Preservation & PREMIS

ARD Prasad

DRTC, ISI

What is Digital Preservation

- **Digital preservation** - management of digital information over time
- To preserve and maintain readability and accessibility
- It is a set of processes and activities
- **Preservation Metadata is essential**

Advantages

- Ensures long-term access and availability
- Older document can be retrieved at a later stage
- Access anywhere (Internet)
- Cross reference and other document can also be retrieved (links)

Limitations/Problems

- For some libraries cost factor is high
- Data can become corrupt
- Security
- File Formats longevity

Issues and challenges

Organisational & Managerial

- Digital preservation does not yet form an integral part of the institution's corporate / information strategy – *lack of organisational infrastructure and skilled staff*
- *Core funding* for institutions does not grow in line with information growth; many institutional repositories rely on short-term project funding
- *Costs for preservation* are in general difficult to calculate and are poorly understood (difficult to segregate costs for preservation from costs for access)
- Organisational model – relationship between institutional repositories and external preservation agencies

Issues and challenges

Technical

- Digital Media
- Standards settling down but can not remain forever
- Little preservation metadata is currently being collected for content within the institutional repositories

Issues/Challenges

- Hardware and Software are becoming obsolete in very short periods of time
- Incompatibility of different versions of hard- and software
- Fading knowledge of how to use older hardware and software
- Aging and decaying storage media
- **Loss of Information**

Example – Loss of Information

J. reine angew. Math. 494 (1998), 1–34

Journal für die reine und angewandte Mathematik
© Walter de Gruyter
Berlin · New York 1998

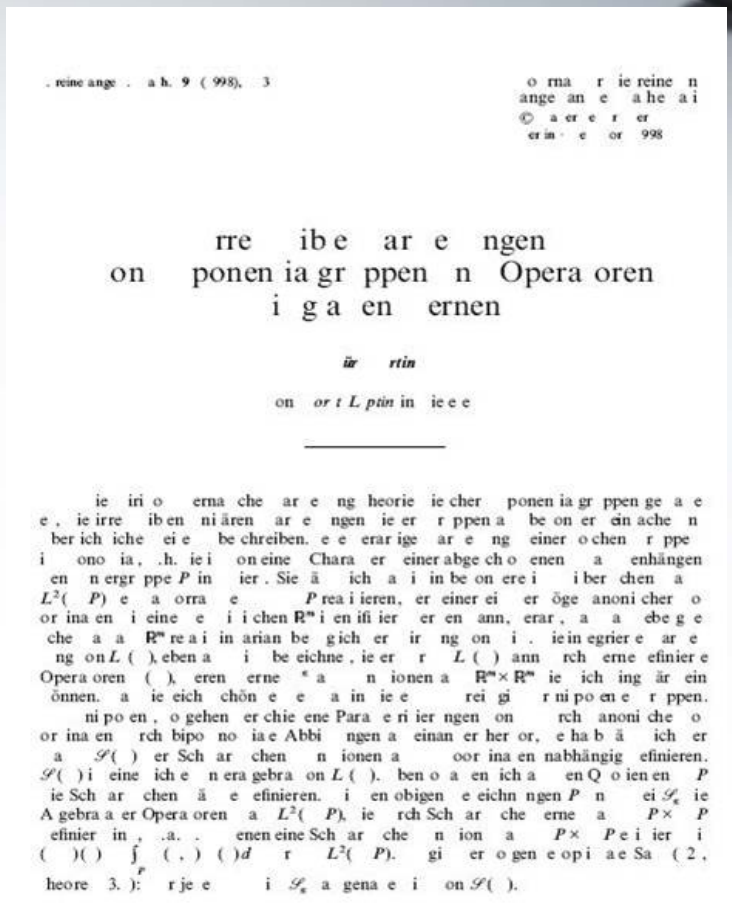
Irreduzible Darstellungen
von Exponentialgruppen und Operatoren
mit glatten Kernen

Für Martin

Von *Horst Leptin* in Bielefeld

Die Kirillow-Bernatsche Darstellungstheorie Liescher Exponentialgruppen gestattet es, die irreduziblen unitären Darstellungen dieser Gruppen auf besonders einfache und übersichtliche Weise zu beschreiben. Jede derartige Darstellung π einer solchen Gruppe G ist monomial, d.h. sie ist von einem Charakter χ einer abgeschlossenen zusammenhängenden Untergruppe P induziert. Sie läßt sich damit insbesondere im Hilbertschen Raum $L^2(G/P)$ des Faktorraumes $\overline{G} = G/P$ realisieren, der seinerseits vermöge kanonischer Koordinaten mit einem euklidischen $\mathbb{R}^m$ identifiziert werden kann, derart, daß das Lebesguesche Maß auf $\mathbb{R}^m$ relativ invariant bezüglich der Wirkung von G ist. Die integrierte Darstellung von $L^1(G)$, ebenfalls mit π bezeichnet, liefert für $f \in L^1(G)$ dann durch Kerne definierte Operatoren $\pi(f)$, deren Kerne K_f^π als Funktionen auf $\mathbb{R}^m \times \mathbb{R}^m$ ziemlich singulär sein können. Das vielleicht schönste Resultat in diesem Umkreis gilt für nilpotente Gruppen. Ist G nilpotent, so gehen verschiedene Parametrisierungen von G durch kanonische Koordinaten durch bipolynomiale Abbildungen auseinander hervor, deshalb läßt sich der Raum $\mathscr{S}(G)$ der Schwartzschen Funktionen auf G koordinatenunabhängig definieren. $\mathscr{S}(G)$ ist eine dichte Unteralgebra von $L^1(G)$. Ebenso lassen sich auf den Quotienten G/P die Schwartzschen Räume definieren. Mit den obigen Bezeichnungen P und π sei $\mathscr{A}_\pi$ die Algebra aller Operatoren $\tilde{K}$ auf $L^2(G/P)$, die durch Schwartzsche Kerne K auf $G/P \times G/P$ definiert sind, m.a.W. zu denen eine Schwartzsche Funktion K auf $G/P \times G/P$ existiert mit $(\tilde{K}\varphi)(x) = \int_{G/P} K(x,y)\varphi(y)dy$ für $\varphi \in L^2(G/P)$. Es gilt der folgende optimale Satz ([2], Theorem 3.4): Für jedes $\pi \in \hat{G}$ ist $\mathscr{A}_\pi$ das genaue Bild von $\mathscr{S}(G)$.

