

Trusted Digital Archives requirements for full text scholarly publications

One or more archives might be necessary to respond to the needs listed below (such as, only by way of example, an integration between NERA and LOCKSS). For simplicity's sake we shall speak of archives in the singular.

Needs analysis by priority.

- 1) The Archive must ensure proper performance – suitable bandwidth, connection speed, a stable and reliable software, completeness and promptness when loading content - in order to provide full-text access to documents Italian universities have signed perpetual access agreements for with the relevant copyright holders. Archive must provide access compliant with those agreements. Information retrieval must take place in each of the following ways.
 - i. Browsing the Archive web site.
 1. A to Z list for titles, both by contents in their entirety and by limiting the list by publisher and/or type of documents (journals; books);
 2. As to each document, browsing must work:
 - a. For each journal by year/volume and then issue and page
 - b. For each book, by chapter when publisher has provided for the chapter division of the digital book.
 - ii. Research by index on Archive web site, according to fields such as for example:
 1. Journals: Publisher, Journal title, ISSN-print; ISSN-electronic; ISSN-L; Year; Volume; Issue.
 2. Books: Publisher, Book title; Author; Year of publication; ISBN-print; ISBN-electronic.
 - iii. Using link resolver. Archive must be set up as target for the most popular Link resolvers used by libraries that are compliant with OpenURL standard.
 - iv. Integration with OPAC and similar systems. Archive must enable metadata export using KBART schemas. "title_url" values must be of the type "Persistent URL".
- 2) As to preservation of scholarly publications published in Italy, the Archive must ensure proper integration with the national initiative for the Digital Legal Deposit (DLD), named "Magazzini Digitali", at National Central Library Florence (BNCF), to comply with Italian law on DLD. Therefore the Archive must set up a proper automated data exchange system to become a contributor of "Magazzini Digitali", with low expenses and with minimum effort (only one workflow) for Italian publishers that wish to be compliant with Long Term Preservation-Perpetual Access-DLD.
- 3) The Archive must allow metadata and content sharing with systems for Text and data Mining (TDM) activities.

If not already certified, the Archive must plan to gain the trustworthy certification within 5 years from the beginning of its activity.

Access to the Archive must take place under the following circumstances.

- 1) To provide access to the preserved full-text content when:
 - a. Publisher is no longer in the condition to provide access on its platform.
 - b. Publishers' servers are temporarily not available.
 - c. For every content (journals, books, etc.) that is not in the front list:
 - i. Journals ceased;
 - ii. Journals transferred to other publisher;
 - iii. Journals that have changed titles (access will be to the content published using the title no longer in use)
 - iv. other.
- 2) To provide access for TDM activities according to the agreements signed with the copyright holders.

Authorized users of the Archive.

Authorized users that belong to the Italian Institutions that have bought perpetual access rights according to the agreements signed with the copyright holders.

As to “Magazzini Digitali”, authorized users are Italian citizens but only for the scholarly content published in Italy, according to Italian laws and to the agreements with Italian publishers when applicable. Content for DLD are accessible only at the premises established by the BNCf as specified by the Italian Estate rules and by the agreements with the publishers when applicable.

Gruppo 404, November 2014