRAC.”

The “Trustworthy Repository Audit and Certification: Criteria and Checklist”

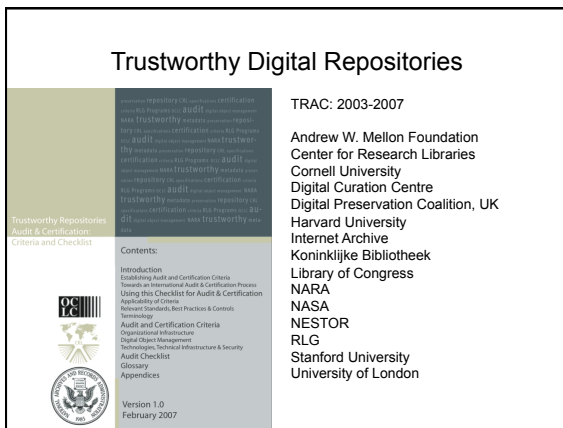
from 2007.

It was the predecessor to TDR (2009).

---



---



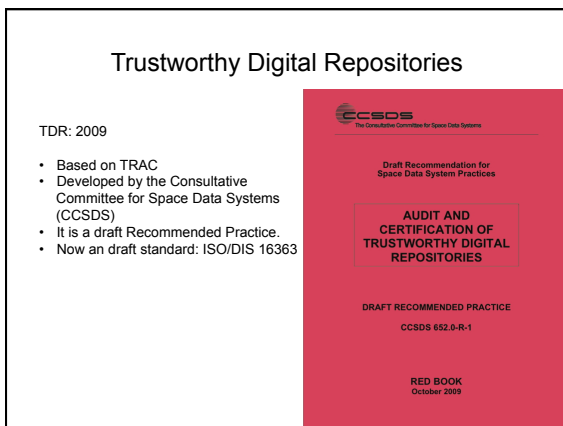
TRAC actually dates back to 2000 when RLG and OCLC began a collaboration to establish attributes of a digital repository for research organizations, based on OAIS.

TRAC itself began development in 2003 when RLG and the NARA created a joint task force to specifically address digital repository certification.

---



---



TDR is still considered a DRAFT, but it has already been through a test phase and will likely replace TRAC officially very soon.

---



---



---



---

**TDR**

Provides 109 “**metrics**” for measuring conformance to OAIS.

---

---

---

---

---

---

---

**TDR Metrics**

- Metric requirement
- Supporting Text
- Examples of Ways the Repository Can Demonstrate It Is Meeting This Requirement
- Discussion

---

---

---

---

---

---

---

**3.3.5 The repository shall define, collect, track, and appropriately provide its information **integrity measurements**.**

**Supporting Text**  
 This is necessary in order to provide documentation that it has developed or adapted appropriate measures for ensuring the integrity of its holding.

**Examples of Ways the Repository Can Demonstrate It Is Meeting This Requirement**  
 Written **definition or specification** of the repository’s integrity measures (for example, computed checksum or hash value); **documentation of the procedures** and mechanisms for monitoring integrity measurements and for responding to results of integrity measurements that indicate digital content is at risk; an **audit process** for collecting, tracking and presenting integrity measurements; **Preservation Policy** and **workflow documentation**.

**Discussion**  
 The mechanisms to measure integrity will evolve as technology evolves. The repository may provide documentation that it has developed or adapted appropriate measures for ensuring the integrity of its holdings. **If protocols, rules and mechanisms are embedded in the repository software, there should be some way to demonstrate the implementation of integrity measures.**

The things to understand about this example are:

Not prescriptive  
 - (doesn’t say you have to use a particular measurement of integrity [e.g., MD5 hash])

- It suggests ways you might document that you are meeting the requirement, but does not require any single, specific method of documenting

- It requires what you are doing (the function of ensuring integrity) works. [track, audit, monitor]

**TDR**

Provides 109 “**metrics**” for measuring conformance to OAIS.

Metrics cover 3 broad areas:

- Organizational Infrastructure
- Digital Object Management
- Technical Infrastructure And Security Risk Management

---

---

---

---

---

---

---

---

**TDR**

Two Metrics from “Organizational Infrastructure”

- Governance & Organizational Viability
- Organizational Structure & Staffing
- Procedural Accountability & Preservation Policy Framework
- Financial Sustainability
- Contracts, Licenses, & Liabilities

Since this is a non-technical presentation, I want to highlight 2 metrics from the largely non-technical category of “organizational infrastructure.”

