# DP Knowhow: Audit and Certification in ISO Standard 16363

Matthias Hemmje (FTK), David Giaretta (APA)

APARSEN-EGI-Community-Forum Training on Data Preservation

22nd of May 2014

Helsinki, Finland

# Why Audit?

- Assure higher management that valuable information is safe
- Identify improvements
- Justify resources needed
- Contractual requirements (in the future)

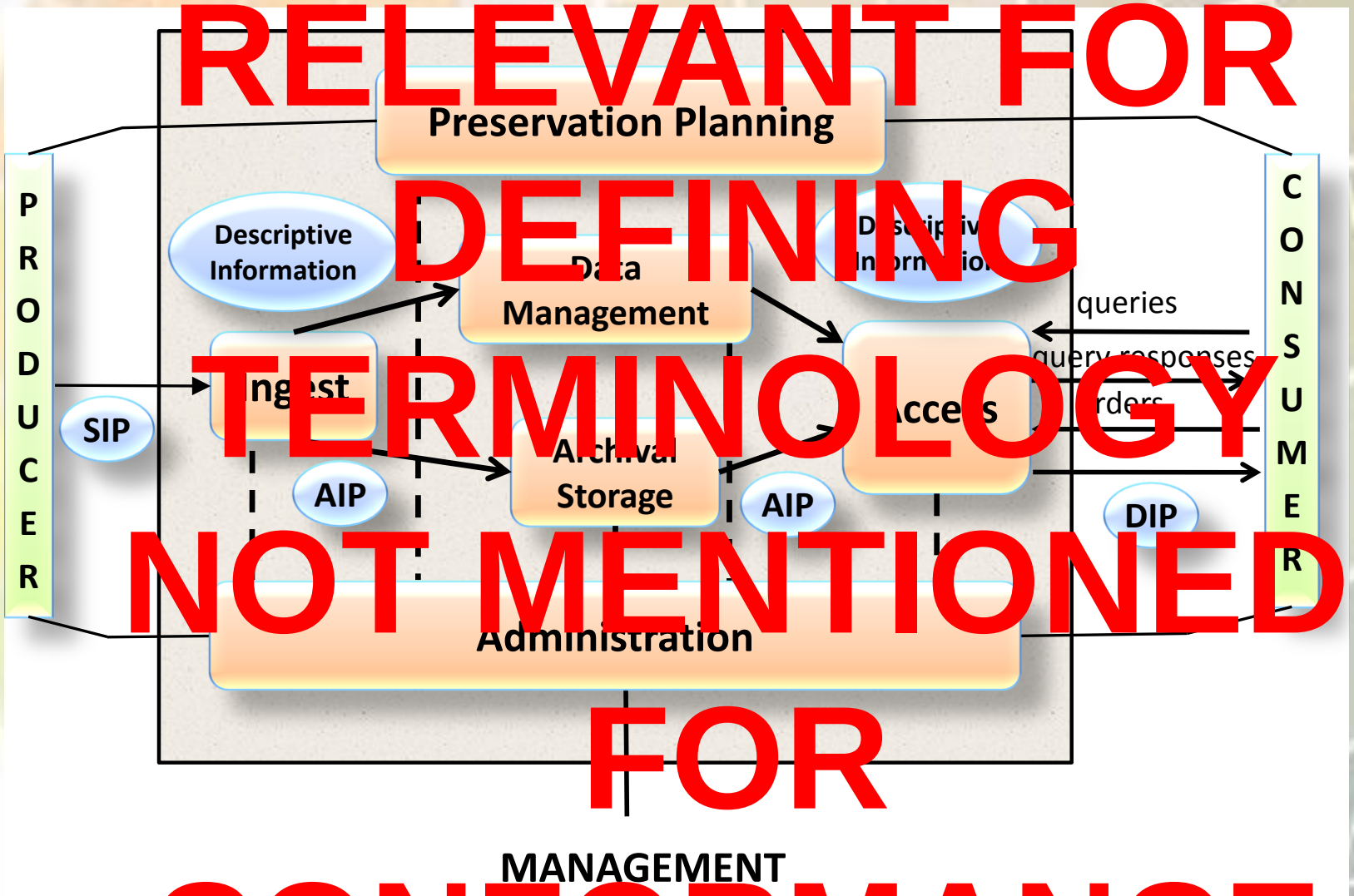
# Why ISO Audit?

