Chapter 5
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 OAIS 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 OAIS Reference Model which is underway will be provided; OAIS 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 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 faith

fully 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.

1 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.

2 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.

2.1 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 Figure 1 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

Figure 1 shows that Representation Information may contain references to other Representation Information. When this is coupled with the fact that Representation Information is an Information Object that may have its own Digital Object and

Figure 1. OAIS information model
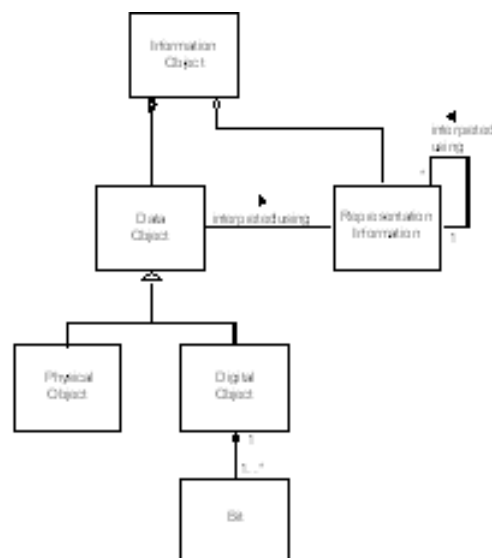
11 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/digital-preservation-challenges-infrastructures-evaluations/70972

Related Content

An Eastern Approach to the Global Challenge of Corruptibility

Dennis P. Heatonand Ravi Subramaniam (2012). *Ethical Models and Applications of Globalization: Cultural, Socio-Political and Economic Perspectives (pp. 89-99).*

www.irma-international.org/chapter/eastern-approach-global-challenge-corruptibility/60422

Reviewing the Ethics and Philosophy Behind Social Media's Crowdsourced Panopticon

Amanda Furiasse (2022). *International Journal of Technoethics (pp. 1-4).*

www.irma-international.org/article/reviewing-the-ethics-and-philosophy-behind-social-medias-crowdsourced-panopticon/302627

A Technoethical Approach to the Race Problem in Anthropology

Michael S. Billinger (2009). *Handbook of Research on Technoethics (pp. 44-68).*

www.irma-international.org/chapter/technoethical-approach-race-problem-anthropology/21571

Web Analytics and Online Retail: Ethical Perspective

Gabriel Ayodeji Ogunmolaand Vikas Kumar (2020). *International Journal of Technoethics (pp. 18-33).*

www.irma-international.org/article/web-analytics-and-online-retail/258967

Fair Use, Copyright, and Academic Integrity in an Online Academic Environment

Kristine Holloway (2013). *Digital Rights Management: Concepts, Methodologies, Tools, and Applications (pp. 917-928).*

www.irma-international.org/chapter/fair-use-copyright-academic-integrity/71011