Acrobat 5



. reine ange . a h. 9 (998), 3

o rna r ie reine n
ange an e a he a i
© a er e r er
er in · e or 998

rre ib e ar e ngen
on ponen ia gr ppen n Opera oren
i g a en ernen

ür rtin

on *or t L ptin* in ie e e

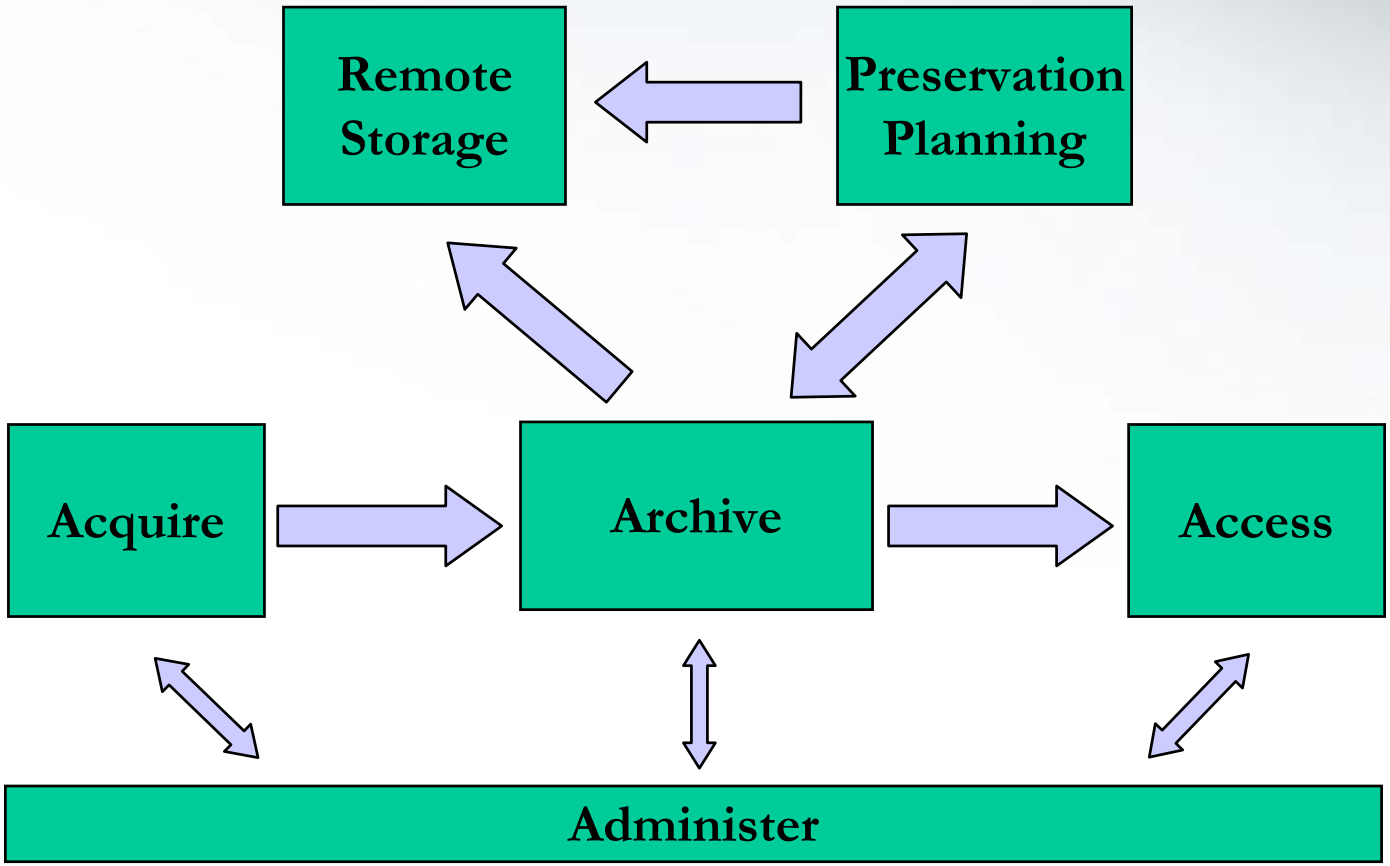
ie iri o erna che ar e ng heorie ie cher ponen ia gr ppen ge a e
e , ie irre ib en ni ären ar e ngen ie er r ppen a be on er ein ache n
ber ich iche ei e be chreiben. e e erar ige ar e ng einer o chen r ppe
i ono ia , .h. ie i on eine Chara er einer abge ch o enen a enhängen
en n ergr ppe P in ier . Sie ä ich a i in be on ere i i ber chen a
$L^2($ $P)$ e a orra e P rea i ieren, er einer ei er öge anoni cher o
or ina en i eine e i i chen $\mathbb{R}^m$ i en ifi ier er en ann, erar , a a ebe g e
che a a $\mathbb{R}^m$ re a i in arian be g ich er ir ng on i . ie in egrier e ar e
ng on L ( ), eben a i be eichne , ie er r L ( ) ann rch erne efinier e
Opera oren ( ), eren erne * a n ionen a $\mathbb{R}^m \times \mathbb{R}^m$ ie ich ing är ein
önnen. a ie eich chön e e a in ie e rei gi r ni po en e r ppen.
ni po en , o gehen er chie ene Para e ri ier ngen on rch anoni che o
or ina en rch bipo no ia e Abbi ngen a einan er her or, e ha b ä ich er
a $\mathscr{S}($) er Sch ar chen n ionen a oor ina en nabhängig efinieren.
$\mathscr{S}($) i eine ich e n era gebra on L ( ). ben o a en ich a en Q o ien en P
ie Sch ar chen ä e efinieren. i en obigen e eichn ngen P n ei $\mathscr{A}_\pi$ ie
A gebra a er Opera oren a $L^2($ $P)$, ie rch Sch ar che erne a $P \times P$
efinier in , .a. . enen eine Sch ar che n ion a $P \times P$ e i ier i
( )( ) $\int_P$ (,) ( )d r $L^2($ $P)$. gi er o gen e op i a e Sa (2,
heore 3.): r je e i $\mathscr{A}_\pi$ a gena e i on $\mathscr{S}($).

Acrobat 7

Statesgies

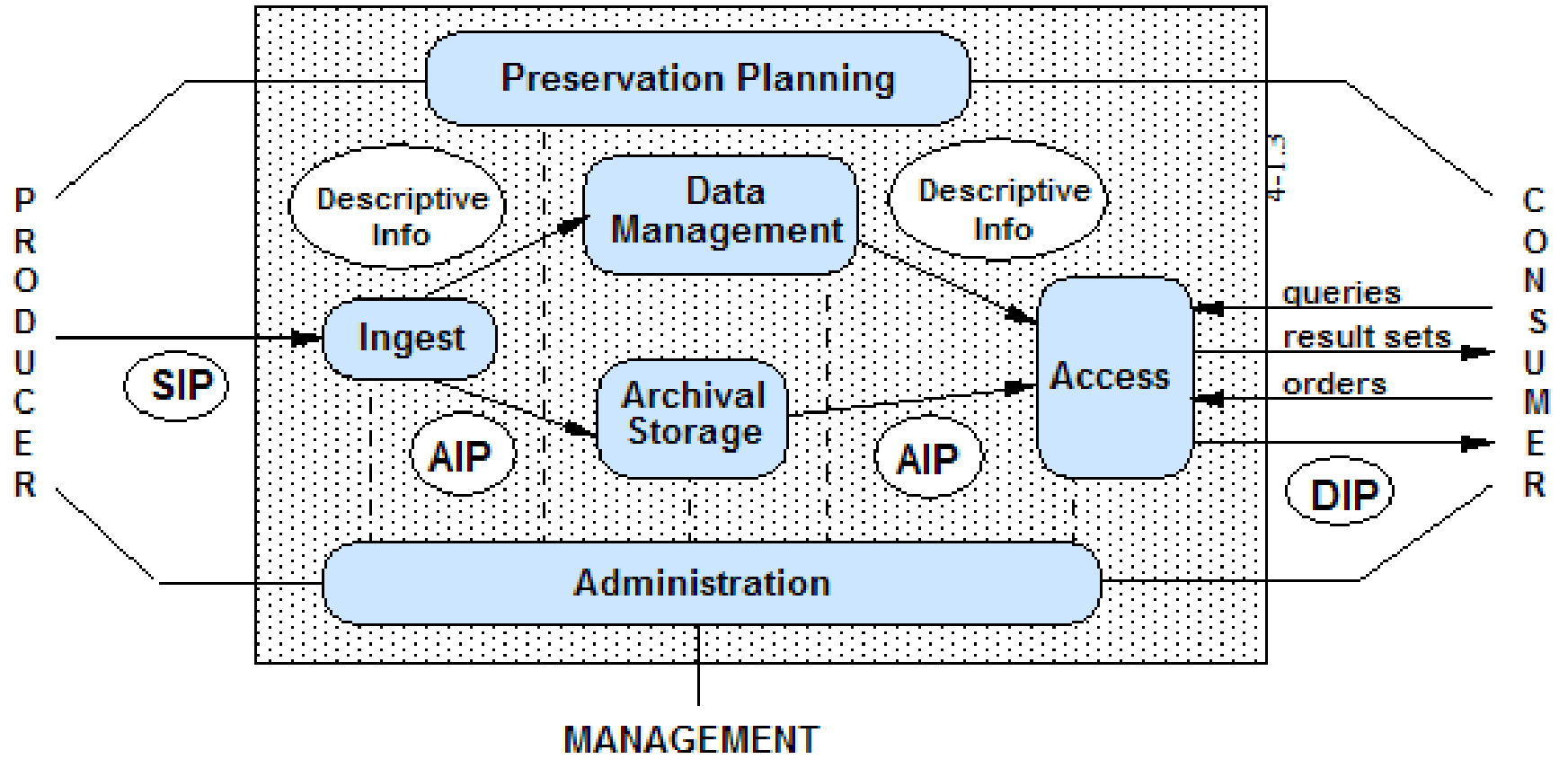
- *Replication*
- *Migration*
- *Emulation*
- *Metadata attachment*