- Well established, well accepted international process
- Need for checking compliance to prescribed standards – conformity assessment – inspection/testing/ certification
- Confidence in conformity assessment
- International acceptability for facilitating trade - Need for recognition of inspection/testing/ certification across borders
- Accomplished through accreditation

# OAIS

- Conformance requirements
  - Information model
    - *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.*
  - Mandatory responsibilities
    - *A conforming OAIS Archive shall fulfill the responsibilities listed in 3.1. Subsection 3.2 provides examples of the mechanisms that may be used to discharge the responsibilities identified in 3.1. These mechanisms are not required for conformance. A separate standard, as noted in 1.5, has been produced on which accreditation and certification processes can be built.*
- OAIS does not cover everything e.g. financial aspects

# OAIS Functional Model




RELEVANT FOR DEFINING TERMINOLOGY NOT MENTIONED FOR CONFORMANCE

# OAIS Information model: Representation Information

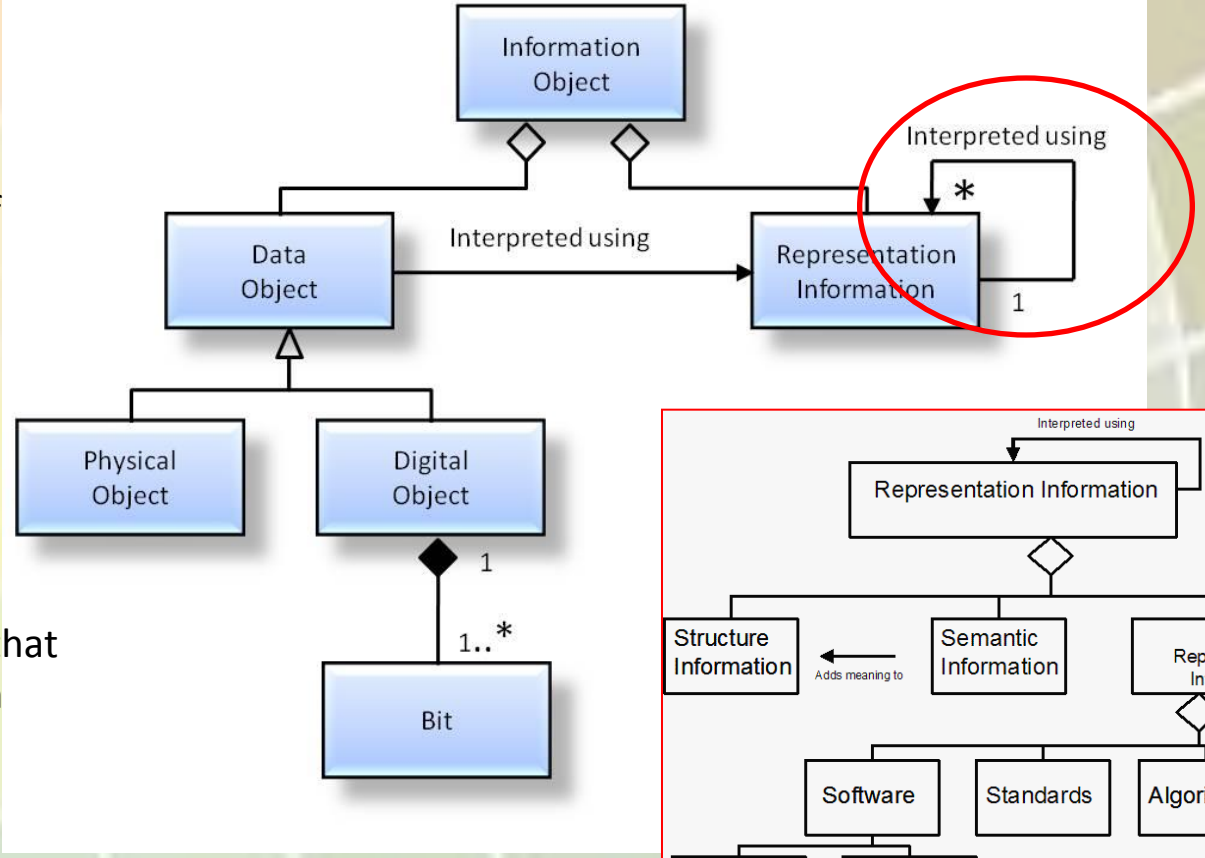
## The Information Model is key

Recursion ends at KNOWLEDGEBASE of the DESIGNATED COMMUNITY

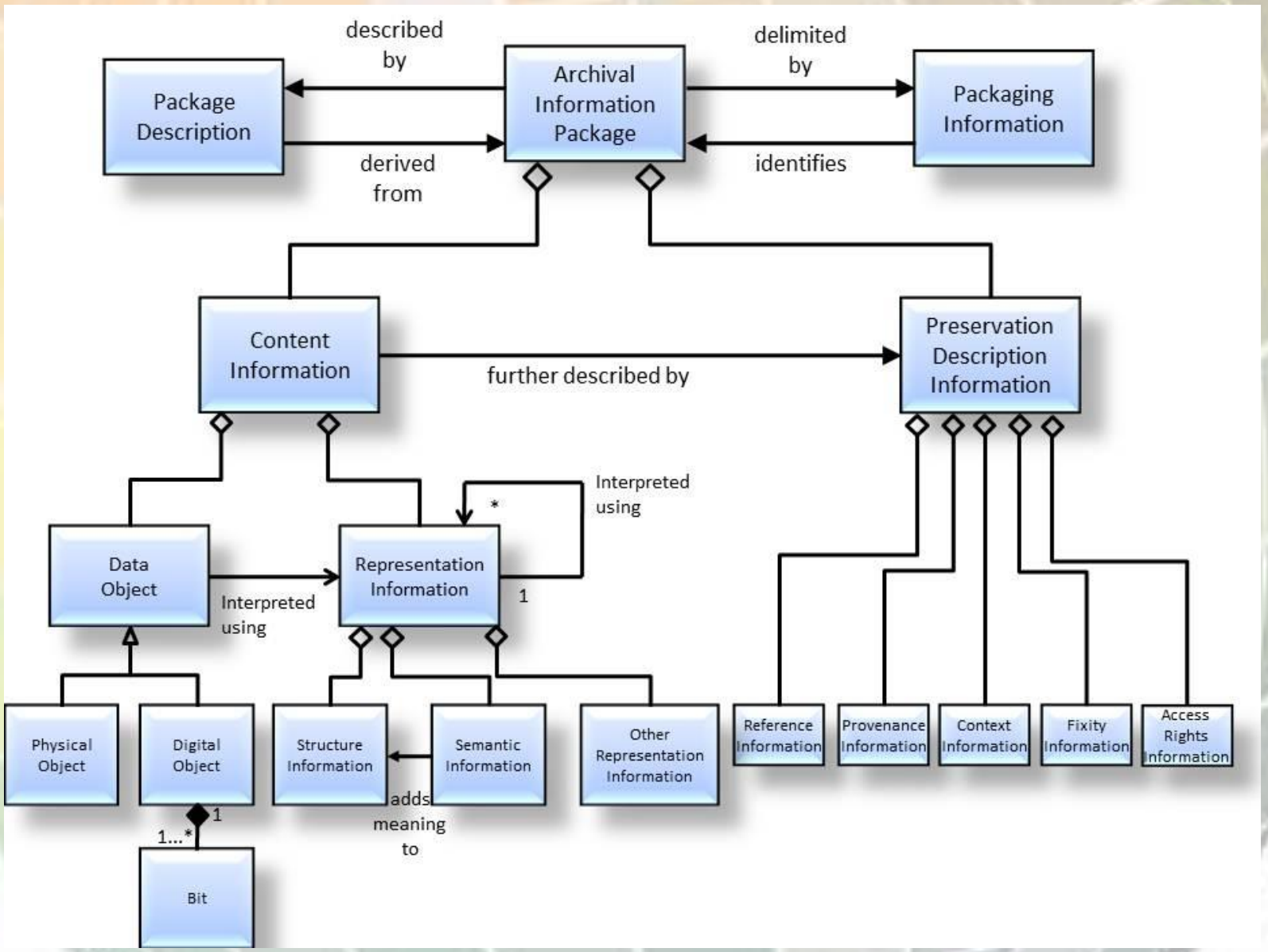
(this knowledge will change over time and region)

Does not demand that ALL Representation Information be collected at once.