This category covers the areas you see here: (governance, staffing, policies, finances, contracts, etc.)

This category includes 25 individual metrics. I’ll describe 2 of those 25.

---

**TDR**

**Sustainability**

**3.4.1 The repository shall have short- and long-term business planning processes in place to sustain the repository over time.**

This is necessary in order to ensure the viability of the repository over the period of time it has promised to provide access to its contents for its Designated Community.

Now, if you remember, OAIS is about the permanence of the Information, not of the Institution.

Nevertheless, TDR wants the institution to be sufficiently stable that it won’t disappear from lack of attention to its own sustainability.

---

---

# TDR

## Succession Plan

3.1.2.1 The repository shall have an appropriate, formal succession plan, contingency plans, and/or escrow arrangements in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope.

This is necessary in order to preserve the information content entrusted to the repository by handing it on to another custodian in the case that the repository ceases to operate.

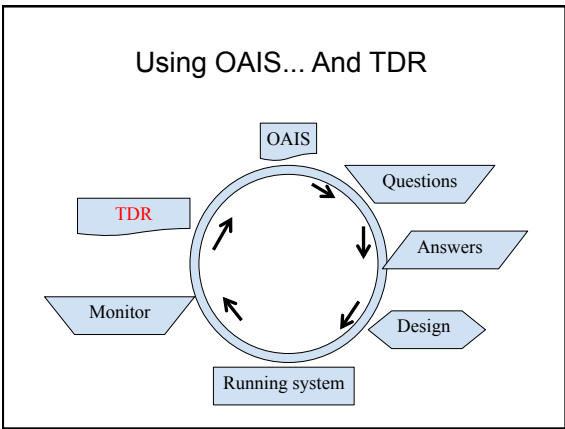
No organization can guarantee it will survive forever, and OAIS doesn't insist that an archive guarantee that.

TDR, therefore, requires that an archive have a plan in place in case the archive becomes unsustainable.

I'll come back to these in a moment, but for now, let's move to how a library might actually use OAIS and TDR...

---

## Using OAIS... And TDR



Now, if we take our earlier diagram of using OAIS and TDR, we can add ....

---

---

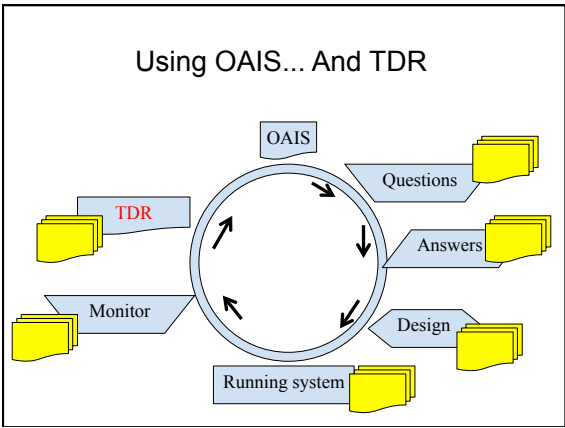
---

---

---

---

## Using OAIS... And TDR



... the documentation that the archive can generate to facilitate a certification audit that uses TDR metrics.

---

---

---

---

---

---

### Documenting for TDR

- Policy Documents
- Implementation documentation
- Evidence of accomplishment
- Demonstrations

What kind of documentation does a library need to have to be certified by TDR?

TDR gives “examples” but does not prescribe.

From my experience, I believe it is useful to conceptualize the kind documentation an archive produces (or can produce) to facilitate the certification process in these 4 categories.

---

---

---

---

---

---

---

---

### Part 3

What does all this mean to me, my library, and FDLP?

### For Every Librarian...

OAIS gives us a consistent terminology that we can use to talk across “domains.”

The more you know, the better you can communicate with:

- technologists
- managers
- producers
- users
- other libraries

OAIS helps you bridge the gap between technologists and libraries and between libraries and archives and between management and front-line librarians.

---

---

---

### For you and your library

- Every library decision should assess the impact of digital issues.
- The more you know, the better you can participate in decision making.
- With a strong grasp of OAIS functions and terminology, you can become a *leader* in library decision making.