Simplified Model for Access and Preservation



Adapted from http://www.dlib.org/dlib/july04/beagrie/07beagrie.html

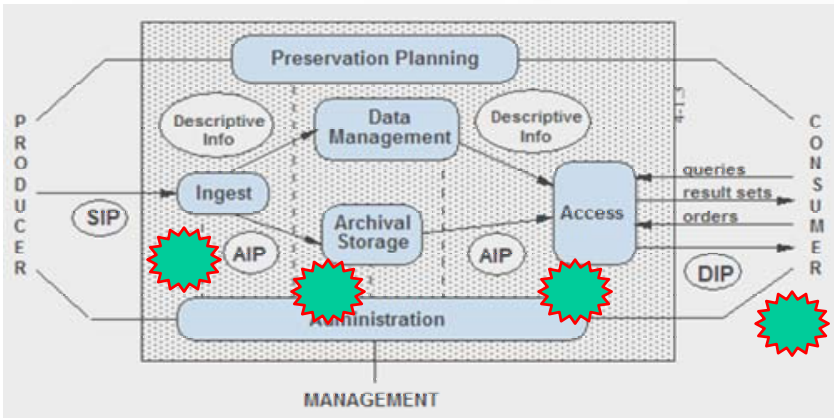
OAIS Functional Model



Reference Model for an Open Archival Information System (OAIS)

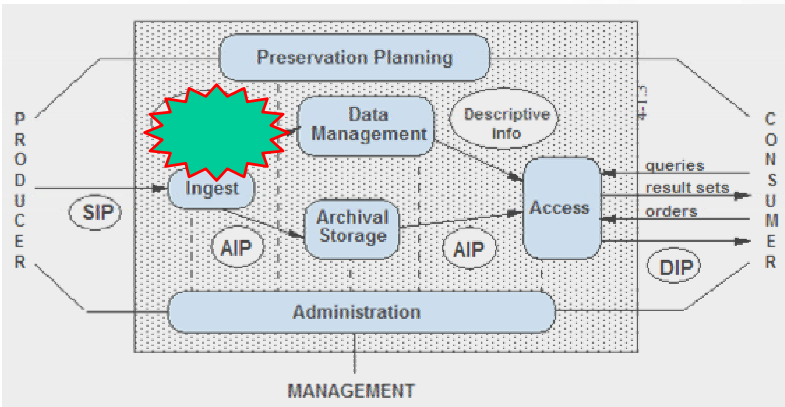
Information Packages

- Submission Information Package (SIP)
  - Producer Archive
  - Accepted Formats
- Archival Information Package (AIP)
  - Internal Storage
  - Standardized Format(s)
- Distribution Information Package (DIP)
  - Archive User
  - Output Format



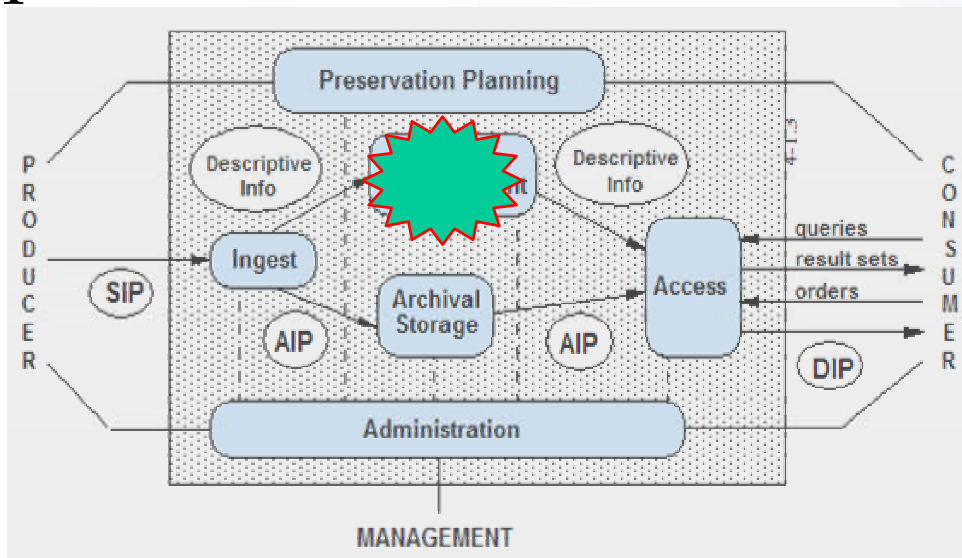
Ingest Activities

- Accepts content from producers
- Packages contents: storage, preservation, access
  - Based on formatting standards, metadata templates
  - Extract data for inclusion in archive database
- Performs quality assurance testing
- Coordinates information updates
  - Archival Storage
  - Data Management
- Generates reports



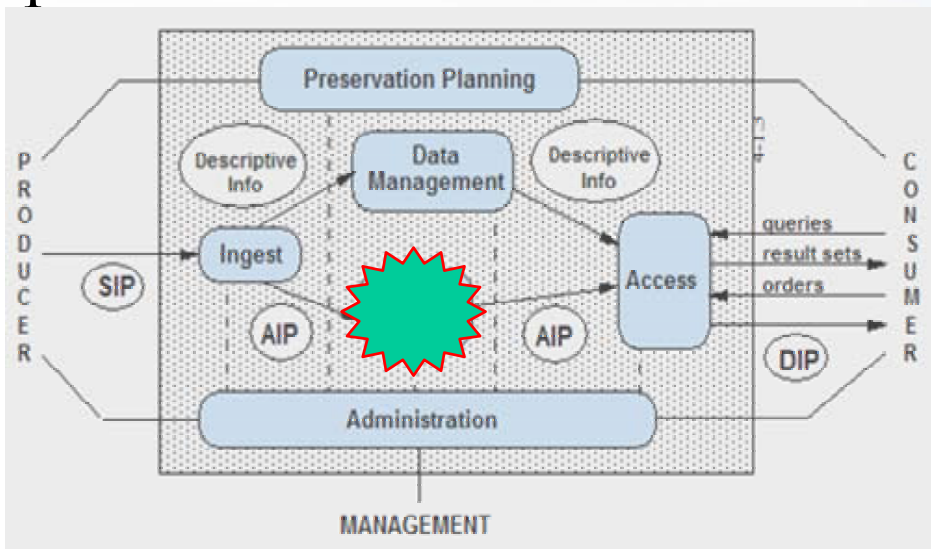
Data Management Activities

- Administers the archive database
  - Maintain schema and definitions
- Performs database updates
- Performs data management queries
- Provides management reports



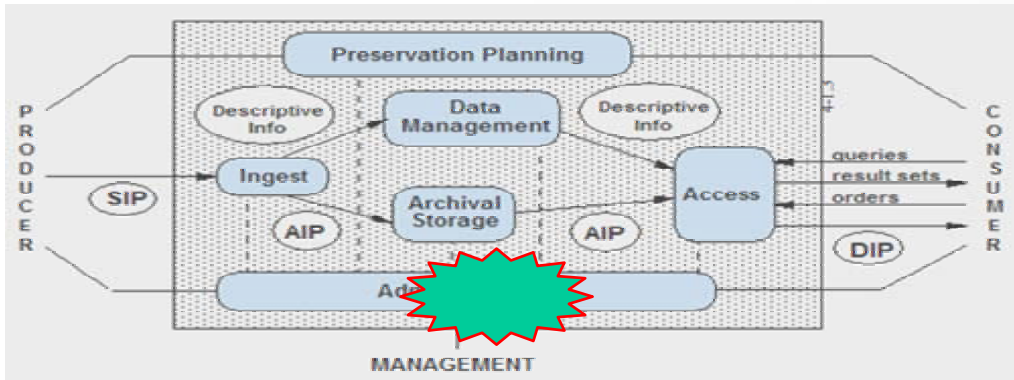
Archival Storage Activities

- Adds archival packages to permanent storage
- Manages the archival storage
- Refreshes archival media and metadata
- Performs error checking on items
- Provides disaster recovery capabilities
- Provides items to fill Access requests