A process which can be tested



# OAIS Information Model



# Mandatory responsibilities (1)

- 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, thereby defining its Knowledge Base.

# Mandatory responsibilities (2)

- Ensure that the information to be preserved is Independently Understandable to the Designated Community. In particular, the Designated Community should be able to understand the information without needing special resources such as 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, including the demise of the Archive, ensuring that it is never deleted unless allowed as part of an approved strategy. There should be no ad-hoc deletions.
- Make the preserved information available to the Designated Community and enable the information to be disseminated as copies of, or as traceable to, the original submitted Data Objects with evidence supporting its Authenticity.

# What OAIS does not cover

- standard(s) for the interfaces between OAIS type Archives;
- standard(s) for the submission (ingest) methodology used by an Archive:
  - ISO 20652:2006 Space data and information transfer systems—Producer-Archive Interface—Methodology Abstract Standard (the more specific Producer-Archive Interface Specification (PAIS) is under preparation);
- standard(s) for the submission (ingest) of digital data sources to the Archive;
- standard(s) for the delivery of digital sources from the Archive;
- standard(s) for the submission of digital metadata, about digital or physical data sources, to the Archive:
  - PREMIS Data Dictionary for Preservation Metadata. Version 2.0, PREMIS Editorial Committee, March 2008;
  - 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 21962:2003 Space data and information transfer systems—Data Entity Dictionary Specification Language (DEDSL)—PVL syntax;
  - ISO 22643:2003 Space data and information transfer systems—Data Entity Dictionary Specification Language (DEDSL)—XML/DTD;
  - ISO 13527:2010 Space data and information transfer systems—XML formatted data unit (XFDU) structure and construction rules;
- syntax standard(s) for the identification of digital sources within the Archive;
- protocol standard(s) to search and retrieve metadata information about digital and physical data sources;
- standard(s) for media access allowing replacement of media management systems without having to rewrite the media;
- standard(s) for specific physical media;
- standard(s) for the migration of information across media and formats;
- standard(s) for recommended archival practices:
  - ISO 15489-1:2001 Information and documentation—Records management. Part 1: General;
  - ISO/TR 15489-2:2001 Information and documentation—Records management. Part 2: Guidelines;
  - ISO 23081-1:2006 Information and documentation—Records management processes—Metadata for records—Part 1: Principles;
  - ISO/TS 23081-2:2007 Information and documentation—Records management processes—Metadata for records—Part 2: Conceptual and implementation issues;
- standard(s) for certification of Archives:
  - CCSDS 652.0-M-1, Audit and Certification of Trustworthy Digital Repositories (Magenta Book, Issue 1), also available as ISO 16363:2011

....

- Specific hardware/ software
- Funding
- Training
- Organisational commitment

# Standards based Repository Audit and Certification (ISO 16363)

**OAIS (ISO 14721)**

Trusted Digital Repositories: Attributes and Responsibilities

TRAC

**Requirements For Bodies Providing Audit And Certification (ISO 16919 )**

**There is a hierarchy of ISO standards concerned with good auditing.**
**ISO 16919 is positioned within this hierarchy in order to ensure that these good practices can be applied to the evaluation of the trustworthiness of digital repositories using ISO 16363.**
**It covers principles needed to inspire confidence that third party certification of the management of the digital repository has been performed with impartiality, competence, responsibility, openness, confidentiality, and responsiveness to complaints**

**Audit and Certification of Trustworthy Digital Repositories (ISO 16363 )**

Audit by external, **accredited**, auditors

**Certification**

**Metrics concerning:**

- **Organizational Infrastructure**
  - **e.g. The repository shall have a documented history of the changes to its operations, procedures, software, and hardware.**
- **Digital Object Management**
  - **e.g. The repository shall have access to necessary tools and resources to provide authoritative Representation Information for all of the digital objects it contains.**
- **Infrastructure and Security Risk Management**
  - **eg. The repository shall have procedures in place to evaluate when changes are needed to current software.**

