briefing paper

Digital Preservation and Open Access Archives
Persistent access to open access digital assets

Open Access repositories promote the widespread dissemination of scientific and scholarly production. Researchers and teachers publish free on line digital assets for claiming their activity and for sharing research results with other researchers. In particular universities, research centres, libraries and, for limited subsets of their collections, museums, administrative archives and other cultural institutions are promoting open access.

In the future, a considerable section of scholarly, academic and cultural institutions memory will be formed by born–digital assets, stored in open access archives. Their digital collections will have an ever growing relevance in making up the scientific and information heritage of the next generations. In order to ensure that these objects will survive and continue to be cited, scholarly and academic communities should be committed to the long term preservation of their repositories.

Context and objectives

According to JISC digital preservation is "the series of actions and interventions required to ensure continued and reliable access to authentic digital objects for as long as they are deemed to be of value". Three key concepts can be underlined in the context of Open Access archives
• OA repositories store digital objects that are prevalently "born-digital" assets without an equivalent analogue original.
• OA assets reliability and availability are relevant for their authors, the scholarly and scientific community.
• OA assets continued access - that is the ongoing usability of a digital asset , the possibility to cite, discover, deliver and use that resource after its creation and deposit into the archive.

Access must be ensured as long as it is necessary by preventing repository or media failure and controlling the effects of technological change. Preservation processes should rely on well defined policies, organisation and strategies.

Models

The OAIS Reference Model defines an Open Archival Information Systems (OAIS) as "an archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community". This model provides a complete functional and information specification of a repository and "establishes mandatory responsibilities that an organization must discharge in order to operate an OAIS archive".

The application of OAIS principles and in particular the implementation of Open Access archives in accordance with OAIS functional and information models is a key prerequisite to establish trusted repositories and to ensure long time preservation of their digital assets.

Properties

Digital preservation can ensure authors and users trust in Open Access digital repositories. A trusted repository must maintain the following properties as a minimum:

• **Authenticity**. The certainty that a digital asset was created by the individual who claims to have created it. Authenticity provides the certainty that the creator of a digital asset cannot deny they are the creator. Digital signatures and digital watermarking are techniques that guarantee the authenticity of digital objects.

• **Integrity**. The ability to maintain data correctness and completeness by preventing accidental or malicious changes (i.e. data corruption). Deriving and saving a bit/byte checksum, such has MD_5, is a basic technique to detect if any modification has affected a digital object after its initial deposit in the archive.

• **Reliability and Availability**. Reliability refers to the ability of hardware and software components to perform according to its specification free of errors or bugs (in theory totally, in practice ensuring a high percentage). Availability is the percentage of time a system, application or component is regularly in function respect to the total time it is required to function. Backups, antivirus software, firewalls, operating system patches, application software upgrades, redundant and fault tolerant hardware components are some of the more common techniques used to guarantee high level percentages of reliability and availability.

• **Capacity for reuse**. The ability to access a digital resource for as long as the holding archive/institution decides to support it. Scholarly and scientific digital assets with lasting value must be properly retrieved and (re)used, even for a long time (e.g. by persistent identifiers and maintaining media and formats).

Digital preservation must include all the data components of open access digital repositories infrastructure, not only digital objects but also their metadata, identifiers and resolvers metadata.

Policies and strategies

• Reliable citation systems based on persistent and unique identifier: trusting in the integrity and authenticity of cited digital assets, ensured by long-term preservation strategies, and avoiding discontinued links (i.e. the well known "404 Not Found" error) by technical and organisational solutions that go beyond the limits of url based citations.

• Organizational policy statements must not only select the digital assets to be preserved and to ensure financial and organisational support for this, but also identify who is responsible for digital preservation. Roles and responsibilities must be accurately defined, expertise must be carefully supported and training has to be constantly promoted.

• Digital preservation of a multidisciplinary knowledge must be grounded on the close cooperation of professionals and experts belonging to different areas.

• Increasing the authors' awareness about following well defined preservation policies and strategies for guaranteeing the access continuity to their digital assets.

• Metadata ensures continued access to digital assets and their availability: digital preservation planning must include strategies and processes for metadata preservation too.

• Constant auditing and assessment efforts to account for changes in organizational, policy, legal and technical components.

Valdo Pasqui *University of Florence*
valdo.pasqui@unifi.it