Outline

- Role of Alliance for Permanent Access (APA)
- Overview of challenges
- Key ideas - OAIS
- Development of the ISO standards
- Organisation and process for audit and certification
Role of APA

• APA is a membership organisation for those interested in digital preservation - for any organisation with a significant interest in digital preservation:
  research organisations, funding organisations, major libraries and repositories, museums and archives, publishers in all fields of research, scholarship and technology, or their umbrella organization, national coalitions of organisations engaged in the curation and preservation of digital records, Universities and academic institutions, commercial organisations and vendors, governmental organisations.

  In addition, individuals may become personal members

• Many activities including promotion and setting up e-infrastructure components to assist digital preservation (e.g. APARSEN and SCIDIP-ES)

• ISO Audit and Certification is one part of the strategic plan
Digital Preservation...

• Easy to do...
• ...as long as you can provide money forever
• Easy to test claims about repositories...
• ...as long as you live a long time
What is wanted

• By Funders
  – How to tell – independently – that money is been spent well
    • Otherwise risk money being wasted and data lost
    • Might not know until too late
  – International standard – preferably ISO

• By Repositories:
  – Comfortable
  – Low cost
  – Low trouble
  – Something to confirm they are going a good job
Challenges

• There is not enough experience (by anyone) of long term preservation of massive amounts of data
• How can audit/certification provide any kind of judgement?
Digital Preservation

- Ensure that digitally encoded information are understandable and usable over the long term
  - Things become “unfamiliar” – hardware/software/tacit knowledge
  - Could be important after just a few years
- Ability to re-use unfamiliar digitally encoded information is key requirement
  - Automation desired (formal detailed structural and semantic descriptions) but
  - The perfect should not be the enemy of the “good enough”
- Reference Model for Open Archival Information System (ISO 14721)
  - The basic standard for work in digital preservation
  - Defines terminology and compliance criteria
PROVIDES USEFUL TERMINOLOGY
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
Repository Audit and Certification Working group

• Closely related to OAIS Reference Model
  – Certification was identified as a follow-on standard
  – Following route of OAIS
    • CCSDS is the “working arm” of an ISO SC
  – Initial (closed) work lead by NARA & RLG led to TRAC, based on
    – OAIS
    – Trusted Digital Repositories: Attributes and Responsibilities (RLG-OCLC)
    – Approx 5 iterations

• Always intended to move to CCSDS Working Group for finalisation/improvement/standardisation

• Mostly virtual meetings, notes and documents openly available
Metrics (published as ISO 16363)

• Available from
  – http://www.digitalrepositoryauditandcertification.org
  – http://www.digitalrepositoryauditandcertification.org

• Overall Structure:
  – Section A: Organisational Infrastructure
  – Section B: Digital Object Management
  – Section C: Infrastructure and Security Risk Management

• Metrics and their structure: Aims to make self-audit practical
  – Statement of requirement
  – Supporting text
  – Examples of Ways the Repository can Demonstrate it is Meeting this Requirement
  – Discussion
Metrics: too many or too few?

- Impossible to anticipate all possibilities
- Other standards (e.g. ISO 2700x security standards) are quite brief
- Should be regarded as a “guide” for auditors
  - the areas to focus on
  - sub-metrics pick out more specific areas
- Fundamentally depends on auditors’ experience/judgement
- Need to try to guarantee consistency of judgements
Who does the certification?

• There is a hierarchy of ISO standards concerned with good auditing practices
  – ensure that these good practices can be applied to the evaluation of the trustworthiness of digital repositories using ISO 16363.
  – to inspire confidence in
    • impartiality, competence, responsibility, openness, confidentiality, and responsiveness to complaints
• “Requirements for Bodies providing Audit and Certification” defines how the audit/certification organisation operates
  • published as ISO 16919
• Defines the Primary TDR Authorisation Body (PTAB) and process for certifying auditors and creating (national) accrediting bodies
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.