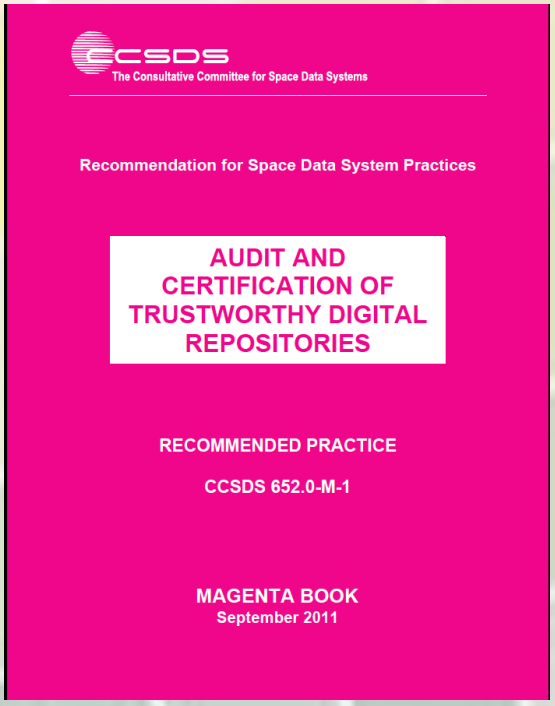
See http://wiki.digitalrepositoryauditandcertification.org and
http://www.alliancepermanentaccess.org/membership/member-resources/audit-and-certification
Standards will be available free from http://www.ccsds.org

# ISO 16363 (25 mins)

- Overall Structure:
  - **Section A: Organisational Infrastructure**
  - **Section B: Digital Object Management**
  - **Section C: Infrastructure and Security Risk Management**
- Metrics and their structure:
  - Statement of requirement
  - Supporting text
  - Examples of Ways the Repository can Demonstrate it is Meeting this Requirement
  - Discussion

# Get Your Copy for Free!

- You can download the Magenta Book for ISO 16363 here: http://public.ccsds.org/publications/archive/652x0m1.pdf



| | Metric |
|---|---|
| 4.2.1 | **THE REPOSITORY SHALL HAVE FOR EACH AIP OR CLASS OF AIPS PRESERVED BY THE REPOSITORY AN ASSOCIATED DEFINITION THAT IS ADEQUATE FOR PARSING THE AIP AND FIT FOR LONG-TERM PRESERVATION NEEDS.** |
| 4.2.1.1 | *The repository shall be able to identify which definition applies to which AIP.* |
| 4.2.1.2 | *The repository shall have a definition of each AIP that is adequate for long term preservation, enabling the identification and parsing of all the required components within that AIP.* |
| 4.2.2 | **THE REPOSITORY SHALL HAVE A DESCRIPTION OF HOW AIPS ARE CONSTRUCTED FROM SIPS.** |
| 4.2.3 | **THE REPOSITORY SHALL DOCUMENT THE FINAL DISPOSITION OF ALL SIPS** |
| 4.2.3.1 | *The repository shall follow documented procedures if a SIP is not incorporated into an AIP or discarded and shall indicate why the SIP was not incorporated or discarded.* |
| 4.2.4 | **THE REPOSITORY SHALL HAVE AND USE A CONVENTION THAT GENERATES PERSISTENT, UNIQUE IDENTIFIERS FOR ALL AIPS** |
| Note: In particular the following aspects must be checked: | |
| 4.2.4.1 | *The repository shall uniquely identify each AIP within the repository.* |
| 4.2.4.1.1 | *The repository shall have unique identifiers.* |
| 4.2.4.1.2 | *The repository shall assign and maintain persistent identifiers of the AIP and its components so as to be unique within the context of the repository.* |
| 4.2.4.1.3 | *Documentation shall describe any processes used for changes to such identifiers.* |
| 4.2.4.1.4 | *The repository shall be able to provide a complete list of all such identifiers and do spot checks for duplications.* |
| 4.2.4.1.5 | *The system of identifiers shall be adequate to fit the repository's current and foreseeable future requirements such as numbers of objects.* |
| 4.2.4.2 | *The repository shall have a system of reliable linking/resolution services in order to find the uniquely identified object, regardless of its physical location.* |

