

The logo for Gruppo 404.it is displayed within a dark blue rectangular box. The text "Gruppo" is in a white, sans-serif font, while "404.it" is in a larger, white, stylized font with a slight shadow effect. Below the box, there is a faint, light blue shadow.

Gruppo 404.it

**Italy and the Trustworthy Digital Archives for
scholarly full text publications. Final Report.
English Summary**

Gruppo 404

June 30 2015

Table of Contents

Executive Summary 2

List of Recommendations..... 4

Executive Summary

This report contains the results of the work done in twelve months, from July 2014 to June 2015, by the working group called "Group 404". Group 404 was set up in July 2014 with the aim to develop a shared vision on the initiatives of digital archive of scientific publications (both toll and open access) to be undertaken in Italy for the benefit of the Italian academic and research community.

The group, consisting of specialists, librarians and computer scientists with long experience in the field of perpetual access and long-term preservation of electronic information resources, has drawn up this final report consisting of two parts. The first part outlines the strategic vision and issues the first four recommendations. The second part focuses on the analysis of the needs, on national infrastructures and services both in Italy and abroad (France, Great Britain, United States, and Brazil) and brings about five subsequent recommendations.

The study started from the need to outline a shared vision. As a matter of fact, only a shared strategic vision and a collaborative and active participation of the actors involved (universities, research centers, ministries, national libraries) make possible to undertake sustainable initiatives and concrete actions that will satisfy the needs of the Italian scientific and research community regarding perpetual access (PCA, Post Cancellation access), digital preservation (DP) and digital legal deposit (DLD) of scientific publications. Once a shared view is reached and intentions and a clear definition of needs, priorities and objectives are outlined, one can define recommendations on actions to be taken. The report analyzed current initiatives in Italy: NERA Archive at CINECA, two installations of Global LOCKSS Network, CLOCKSS archive, Magazzini Digitali at BNCf, Portico archive. This work generated the initial recommendations: the first one stresses the need for a more effective involvement of the governance and management of the Italian academic and research system, in particular the relevant ministries (MIUR, i.e. Ministry of Education, University and Research and MIBACT, i.e. Ministry of Cultural Heritage and Tourism), the CRUI, each university, the major scientific research centers. The second recommendation focuses on the need for the national electronic resources license model to require specific clauses to substantiate the strategic vision on perpetual access regarding Toll Access publications. The third recommendation concerns open access publications, that is the need to adopt appropriate solutions in terms of DP, safeguarding authors as well. The fourth recommendation faces the sustainability and cost reduction within the DP, the PCA and the DLD both for publishers and libraries. This can be achieved by arranging agreements and automatic mechanisms in advance, in order to enable the Italian National Library in Florence to harvest both open access (OA) and Toll access (TA) publications for DLD directly from the archives where Italian publishers preserve contents for DP and / or for PCA (e.g. Portico, CLOCKSS, etc.).

During its meetings, the group has carried out an in depth analysis of the requirements within the Italian context, giving them a priority order. Perpetual access has been identified as the highest priority need and consequently it was deemed appropriate to investigate on the main National Hosting initiatives worldwide, either established or still under construction.

The second part of the report focuses on these issues. Therefore, personal meetings with representatives of some of these initiatives were organized, in particular with our French colleagues of Couperin on the ISTEX project and with the Executive Director and the Chief Scientist of LOCKSS Program. In order to broaden the view and get in touch with other experiences a questionnaire was drawn and was sent to those archives offering National Hosting (ISTEX in France, Scholars Portal in Canada, EJC and EBC of

OhioLINK consortium for Ohio), to interesting pilot projects such as testing SafeNet in UK, to digital archives such Portico, the Cariniana network in Brazil, the Global LOCKSS Network program.

According to the responses received and in some cases to the correspondence with the heads of these services, each initiative is presented in highlights (technology used, model of governance and financing etc.). It is also compared with the strategic vision and the needs outlined by the Group 404 and ends with brief final considerations reflecting each initiative as a whole, bearing in mind how as such a solution could apply to the Italian context and/or what possible forms of cooperation could be established with these realities.

Following the analysis of the Italian and international context we have formulated five recommendations in addition to the four ones of the first part.