Standards based Repository Audit and Certification (ISO 16363)

OAIS (ISO 14721)

TRAC

Requirements For Bodies Providing Audit And Certification (ISO 16919)

Audit and Certification of Trustworthy Digital Repositories (ISO 16363)

Audit by external auditors

Monitored self-audit using ISO 16363 (or DIN31644 in Germany)

Basic Certification

Monitor self-audit using DSA metrics

Extended Certification

Formal Certification

EUROPEAN FRAMEWORK FOR AUDIT AND CERTIFICATION OF DIGITAL REPOSITORIES

to be promoted by the EU

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
  - e.g. The repository shall have procedures in place to evaluate when changes are needed to current software.

Standards will be available free from http://www.ccsds.org
Testing the standards

• Test audits were carried out to test
  – the PTAB’s common understanding of the metrics
  – the usability of the metrics document by repositories
  – 3 in Europe (UKDA, CINES, DANS)
  – 3 in USA (NSSDC, SEDAC, KA State Archive)

• Final versions of the standards are now published by ISO (free from CCSDS)

• PTAB formal body created
The audits

• Self-audit process covering all metrics, providing information to audit team:

• Considered all metrics including 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?

• Learning process – over several audit cycles
What would Certification look like?

• Not a simple statement that “Yes this repository is perfect”!

• Should be regarded as part of a process of **improvement**
  – 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

• Maturity levels
Next Steps in process

• Accredit new auditors
  – accredited training courses
  – conduct audits with existing auditors
  – boot-strapped by PTAB
• Set up **national** accreditation boards

We are working on the assumption that there will be a significant demand from public and **private/commercial** organisations/service providers. Process is designed to scale up to meet the need.
A vision

• Continued improvements in digital preservation capabilities
  – Public funded, Private/commercial
• Certification required by people with the money
• Trained, consistent set of auditors, consultants across the globe
  – National accreditation bodies
• Help to build market for vendors, systems builders and service providers
• Society benefits from its digitally encoded capital
• **ISO Audit**
  
  • [http://www.iso16363.org/](http://www.iso16363.org/)
    
    • [http://wiki.digitalrepositoryauditandcertification.org](http://wiki.digitalrepositoryauditandcertification.org)

• **OAIS Reference Model**
  
  – *Original version available from*
    
    [http://public.ccsds.org/publications/archive/650x0b1.pdf](http://public.ccsds.org/publications/archive/650x0b1.pdf)

  – *Updated version will be available from CCSDS*

  – **Alliance for Permanent Access**
    
    – [http://www.alliancepermanentaccess.org](http://www.alliancepermanentaccess.org)
    
    – *Information about SCIDIP-ES and APARSEN at*
      

    – *Additional OAIS and ISO Audit information will be at*
      
Riding the wave

How Europe can gain from the rising tide of scientific data

Final report of the High Level Expert Group on Scientific Data
A submission to the European Commission

October 2010

David Giaretta
Advanced Digital Preservation

There is a growing recognition of the need to address the longevity of digital information, and a wide variety of factors depend on how we preserve all aspects of digital life. This has been discussed in many books and articles on digital preservation, so why is there a need to do more? Because, for the most part, the current practices in science and research, image and video production, the internet and web p---collected digital information has to be simply displayed by software or a human viewer. The real need is to make more types of digital objects that may need to be preserved, such as databases, scientific data and software data.

David Giaretta, Director of the Alliance for Permanent Access, and his team work to explain why this is so and develop innovative solutions for improving this important aspect of how we use digital tools and technologies. The book is an essential reference for those who are interested in this area.

Various examples of digital objects from many sources are used to explain the tools and techniques presented. The practical and technical issues are explained in a way that is accessible to anyone who wants to learn more about digital preservation.

To further aid understanding, the book is supported by many hours of videos and presentations from the CASPAR project and by a set of open source software.

I am very excited by the potential of CASPAR L's technical techniques to ensure sustainable quality of access to digital objects in the future.

Martine Vermeulen, Member of the European Commission