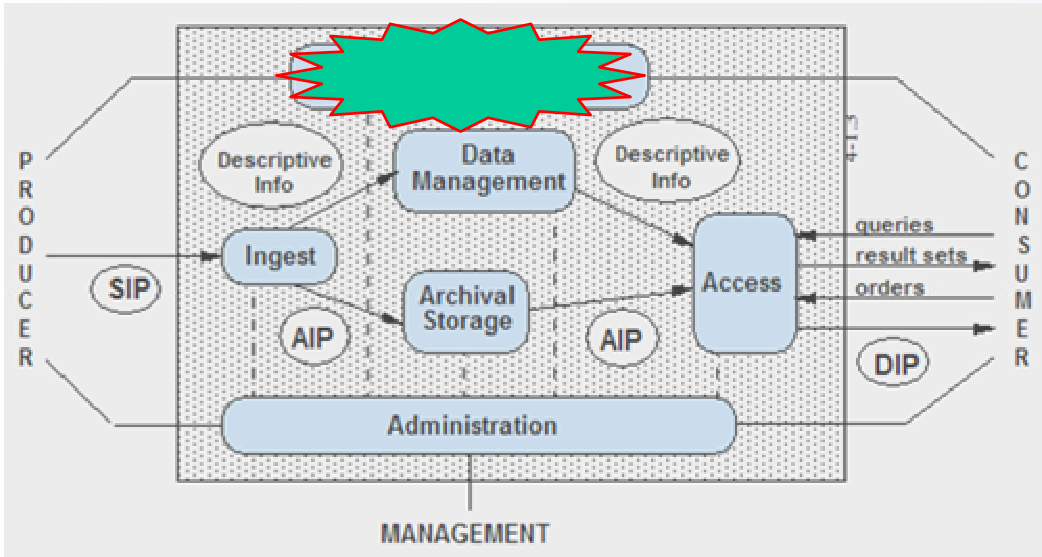
Administration Activities

- Provides for overall archive operations
- Solicits and negotiates submission agreements
- Audits & ensures submissions meet standards
- Monitors and improves archive operations
- Maintains system configuration
- Establishes archive standards, policies, services
- Inventories/reports on contents
- Migrates archive contents



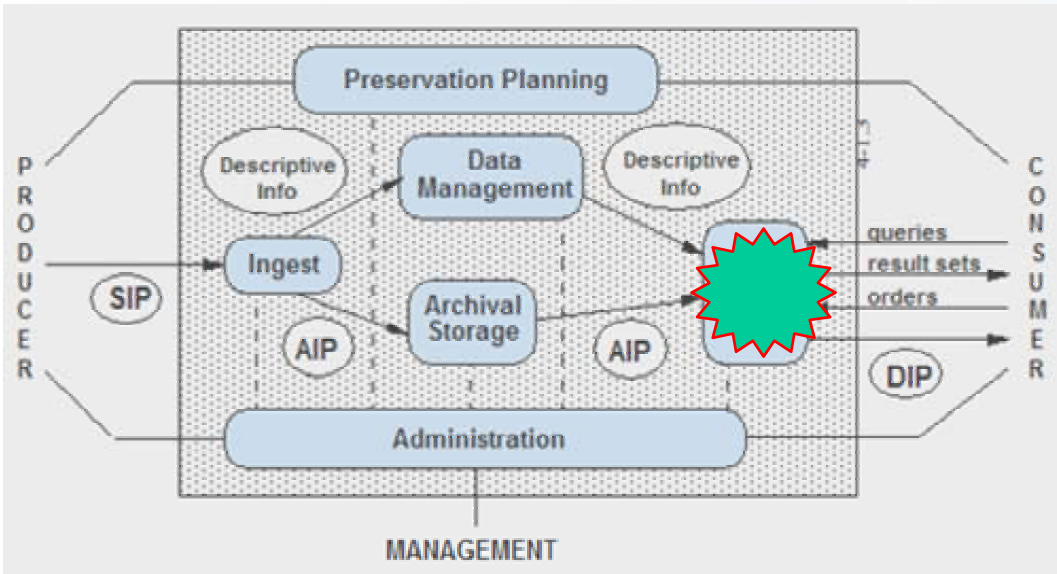
Preservation Planning Activities

- Ensures information remains accessible
- Evaluates archive contents, recommend updates
- Recommends & develops standards, policies
- Monitors technology environment changes
- Develops migration plans, prototypes, test plans
- Designs / assists with info package templates



Access Activities

- Helps users determine information available
- Coordinates information requests
- Applies controls and access limits
- Generates and delivers responses
- Ensures user satisfaction



Digital preservation technologies

Tools

- JHOVE
- Format registries
- METS
- Web Archives Workbench

Applications and Initiatives

- E-Depot (KB)
- Portico (E-Archive)
- PANDORA (NLA)
- FCLA Digital Archive (DAITSS)

Standards and Frameworks

- OAIS
- Z39.87
- PREMIS
- PDF/A

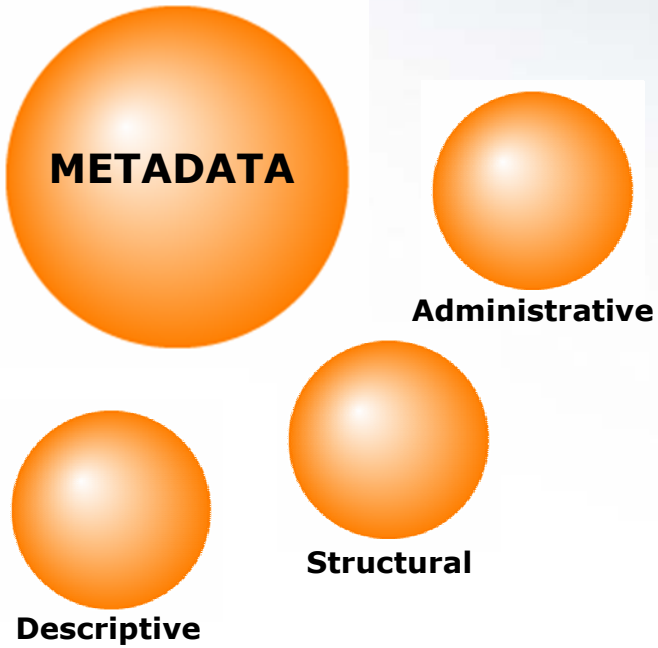
Preservation Strategies

- Media migration
- Format migration
- Emulation
- Universal Virtual Computer

PREMIS

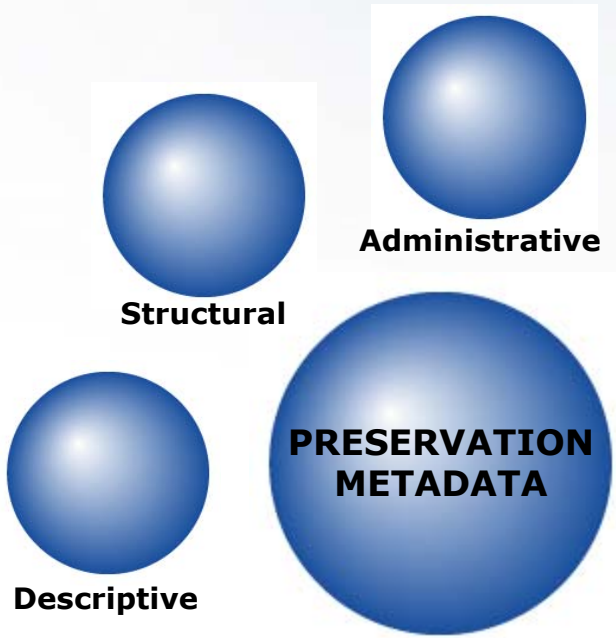
PREservation Metadata: Implementation Strategies

Metadata and Preservation Metadata



"Structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource"

"Information that supports and documents the digital preservation process"



Preservation Metadata

- Preservation metadata is sometimes considered a subset of technical or administrative metadata
- Preservation metadata stores technical details on the format, structure and use of the digital content, the history of all actions performed on the resource including

Preservation Metadata Functions

- Information that supports and documents the digital preservation process:
  - Establish provenance: track chain of custody and alterations over time
  - Details authenticity
  - Describes technical details of object
  - Documents technical processes object has undergone
  - Describes the environment from which it originated
  - Specify rights management information

Back ground

- March 2000: OCLC, RLG jointly sponsored international working group on preservation metadata
  - Identify key issues, seek consensus
- White paper (January 2001)
  - Defined preservation metadata; role in preservation process
  - Reviewed/synthesized existing preservation metadata initiatives
- Preservation metadata framework (June 2002)
  - Comprehensive description of types of information constituting preservation metadata
  - Based on OAIS information model
  - Set of "prototype" preservation metadata elements

Objectives

- Define “core” set of preservation metadata elements, with supporting data dictionary, applicable to broad range of digital preservation activities
- Identify and evaluate alternative strategies for encoding, storing, managing, and exchanging preservation metadata