# Metrics: too many or too few?

- Impossible to anticipate all possibilities
- Should be regarded as a “guide” for auditors
- Fundamentally depends on auditors’ experience/judgement
- Need to try to guarantee consistency of judgements

# ISO Process for Audits

- Comprehensive view of repository's capabilities
- Preparatory work by repository
- First audit and resulting certification
  - Identifies improvements needed
  - Repository prepares and implements improvement plan
- Surveillance audit after 18 months
- Re-certification

# The audits

- Self-audit process covering all metrics, providing information to audit team:
- The basics:
  - Are the bits safe?
  - Are the data understandable/usable by the **Designated Community**?
  - Is **authenticity** safeguarded (evidence based)
    - E.g. Is the information really what it is claimed to be?
  - Can the digital holdings be handed over to another repository if/when necessary?
- The repository must try to provide evidence
  - Why do they think people (including their funders) should trust them?

# What would Certification look like?

- **NOTE: The audit and certification would be undertaken by an accredited organisation**
- Not a simple statement that "Yes this repository is perfect"!
- Should be regarded as part of a process of **<u>improvement</u>**
  - Audit/certification provides information on which an organization can act to improve its performance
  - Improvement plan
    - "repository OK as long as …."
  - Cycle of certification/ surveillance audit/ re-certification

# Example Issues from ISO 16363 Metrics

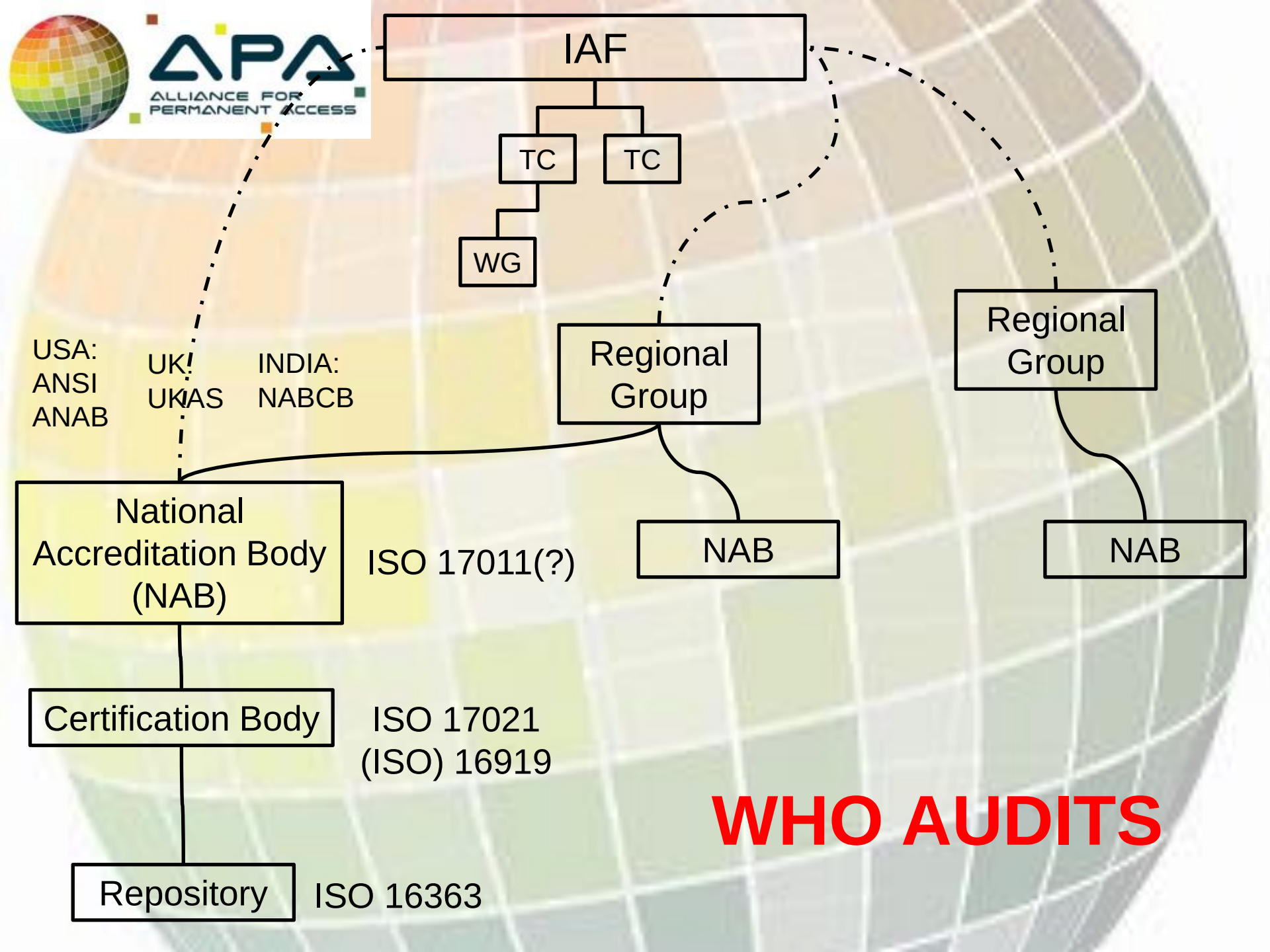
- Metric 3.3.1 THE REPOSITORY SHALL HAVE DEFINED ITS DESIGNATED COMMUNITY AND ASSOCIATED KNOWLEDGE BASE(S) AND SHALL HAVE THESE DEFINITIONS APPROPRIATELY ACCESSIBLE.
  - *The metric was not satisfied. There were no records or documentary evidence presented to describe any Designated Community, nor any AIP associations with any Designated Community.*