---



---



---



---

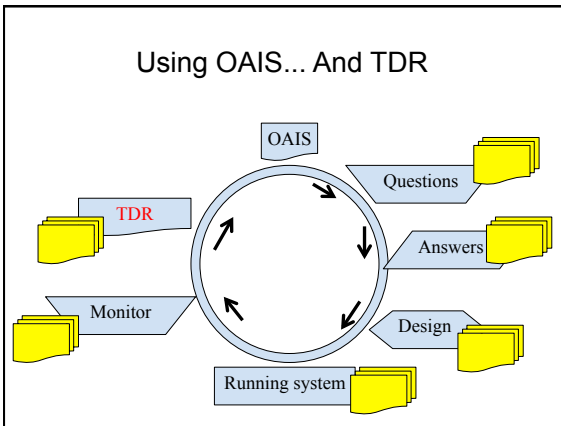


---



---

### Using OAIS... And TDR



Let me illustrate this by returning to this diagram. Most repositories don't look like this.

---



---



---

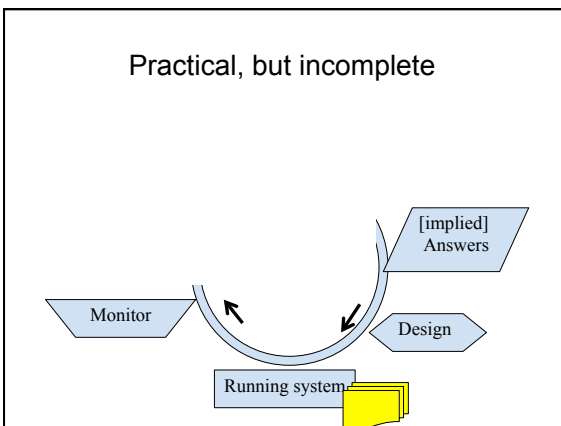


---



---

### Practical, but incomplete



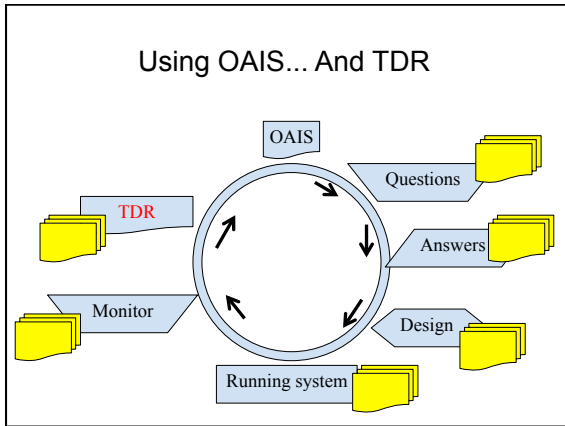
They look a lot more like this. Implied or assumed answers, maybe some documentation, maybe not... Very practically oriented, but very incomplete. This is bad...

Those who are knowledgeable of OAIS and TDR can turn this....

---



---



Into this.

And this is true even if you don't have a "Running system" or a "data center."

This can help you make good policy decisions whenever there is digital content involved (which is almost always...)

And that brings us to...

For FDLP

We can better understand what it means to design an FDLP that guarantees long-term preservation and access.

- Functional and Information Models
- Understandability for the Long-term
- The role of the "archive" and how it differs from the role of the Producer.

...how this can help as we discuss the future of the FDLP.

We can understand key concepts that will help us ensure good policies for the system as a whole, not just for our individual libraries.

---



---



---



---

Examples

Three broken models

I want to show you three models that are real and in use today and that would be different if they had been designed with an understanding of OAIS.

I use the term "model" here very loosely -- as you will see.

and i use the term "broken" to be provocative and mean "does not conform to OAIS."

---



Broken Model 1: Medium not Message



One thing that we learn from OAIS is that information packages must be functional.

The bits we deal with in a digital world have to be preservable, readable, understandable, usable, and so forth. We should be focused on the "message" not the "medium"

---

---

---

Broken Model 1: Medium not Message



These Packages must be Functional:

- Is this SIP preservable, or can I make it into a preservable AIP?
- Is this AIP deliverable and usable, or can I make it into a deliverable, usable DIP?

What matters in the OAIS information model is the 3 information packages and their ability to fulfill their respective functions.

The red arrows in between them indicate a transfer of bits from one package to the next.

The medium of that transfer does not matter except as it affect the functionality of the resulting package.

The physical medium that we use to transfer bits from SIP to AIP to DIP is just that: a medium for transportation of bits.

Broken Model 1: Medium not Message



U.S. Superintendent of Documents. *SOD 301, Dissemination/Distribution Policy for the Federal Depository Library Program*. Sept. 28, 2006.

Determines  
"the format of materials  
disseminated/distributed  
to Federal depository libraries."