How PREMIS defines Preservation metadata

- "The information a repository uses to support the digital preservation process"
- Metadata that supports
  - viability
  - renderability
  - understandability
  - authenticity
  - identity
- Mandatory elements represent "the minimum amount for a second repository to accept custody of a digital object and assume responsibility for its long-term preservation"

PREMIS goals

- Build on the OAIS reference model
- Be implementation independent
- “Provide a starting point for improvements and enhancements based on community experience and feedback”

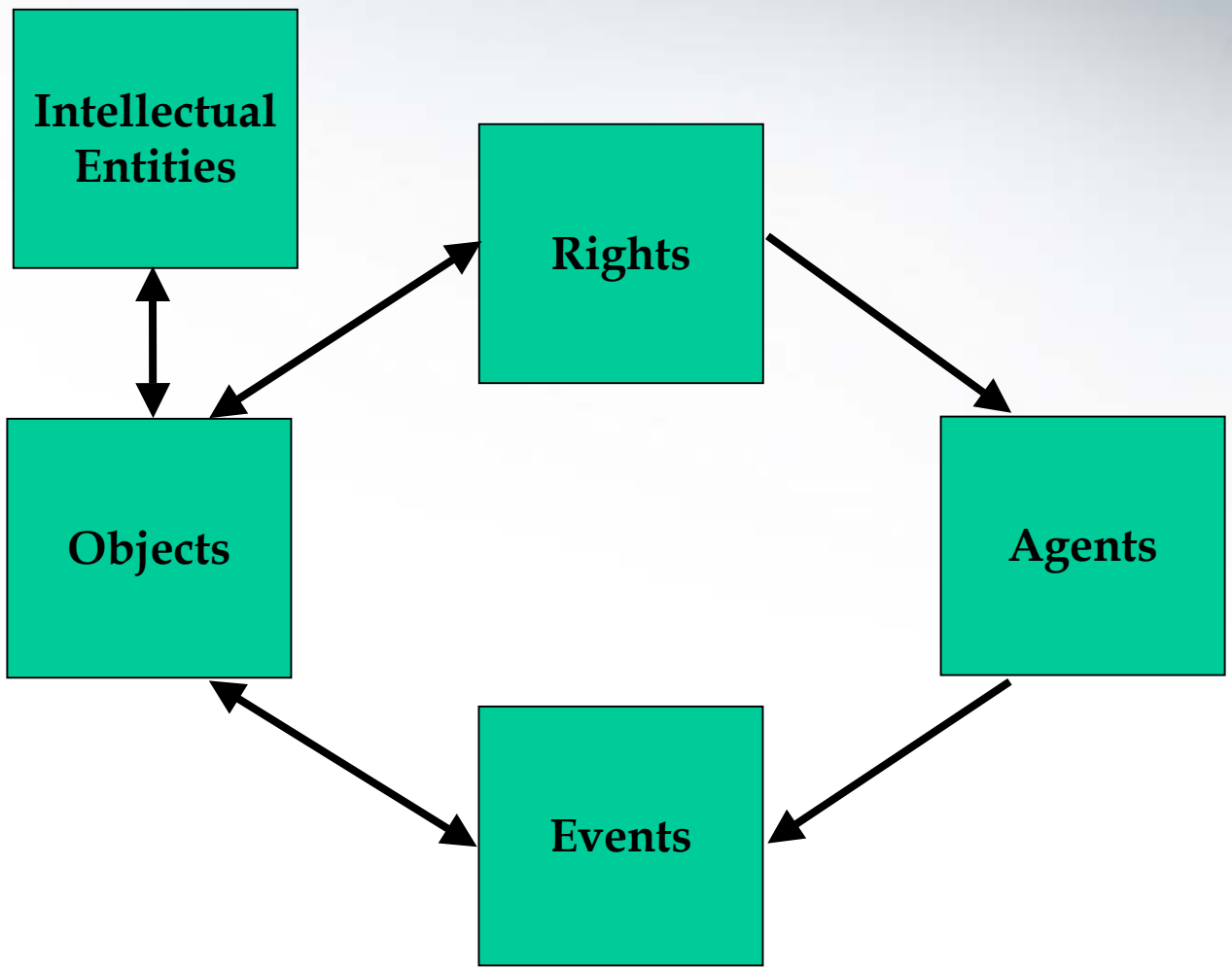
Development strategies

- Particular attention to documenting
  - digital provenance
  - relationships
- "Whenever possible the group defined elements that do not require human intervention to supply or analyze," but did not limit to these
- Defined "semantic units" rather than "metadata elements"

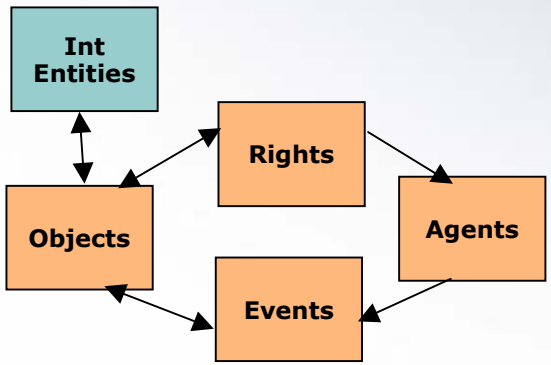
The PREMIS Data Model

- Data model includes:
  - Entities: "things" relevant to digital preservation that are described by preservation metadata (Intellectual Entities, Objects, Events, Rights, Agents)
  - Properties of Entities (semantic units)
  - Relationships between Entities

PREMIS Data Model



Intellectual Entity



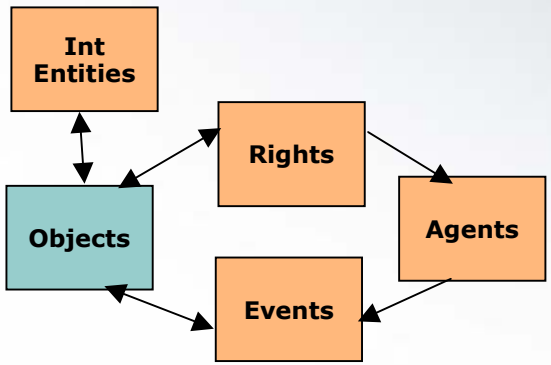
Examples:

- Rabbit Run by John Updike (a book)
- "Maggie at the beach"

  (a photograph)
- The Library of Congress Website (a website)
- The Library of Congress: American Memory Home page (a web page)

- Set of content that is considered a single intellectual unit for purposes of management and description (e.g., a book, a photograph, a map, a database)
- May include other Intellectual Entities (e.g. a website that includes a web page)
- **Has one or more digital representations**
- Not fully described in PREMIS DD, but can be linked to in metadata describing digital representation