# Example Issues from ISO 16363 Metrics

- Metric 4.2.5.2 THE REPOSITORY SHALL HAVE TOOLS OR METHODS TO DETERMINE WHAT REPRESENTATION INFORMATION IS NECESSARY TO MAKE EACH DATA OBJECT UNDERSTANDABLE TO THE DESIGNATED COMMUNITY.
  - *This metric was not satisfied. In our discussions, we perceived that <SITE>s view of their designated communities was ambiguous. Consequently, there were few indications that representation information was being organized to support the long term use and understanding of the AIPs. While a great deal of representation information, and indeed, metadata in general is being collected, there appear to be few opportunities within <SITE> for explicit review of those ancillary data in light of changes in long-term preservation needs of one or more Designated Communities.*

# Example Issues from ISO 16363 Metrics

- METRIC 3.1.2 THE REPOSITORY SHALL HAVE A PRESERVATION STRATEGIC PLAN THAT DEFINES THE APPROACH THE REPOSITORY WILL TAKE IN THE LONG-TERM SUPPORT OF ITS MISSION
  - *Currently there is no preservation strategic plan separate from the overall strategic plan of the organization. However, within the new strategic plan 2011-2015 attention is given to preservation aspects. This Strategic Plan will be published in June. The metric was not satisfied.*


IAF
TC
TC
WG
Regional Group
Regional Group
USA:
ANSI
ANAB
UK:
UKAS
INDIA:
NABCB
National Accreditation Body (NAB)
ISO 17011(?)
NAB
NAB
Certification Body
ISO 17021
(ISO) 16919
Repository
ISO 16363
WHO AUDITS

# Auditors - new CASCO approach

- Instead of qualifications and particular experience now the focus is on **COMPETENCIES**

# Example of Competencies

- *Possesses the knowledge to assess the TDRMS' procedures and processes when creating Archival Information Packages (AIPs),and its ability to :*
  - *assess the level of detail to which an AIP should be described*
  - *determine the functions of the various components of an AIP and how they may be implemented*
  - *identify the range of provenance information that should be collected*
  - *identify the difference between SIP and AIP and ways in which the former may be converted to the latter*
  - *identify and assess workflows and whether they reliably achieve what they purport to do*
  - *assess the relationship between the various identifiers used within a repository*
  - *assess ways of defining Designated Communities and how the appropriate amount of Representation Information may be obtained*
  - *identify (or assess) possible changes in the Designated Community and its knowledge base and impacts on understandability*
  - *assess the risks to the integrity of digital holdings in various circumstances - both technical and non-technical.*