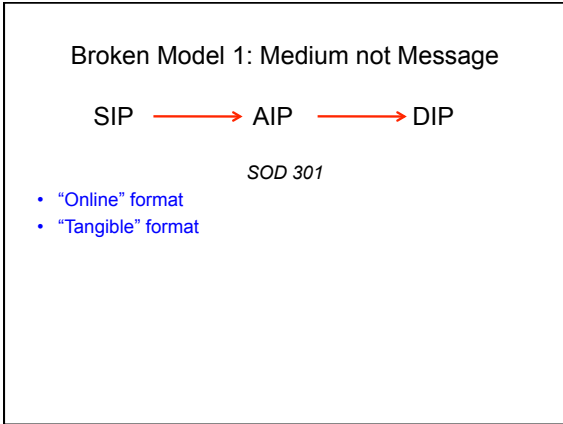
Which brings us to this superintendent of document policy.

This policy determines what it calls the "format" of materials for what it calls "dissemination/distribution" to Federal depository libraries.

---

---

---



Although S.O.D 301 has some useful criteria and considerations (including "functionality" and "permanent public access")...

...it filters all those criteria down to two alternatives: "online" or "tangible"

These alternatives tell you all you need to know about their conformance to OAIS:

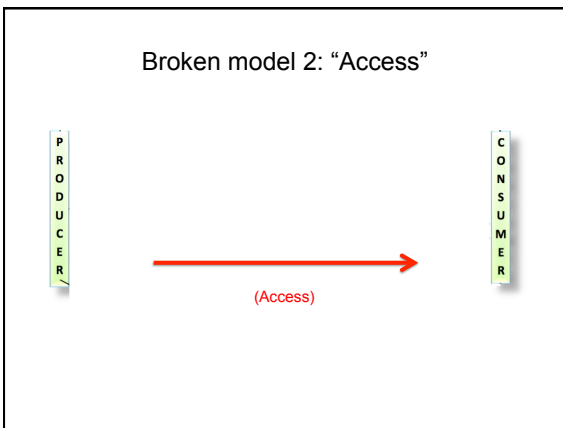
"format" does not guarantee preservability or usability.

"Online" and "tangible" are only methods of transporting bits.

They are literally the Medium and not the Message.

This policy is broken because it assumes that the medium is more important than the message.

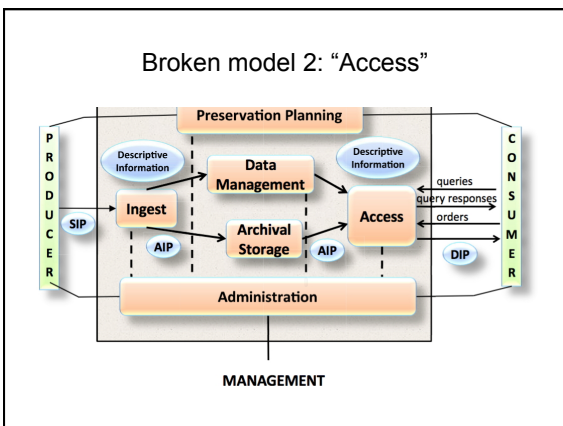
We should be paying attention to the SIP and AIP and DIP information packages and their functionality, not the medium.



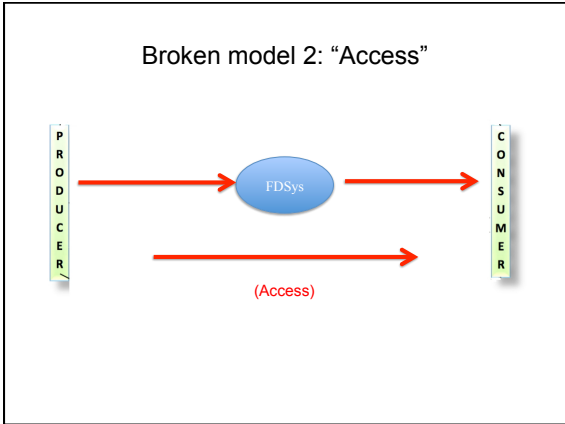
But, you might say, we have online access.

Which brings us to our second broken model: over-reliance on access.

We all love instant online access but OAIS teaches us that access from a producer without an archive is an incomplete model.



What we need is an archive role in addition to a producer/distributor role.



We do have Fdsys and some other government archives and even agency-producers who value preservation.

But what about the title 44 fugitives?  
 And what about the non-title 44 materials?  
 Do we even know what is slipping through the cracks?