Object

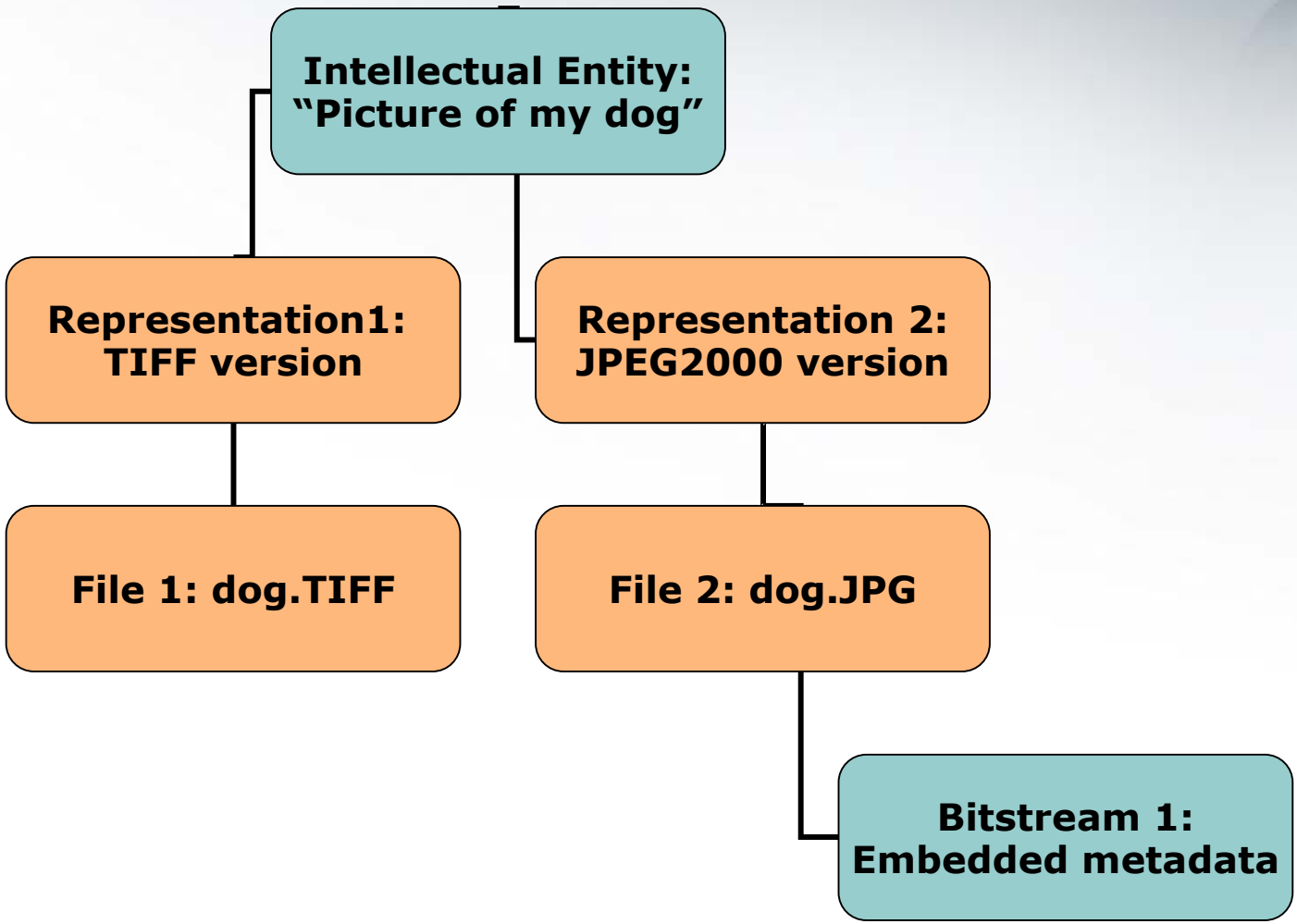


Examples:

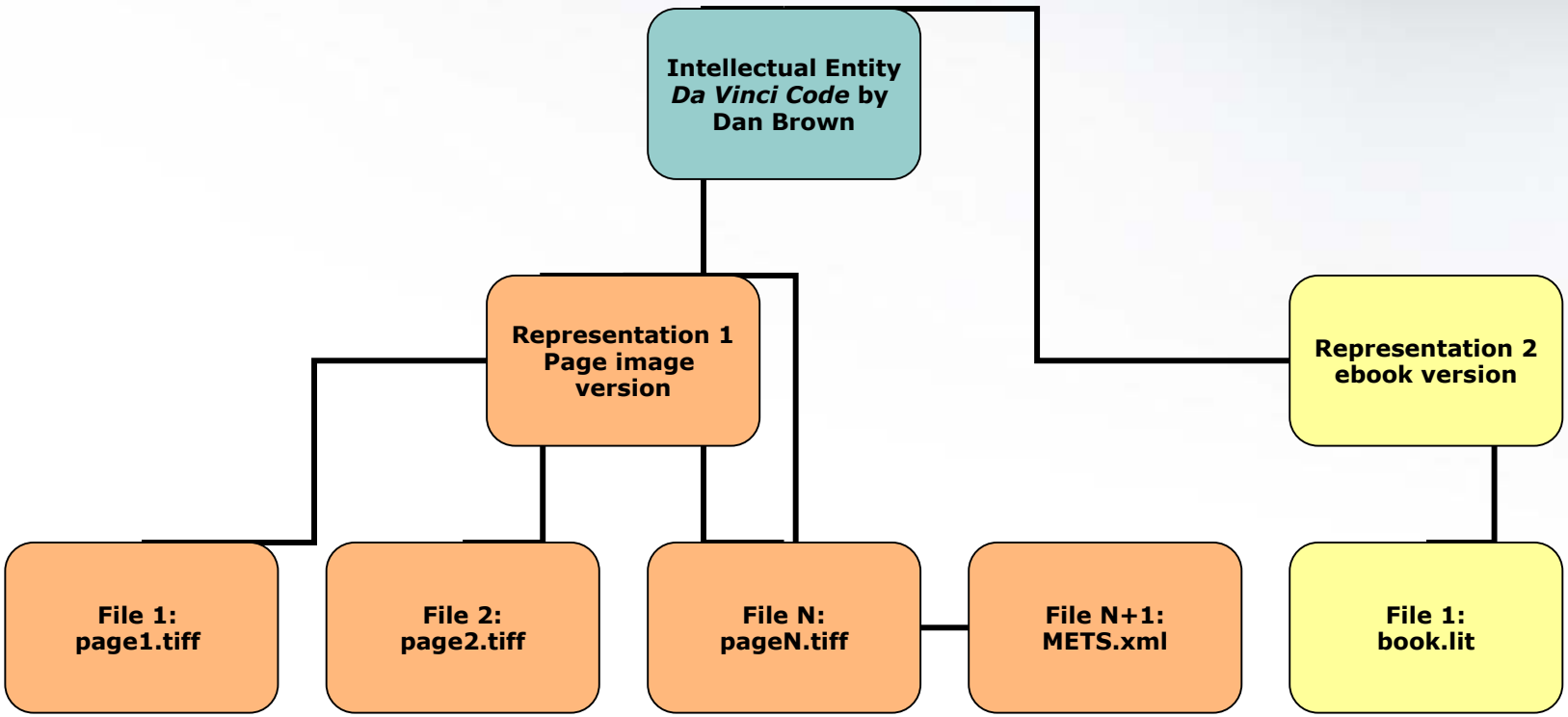
- chapter1.pdf (a file)
- chapter1.pdf + chapter2.pdf + chapter3.pdf (representation of a book w/3 chapters)
- TIFF file containing header and 2 images (2 bitstreams (images), each with own set of properties (semantic units): e.g., identifiers, technical metadata, inhibitors, …)

- Discrete unit of information in digital form
- **Objects are what repository actually preserves**
- Three types of Object:
  - **FILE:** named and ordered sequence of bytes that is known by an operating system
  - **REPRESENTATION:** set of files, including structural metadata, that, taken together, constitute a complete rendering of an Intellectual Entity
  - **BITSTREAM:** data within a file with properties relevant for preservation purposes (but needs additional structure or reformatting to be stand-alone file)

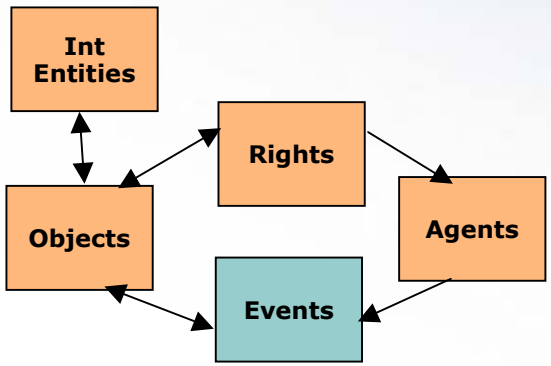
Object Example 1: photo in two formats



Object Example 2: book in two versions



Event

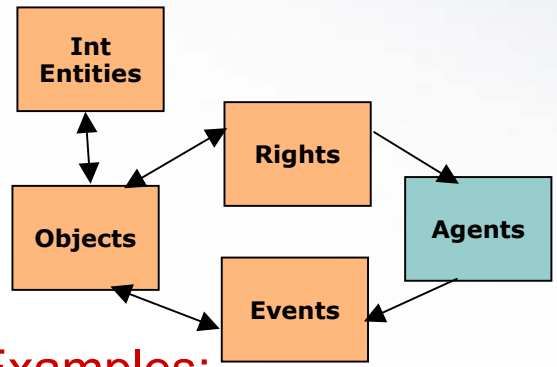


Examples:

- Validation Event: use JHOVE tool to verify that chapter1.pdf is a valid PDF file
- Ingest Event: transform an OAIS SIP into an AIP (one Event or multiple Events?)
- Migration Event: create a new version of an Object in an up-to-date format