# Get Your Copy for Free!

- Standard will be available on http://public.ccsds.org/publications/MagentaBooks.aspx
  - NOTE: Issue 1. November 2011 is the OLD version
  - The revised version will be available in a few days
- Self-assessment template is available on the PTAB website http://www.iso16363.org/preparing-for-an-audit/

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | 3. ORGANIZATIONAL INFRASTRUCTURE | | | | | | |
| 2 | 3.1 GOVERNANCE & ORGANIZATIONAL VIABILITY | | | | | | |
| 3 | | | **Metric** | **Supporting Text** | **Examples of Documents the Repository can use to demonstrate it is Meeting this Requirement:** | **Brief description of evidence (add rows if necessary to list all relevant documents for a metric) Use short titles for documents. Provide detailed description of each document on the Reference tab.** | **Explanation of how the repository addresses this metric** |
| 4 | | 3.1.1 | THE REPOSITORY SHALL HAVE A MISSION STATEMENT THAT REFLECTS A COMMITMENT TO THE PRESERVATION OF, LONG TERM RETENTION OF, MANAGEMENT OF, AND ACCESS TO DIGITAL INFORMATION. | This is necessary in order to ensure commitment to preservation and access at the repository's highest administrative level. | Mission statement or charter of the repository or its parent organization that specifically addresses or implicitly calls for the preservation of information and/or other resources under its purview; a legal, statutory, or government regulatory mandate applicable to the repository that specifically addresses or implicitly requires the preservation of information and/or other resources under its purview. | | |
| 5 | | 3.1.2 | THE REPOSITORY SHALL HAVE A PRESERVATION STRATEGIC PLAN THAT DEFINES THE APPROACH THE REPOSITORY WILL TAKE IN THE LONG-TERM SUPPORT OF ITS MISSION | This is necessary in order to help the repository make administrative decisions, shape policies and allocate resources in order to successfully preserve its holdings. | Preservation Strategic Plan; meeting minutes; documentation of administrative decisions which have been made. | | |
| | | 3.1.2.1 | *The repository shall have an appropriate, formal succession plan, contingency plans, and/or escrow arrangements in* | This is necessary in order to preserve the information content entrusted to the repository by handing it on to another | 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 | | |

# Steps to take

- Self-assessment template is available on the PTAB website www.iso16363.org
- Begin gathering (or creating!) the documentation you will need for certification and audit
- Find accredited auditor
  - Should be available soon!!
- Agree scope of audit

# Self-assessments

- Conduct self-assessments to identify shortcomings and help identify any “surprises.”
- Determine which metrics apply to the repository. If the repository believes that any metrics do not apply, document why the metric does not apply.
- Determine which metrics are being met successfully and provide examples to document how the repository meets each metric.
- Determine which metrics are not being met and what measures need to be implemented to achieve and document success for that metric.
- Determine what resources (additional or reallocated) are required to achieve success.
- Refer to external resources, such as case studies and best practices, as desired and available.
- Incorporate findings from previous audits conducted of the repository or similar institutions if such results are available. These could include information technology security audits, ISO 9000 suite audits, quality assurance audits, risk assessments, and similar evaluations.
- Populate and maintain the Self-Assessment Template for ISO 16363 to better organize and track progress on meeting the metrics and prepare for an ISO 16363 audit when they can be performed.

# Training Process for Repositories Managers

- Handbook for repository being prepared
- Information for managers to understand what the repository staff is attempting to accomplish
- Training for staff to understand what is required
- Helps in preparation for audit and self-audit

# Links

- **ISO Audit**
  - http://www.iso16363.org/
    - *http://wiki.digitalrepositoryauditandcertification.org*
- **OAIS Reference Model**
  - *Original version available from http://public.ccsds.org/publications/archive/650x0b1s.pdf*
  - Updated version at *http://public.ccsds.org/publications/archive/650x0m2.pdf*
- ***Alliance for Permanent Access***
  - *http://www.alliancepermanentaccess.org*
  - *Information about SCIDIP-ES and APARSEN at http://www.alliancepermanentaccess.org/index.php/current-projects/ and www.aparsen.eu and www.scidip-es.eu*
  - *Additional OAIS and ISO Audit information will be at http://www.alliancepermanentaccess.org/index.php/membership/member-resources/*