When we emphasize "access" at the expense of preservation, we can be sure that a lot of material is not being preserved.

That brings us to broken model 3:

Broken model 3: Too Few Archives

TDR 3.4.1  
 The repository shall have processes in place to sustain the repository over time.

TDR tells us an archive needs to have a plan for sustainability.

---

---

---

---

---

---

---

---

Broken model 3: Too Few Archives

Is any government agency sustainable?

On the one hand, an agency can claim that it has the full faith and credit of the government, legal mandates, and (in some cases) the historical precedent of its long-term mission.

On the other hand, agencies come and go, budgets are cut and reallocated, and missions change.

---

---

---

---

---

---

---

---

Broken model 3: Too Few Archives

When an agency is the sole archive for a body of knowledge, what will happen when...

- The agency mission is changed?
- The budget is cut?
- The agency function is transferred to another agency?

---

---

---

---

---

---

---

---

Broken model 3: Too Few Archives

The solution to the sustainability problem is to have a Succession Plan:

The repository shall have an appropriate, formal succession plan. [TDR 3.1.2.1]

---

---

---

---

---

---

---

---

Broken model 3: Too Few Archives

The solution to the sustainability problem is to have a Succession Plan:

The repository shall have an appropriate, formal succession plan. [3.1.2.1]

But, who can succeed Census, or Fdsys, or NARA, or...?

One has to ask..

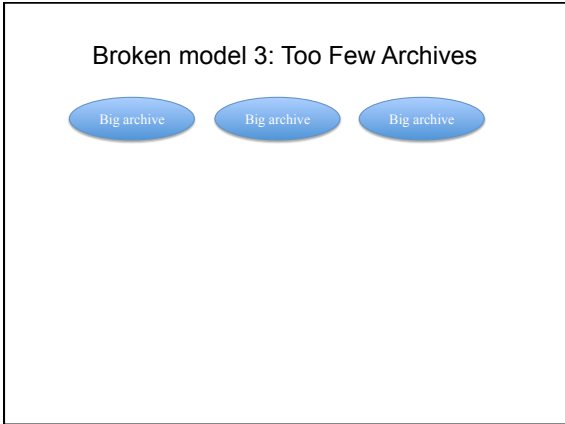
To whom would GPO turn over Fdsys?

Or,

if any agency can no longer keep "everything" available for "online access" who will pick up the slack?

---

---



The model we have today is one of a few agencies that are big and which, for the most part, do not overlap in coverage or mission.

e.g.: Library of Congress, National Library of Medicine, National Agricultural Library and the National Library of Education.  
 e.g.: NARA and GPO and NASA and NTIS

It would be hard, if not impossible for any of these big agencies to serve as a succession partner of any of the others.

With our current, broken model of too few archives, NO agency can claim to have a workable succession plan.

Solutions.

Knowing what we do about OAIS and TDR, what solutions might resolve these 3 problems and provide us with a more OAIS-compliant FDLP?

Here are my 3 suggestions...

---



---



---



---

Solutions.

1. We need **multiple repositories** serving the needs of many different, specific **"Designated Communities"** of users.

---



---



---



---



---



---

Solutions.

- 2. We need **preservable digital objects** that can be deposited into FDLP libraries and cease reliance on the flawed alternatives of "online" vs. "tangible."

---

---

---

---

---

---

---

Solutions.

- 3. We need an information-preservation "ecosystem" consisting of many repositories capable of cooperating with each other's **succession planning**.

---

---

---

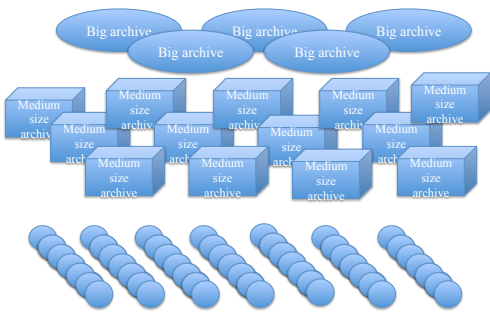
---

---

---

---

Solutions.



Such an "ecosystem" might look like this.

---

---

---

---

---

---

---

Solutions.

- Focus on **Information**, not institutions
- Focus on **Functionality** and **understandability**, not media.
- Focus on **Long-term Preservation for Usability**, not on short-term access.
- Focus on many specific **designated communities** of users.

And it would be based on principles that we learn from OAIS and TDR:

---

---

---

---

---

---