- An action that involves or impacts at least one Object or Agent associated with or known by the preservation repository
- Helps document digital provenance. Can track history of Object through the chain of Events that occur during the Objects lifecycle
- Determining which Events are in scope is up to the repository (e.g., Events which occur before ingest, or after de-accession)
- Determining which Events should be recorded, and at what level of granularity is up to the repository

Agent

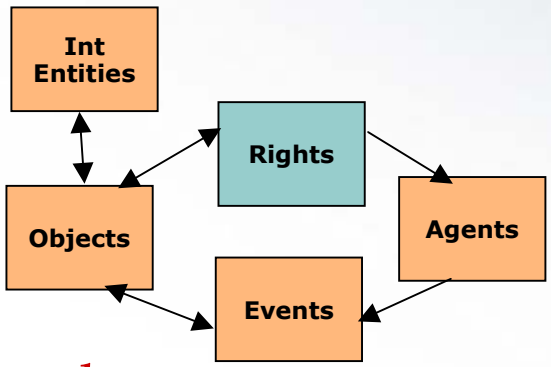


Examples:

- Priscilla Caplan (a person)
- Florida Center for Library Automation (an organization)
- Dark Archive in the Sunshine State implementation (a system)
- JHOVE version 1.0 (a software program)

- Person, organization, or software program/system associated with an Event or a Right (permission statement)
- Agents are associated only indirectly to Objects through Events or Rights
- Not defined in detail in PREMIS DD; not considered core preservation metadata beyond identification

Rights



Example:

- Priscilla Caplan grants FCLA digital repository permission to make three copies of metadata_fundamentals.pdf for preservation purposes.

- An agreement with a rights holder that grants permission for the repository to undertake an action(s) associated with an Object(s) in the repository.
- Not a full rights expression language; focuses exclusively on permissions that take the form:
  - Agent X grants Permission Y to the repository in regard to Object Z.

Relationships between entities

- Between objects
  - Structural relationships
  - Derivation relationships
  - Dependency relationships
- Others defined by data model indicated in data dictionary by linking attributes

The PREMIS data dictionary

- Defines semantic units for:
  - *Objects*
  - *Events*
  - *Agents*
  - *Rights*
- Intellectual Entity is out of scope because it is "well served by descriptive metadata"

Entries include information on

- Name
- Semantic components
- Definition
- Rationale
- Data constraint
- Object category
- Applicability
- Examples
- Repeatability
- Obligation
- Creation/Maintenance notes
- Usage notes

Sample data dictionary entry

Semantic unit	format		
Semantic components	formatDesignation, formatRegistry		
Definition	Identification of the format of a file or bitstream where format is the organization of digital information according to preset specifications.		
Rationale	Many preservation activities depend on detailed knowledge about the format of the digital object. An accurate identification of format is essential. The identification provided, whether by name or pointer into a format registry, should be sufficient to associate the object with more detailed format information.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Repeatability		Not repeatable	Not repeatable
Obligation		Mandatory	Mandatory
Creation/ Maintenance notes	The format of a file or bitstream should be ascertained by the repository on ingest. Even if this information is provided by the submitter, directly in metadata or indirectly via the file name extension, recommended practice is to independently identify the format by parsing the file when possible. If the format can not be identified at the time of ingest, it is valid to record that the format is unknown, but the repository should subsequently make an effort to identify the format, even if manual intervention is required.		
Usage notes	A bitstream embedded within a file may have different characteristics than the larger file. For example, a bitstream in LaTex format could be embedded within an SGML file, or multiple images using different colorspaces could be embedded within a TIFF file. Format must be recorded for every file. When the bitstream format can be recognized by the repository and the repository might want to treat the bitstream differently from the embedding file for preservation purposes, format can be recorded for embedded bitstreams. Either formatDesignation or formatRegistry should be recorded. Both are optional, but since format (the container) is mandatory, one of these must be used. See "Format information," page 4-1.		

Role of a preservation policy

- PREMIS helps a repository to implement a preservation policy; it doesn't *set* that policy
- Policy can be complicated
  - Is descriptive metadata part of an Intellectual Entity?
  - If so, should we treat it as a file?
  - Is PREMIS data itself a file (or a bit stream) that is managed by the repository?
  - etc., ad infinitum…
- The data dictionary is only a starting point, does not include all information needed to preserve an Object

XML Schema

- Literal representations of the semantic units and attributes of the PREMIS data dictionary
- Of use for exchange of preservation objects
- Likely of less use for a repository's internal representation
- 5 separate schema
  - PREMIS container
  - Object entity
  - Event entity
  - Agent entity
  - Rights entity

PREMIS container

```
<?xml version="1.0" encoding="UTF-8" ?>
<!-- Copyright © 2005 OCLC Online Computer Library Center and RLG Research Libraries Group -->
<!-- PREMIS Preservation Metadata Schema: Object -->
<!-- Version 1.1, September 27, 2005 -->
- <!--
    Changes applied to this schema (since Version 1.0, May 17, 2005):
   - September 27, 2005.  Ray Denenberg, Library of Congress.
   - version changed to 1.1.
   - namespace changed (both xmlns and targetNamespace
   from http://www.loc.gov/standards/premis
   to http://www.loc.gov/standards/premis/v1.
   - version attribute added.
   - include declaration changes version e.g. "Object-v1-0" to "Object-v1-1" etc.
-->
- <xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" targetNamespace="http://www.loc.gov/standards/premis/v1"
    xmlns="http://www.loc.gov/standards/premis/v1" elementFormDefault="qualified" attributeFormDefault="unqualified">
   <xs:include schemaLocation="Object-v1-1.xsd" />
   <xs:include schemaLocation="Event-v1-1.xsd" />
   <xs:include schemaLocation="Agent-v1-1.xsd" />
   <xs:include schemaLocation="Rights-v1-1.xsd" />
  - <xs:element name="premis">
    - <xs:complexType>
      - <xs:sequence>
          <xs:element ref="object" minOccurs="1" maxOccurs="unbounded" />
          <xs:element ref="event" minOccurs="0" maxOccurs="unbounded" />
          <xs:element ref="agent" minOccurs="0" maxOccurs="unbounded" />
          <xs:element ref="rights" minOccurs="0" maxOccurs="unbounded" />
        </xs:sequence>
      - <xs:attribute name="version">
        - <xs:simpleType>
          - <xs:restriction base="xs:string">
              <xs:enumeration value="1.1" />
            </xs:restriction>
          </xs:simpleType>
        </xs:attribute>
      </xs:complexType>
    </xs:element>
  </xs:schema>
```

Object Entity

```
<?xml version="1.0" encoding="UTF-8" ?>
<!-- Copyright © 2005 OCLC Online Computer Library Center and RLG Research Libraries Group -->
<!-- PREMIS Preservation Metadata Schema: Object -->
<!-- Version 1.1, September 27, 2005 -->
- <!--
    Changes applied to this schema (since Version 1.0, May 17, 2005):
   - September 27, 2005.  Ray Denenberg, Library of Congress.
   - version changed to 1.1.
   - namespace changed (both xmlns and targetNamespace
   from http://www.loc.gov/standards/premis
   to http://www.loc.gov/standards/premis/v1.
   - version attribute added.
   - element relatedEventIdentification (within element relationship)
   added  minOccurs="0"
   - element inhibitorTarget (within element inhibitors) added attribute type=xs:string
   - corrected xlink namespace URL to "http://www.w3.org/1999/xlink" (removed slash at end)
   - modified dateCreatedByApplication type attribute to type="dateType"
     <xs:union memberTypes="xs:date xs:dateTime">
   - element storageMedium (within element storage) added maxOccurs="1"
   -October 20, 2005. Rebecca Guenther, Library of Congress
   - modified dateCreatedByApplication: take out type="dateType" change to:
          <xs:simpleType>
          <xs:union memberTypes="xs:date xs:dateTime" />
          </xs:simpleType>
-->
- <xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" targetNamespace="http://www.loc.gov/standards/premis/v1"
   xmlns="http://www.loc.gov/standards/premis/v1" elementFormDefault="qualified" attributeFormDefault="unqualified">
  - <xs:element name="object">
    - <xs:complexType>
      - <xs:sequence>
        - <xs:element name="objectIdentifier" minOccurs="1" maxOccurs="unbounded">
          - <xs:complexType>
            - <xs:sequence>
                <xs:element name="objectIdentifierType" minOccurs="1" maxOccurs="1" type="xs:string" />
                <xs:element name="objectIdentifierValue" minOccurs="1" maxOccurs="1" type="xs:string" />
              </xs:sequence>
              <xs:anyAttribute namespace="http://www.w3.org/1999/xlink" />
            </xs:complexType>
          </xs:element>
          <xs:element name="preservationLevel" minOccurs="0" maxOccurs="1" type="xs:string" />
          <xs:element name="objectCategory" minOccurs="1" maxOccurs="1" type="xs:string" />
        - <xs:element name="objectCharacteristics" minOccurs="0" maxOccurs="unbounded">
          - <xs:complexType>
            - <xs:sequence>
                <xs:element name="compositionLevel" minOccurs="0" maxOccurs="1" type="xs:nonNegativeInteger" />
```

