

Chapter 8

Digital Preservation Challenges, Infrastructures and Evaluations

David Giaretta

Science and Technology Facilities Council, UK

ABSTRACT

To preserve digitally encoded information over a long term following the OAI Reference Model requires that the information remains accessible, understandable and usable by a specified Designated Community. These are significant challenges for repositories. It will be argued that infrastructure which is needed to support this preservation must be seen in the context of the broader science data infrastructure which international and national funders seek to put in place. Moreover aspects of the preservation components of this infrastructure must themselves be preservable, resulting in a recursive system which must also be highly adaptable, loosely coupled and asynchronous. Even more difficult is to be able to judge whether any proposal is actually likely to be effective. From the earliest discussions of concerns about the preservability of digital objects there have been calls for some way of judging the quality of digital repositories. In this chapter several interrelated efforts which contribute to solutions for these issues will be outlined. Evidence about the challenges which must be overcome and the consistency of demands across nations, disciplines and organisations will be presented, based on extensive surveys which have been carried out by the PARSE.Insight project (<http://www.parse-insight.eu>). The key points about the revision of the OAI Reference Model which is underway will be provided; OAI provides many of the key concepts which underpin the efforts to judge solutions. In the past few years the Trustworthy Repositories Audit and Certification: Criteria and Checklist (TRAC) document has been produced, as well as a number of related checklists. These efforts provide the background of the international effort (the RAC Working

DOI: 10.4018/978-1-61520-767-1.ch008

Group <http://wiki.digitalrepositoryauditandcertification.org>) to produce a full ISO standard on which an accreditation and certification process can be built. If successful this standard and associated processes will allow funders to have an independent evaluation of the effectiveness of the archives they support and data producers to have a basis for deciding which repository to entrust with their valuable data. It could shape the digital preservation market. The CASPAR project (<http://www.casparpreserves.eu>) is an EU part funded project with total spend of 16MEuros which is trying to faithfully implement almost all aspects of the OAIS Reference Model in particular the Information Model. The latter involves tools for capturing all types of Representation Information (Structure, Semantics and all Other types), and tools for defining the Designated Community. This chapter will describe implementations of tools and infrastructure components to support repositories in their task of long term preservation of digital resources, including the capture and preservation of digital rights management and evidence of authenticity associated with digital objects. In order to justify their existence, most repositories must also support contemporaneous use of contemporary as well as “historical” resources; the authors will show how the same techniques can support both, and hence link to the fuller science data infrastructure.

INTRODUCTION

Much work has been undertaken in the area of digital preservation. It has been said National Science Foundation Cyberinfrastructure Council, (2007) that “the Open Archival Information System OAIS, (2002), now adopted as the ‘de facto’ standard for building digital archives”. The work presented here is firmly based on OAIS.

OAIS REFERENCE MODEL

The OAIS Reference Model provides a number of models for repositories including a Functional Model, to which is relatively easy to map an existing archive system, an Information Model, which is rather more challenging, an Information Packaging Model and federation models, plus preservation perspectives including types of migration and a variety of software related processes. A number of overall strategies, processes and supporting infrastructures may be derived from these.

OAIS Information Model

The Information Model provides the concepts to support the long-term understandability of the preserved data. This introduces the idea of Representation Information.

The UML diagram in Error! Reference source not found. means that

- An Information Object is made up of a Data Object and Representation Information
- A Data Object can be either a Physical Object or a Digital Object. An example of the former is a piece of paper or a rock sample.
- A Digital Object is made up of one or more Bits.
- A Data Object is interpreted using Representation Information

Representation Information is itself interpreted using further Representation Information.

This figure shows that Representation Information may contain references to other Representation Information. When this is coupled with the fact

13 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/digital-preservation-challenges-infrastructures-evaluations/42741

Related Content

Application of Mobile Technologies in Library Service Delivery

Oluwatoyin Oyeyemi Obinyan (2020). *Handbook of Research on Digital Devices for Inclusivity and Engagement in Libraries* (pp. 371-381).

www.irma-international.org/chapter/application-of-mobile-technologies-in-library-service-delivery/234007

A Content-Driven System Architecture for Tackling Automatic Cataloging of Animated Movie Databases

Bogdan Ionescu, Alexandru Marin, Patrick Lambert, Didier Coquinand Constantin Vertan (2012). *Multimedia Storage and Retrieval Innovations for Digital Library Systems* (pp. 101-123).

www.irma-international.org/chapter/content-driven-system-architecture-tackling/64464

Efficient Word Segmentation and Baseline Localization in Handwritten Documents Using Isothetic Covers

Mousumi Dutt, Aishariya Sarkar, Arindam Biswas, Partha Bhowmickand Bhargab B. Bhattacharya (2011). *International Journal of Digital Library Systems* (pp. 1-13).

www.irma-international.org/article/efficient-word-segmentation-baseline-localization/59883

Change Management in the Academic Library: Transition From Print to Digital Collections

Oyeronke Adebayo, Michael Opeoluwa Fagbohun, Ugwunwa Chinyere Esseand Nwanne Mary Nwokeoma (2018). *Digitizing the Modern Library and the Transition From Print to Electronic* (pp. 1-28).

www.irma-international.org/chapter/change-management-in-the-academic-library/188021

Effective and Efficient Browsing of Large Image Databases

Gerald Schaefer (2009). *Handbook of Research on Digital Libraries: Design, Development, and Impact* (pp. 142-148).

www.irma-international.org/chapter/effective-efficient-browsing-large-image/19877