Event entity

```
<?xml version="1.0" encoding="UTF-8" ?>
<!-- Copyright © 2005 OCLC Online Computer Library Center and RLG Research Libraries Group -->
<!-- PREMIS Preservation Metadata Schema: Event -->
<!-- Version 1.1, September 27, 2005 -->
- <!--
    Changes applied to this schema (since Version 1.0, May 17, 2005):
   - September 27, 2005.  Ray Denenberg, Library of Congress.
   - version changed to 1.1.
   - namespace changed (both xmlns and targetNamespace
   from http://www.loc.gov/standards/premis
   to http://www.loc.gov/standards/premis/v1.
   - version attribute added.
   - corrected xlink namespace URL to "http://www.w3.org/1999/xlink" (removed slash at end)
   - modified eventDateTime type attribute to type="dateType"
     <xs:union memberTypes="xs:date xs:dateTime">
   -October 20, 2005. Rebecca Guenther, Library of Congress
   - modified eventDateTime: take out type="dateType" change to:
          <xs:simpleType>
          <xs:union memberTypes="xs:date xs:dateTime" />
          </xs:simpleType>
-->
- <xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" xmlns="http://www.loc.gov/standards/premis/v1"
  targetNamespace="http://www.loc.gov/standards/premis/v1" elementFormDefault="qualified" attributeFormDefault="unqualified">
 - <xs:element name="event">
  - <xs:complexType>
   - <xs:sequence>
    - <xs:element name="eventIdentifier">
     - <xs:complexType>
      - <xs:sequence>
          <xs:element name="eventIdentifierType" type="xs:string" />
          <xs:element name="eventIdentifierValue" type="xs:string" />
        </xs:sequence>
        <xs:anyAttribute namespace="http://www.w3.org/1999/xlink" />
       </xs:complexType>
     </xs:element>
     <xs:element name="eventType" type="xs:string" />
    - <xs:element name="eventDateTime" minOccurs="1" maxOccurs="1">
     - <xs:simpleType>
        <xs:union memberTypes="xs:date xs:dateTime" />
       </xs:simpleType>
     </xs:element>
     <xs:element name="eventDetail" type="xs:string" minOccurs="0" />
    - <xs:element name="eventOutcomeInformation" minOccurs="0" maxOccurs="unbounded">
     - <xs:complexType>
      - <xs:sequence>
```

Agent entity

```
<?xml version="1.0" encoding="UTF-8" ?>
<!-- Copyright © 2005 OCLC Online Computer Library Center and RLG Research Libraries Group -->
<!-- PREMIS Preservation Metadata Schema: Agent -->
<!-- Version 1.1, September 27, 2005 -->
- <!--
    Changes applied to this schema (since Version 1.0, May 17, 2005):
   - September 27, 2005.  Ray Denenberg, Library of Congress.
   - version changed to 1.1.
   - namespace changed (both xmlns and targetNamespace
   from http://www.loc.gov/standards/premis
   to http://www.loc.gov/standards/premis/v1.
   - version attribute added.
   - corrected xlink namespace URL to "http://www.w3.org/1999/xlink" (removed slash at end)

-->
- <xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" targetNamespace="http://www.loc.gov/standards/premis/v1"
    xmlns="http://www.loc.gov/standards/premis/v1" elementFormDefault="qualified" attributeFormDefault="unqualified">
  - <xs:element name="agent">
    - <xs:complexType>
      - <xs:sequence>
        - <xs:element name="agentIdentifier" minOccurs="1" maxOccurs="unbounded">
          - <xs:complexType>
            - <xs:sequence>
                <xs:element name="agentIdentifierType" minOccurs="1" maxOccurs="1" type="xs:string" />
                <xs:element name="agentIdentifierValue" minOccurs="1" maxOccurs="1" type="xs:string" />
              </xs:sequence>
              <xs:anyAttribute namespace="http://www.w3.org/1999/xlink" />
            </xs:complexType>
          </xs:element>
          <xs:element name="agentName" minOccurs="0" maxOccurs="unbounded" type="xs:string" />
          <xs:element name="agentType" minOccurs="0" maxOccurs="1" type="xs:string" />
        </xs:sequence>
        <xs:attribute name="xmlID" type="xs:ID" />
      - <xs:attribute name="version" use="optional">
        - <xs:simpleType>
          - <xs:restriction base="xs:string">
              <xs:enumeration value="1.1" />
            </xs:restriction>
          </xs:simpleType>
        </xs:attribute>
      </xs:complexType>
    </xs:element>
  </xs:schema>
```

Right entity

```
- <xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" targetNamespace="http://www.loc.gov/standards/premis/v1"
    xmlns="http://www.loc.gov/standards/premis/v1" elementFormDefault="qualified" attributeFormDefault="unqualified">
  - <xs:element name="rights">
    - <xs:complexType>
      - <xs:sequence>
        - <xs:element name="permissionStatement" minOccurs="1" maxOccurs="unbounded">
          - <xs:complexType>
            - <xs:sequence>
              - <xs:element name="permissionStatementIdentifier" minOccurs="1" maxOccurs="1">
                - <xs:complexType>
                  - <xs:sequence>
                      <xs:element name="permissionStatementIdentifierType" type="xs:string" minOccurs="1" maxOccurs="1" />
                      <xs:element name="permissionStatementIdentifierValue" type="xs:string" minOccurs="1" maxOccurs="1" />
                    </xs:sequence>
                    <xs:anyAttribute namespace="http://www.w3.org/1999/xlink" />
                  </xs:complexType>
                </xs:element>
              - <xs:element name="linkingObject" minOccurs="1" maxOccurs="unbounded">
                - <xs:complexType>
                  - <xs:simpleContent>
                    - <xs:extension base="xs:string">
                        <xs:attribute name="linkingObjectXmlID" type="xs:IDREF" use="optional" />
                      </xs:extension>
                    </xs:simpleContent>
                  </xs:complexType>
                </xs:element>
              - <xs:element name="grantingAgent" minOccurs="0" maxOccurs="unbounded">
                - <xs:complexType>
                  - <xs:simpleContent>
                    - <xs:extension base="xs:string">
                        <xs:attribute name="GrantAgentXmlID" type="xs:IDREF" use="optional" />
                        <xs:anyAttribute namespace="http://www.w3.org/1999/xlink" />
                      </xs:extension>
                    </xs:simpleContent>
                  </xs:complexType>
                </xs:element>
              - <xs:element name="grantingAgreement" minOccurs="0" maxOccurs="1">
                - <xs:complexType>
                  - <xs:sequence>
                    - <xs:element name="grantingAgreementIdentification" minOccurs="0" maxOccurs="1">
                      - <xs:complexType>
                        - <xs:simpleContent>
                          - <xs:extension base="xs:string">
                              <xs:anyAttribute namespace="http://www.w3.org/1999/xlink" />
                            </xs:extension>
```

Current PREMIS activity

- PREMIS Maintenance Activity hosted at the Library of Congress
- Commissioned report on Rights in the PREMIS Data Model
- Proposals for revisions of two semantic units in public comment period

Conclusion

- Need digital Preservation throughout the useful lifetime of digital data
  - Legal and safety requirements
  - Maximise potential of digital data
  - Maximise investment in digital data
- Plan from the outset for longevity and sustainable access

References

- http://www.loc.gov/standards/premis/
- http://www.digitalpreservation.gov/
- http://en.wikipedia.org/wiki/Preservation_Metadata:_Implementation_Strategies_(PREMIS)
- http://en.wikipedia.org/wiki/Preservation_Metadata
- http://www.diglib.org/preserve.htm
- <http://www.oclc.org/research/projects/pmwg/default.htm
- http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf
- http://www.dpconline.org/graphics/index.html
- http://www.ccsds.org/documents/650x0b1.pdf


Thank you