Therefore, it is recommended to check the possible technological and organizational evolution of NERA archive quantifying the costs to test its ability to respond to the quality requirements of the service, integration with other services operated by universities (Recommendation 5); to undertake PLN LOCKSS with six universities that ensure proper guarantees and availability and work in coordination with CRUI as reference for national negotiations (Recommendation 6); to maintain relationships with other national and international archives in order to exchange experiences and best practices (Recommendation 7); to secure the necessary funds to manage National Hosting, including grant programs applications (from public or private organizations) at national and / or European and / or international level (Recommendation 8); to continue to offer to Italian universities and research institutes a national agreement negotiated by CRUI-CARE to use Portico's services as additional guarantee and not in lieu of National Hosting and negotiate through CRUI-CARE a national agreement for a partnership in the CLOCKSS Archives both to strengthen the presence of the Italian community (libraries and publishers) in this archive, and to bring possible integration with other Italian digital archives such as the National Central Library of Florence for the Digital Legal Deposit (Recommendation 9).

The study ends with the hope that the work done by this group could be used by CRUI-CARE and carried out by setting up an ad hoc team of specialists and launching an experimental project lasting approximately twenty-four months (2016- 2017). Such a project and its testing will put the Italian academic community in a position to decide the final organizational and financing structure to be taken for an Italian sustainable initiative of National Hosting.

List of Recommendations

Recommendation No. 1: The implementation of a strategic vision requires an increasing and more effective involvement of those who, in different ways, participate in the government of the Italian academic and research system, in particular the relevant Ministries (MIUR, i.e. Ministry of Education, University and Research and MIBACT, i.e. Ministry of Cultural Heritage and Tourism), CRUI and every single University (especially the largest ones), the major National Research Centers.

Recommendation No. 2: The strategic vision should bring about proper contract clauses for Toll Access resources, for which PA clauses have been negotiated. In particular, they should be included in the licensing model drawn up by the national group CARE, within the CRUI, which is currently following national negotiations in Italy.

Clarification. Attachment No. 1 (in Italian) shows the clauses that have to be included in the national contracts, in order to translate the strategic vision into agreements with the publishers of academic journals in an adequate way.

Recommendation No. 3: Open Access publications should adopt appropriate solutions in terms of long-term preservation (DP), in order to protect authors as well.

Clarification. Attachment No. 2 (in Italian) shows the elements in detail that originate from this recommendation.

Recommendation No. 4: The activities that Italian libraries and publishers have to implement to ensure long-term preservation (DP), perpetual access (PCA) and digital legal deposit (DLD) should be made less expensive and more sustainable, thanks to agreements between the Italian National Library in Florence and archives (e.g. CLOCKSS, PROTICO, etc.) for automatic mechanisms allowing harvesting both for Open Access and Toll Access publications to populate DLD.

Recommendation No. 5: Check and settle the possible evolution of NERA Archive quantifying its costs in order to test its ability to adequately respond to the actual needs of the service quality, integration with other services operated by universities and its economic sustainability.

Clarification. From the technological point of view, the readiness of Scholars Portal, Portico, ISTEEX, Global LOCKSS Network, attained during the work of the Group 404, could be the starting point to re-engineer the archive. The collaboration with Scholars Portal could be particularly significant because it could help not only with technological choices and production organization, but also to recover contents, because Scholars Portal experienced a genesis similar to NERA - it migrated from ScienceServer.

Timing. Present a proposal within 15 months from the start of the audit.

Recommendation No. 6: Start the experience of a PLN LOCKSS with six Italian universities ensuring sufficient availability and guarantees. This step should possibly take place in coordination with CRUI as reference point for national negotiations.

Clarification. Each of the six Italian universities will provide:

a) the hardware to be used in its network;

b) the staff (a librarian, a systems architect, a developer, all part-time for a total estimated of 1 FT).

If CRUI deems it necessary, it will provide a budget for the needs related to the initiative.

The universities could be selected preferably among those that are already in contact with LOCKSS program staff at Stanford University and that have already joined CLOCKSS archive which uses the same technology.

Timing. Within 12 months from the start of the activities, an alternative perpetual access should be implemented other than the access on the publisher server for at least one major foreign publisher, that doesn't transfer its data to NERA but in GLN (e.g. one among Taylor and Francis, Sage, OUP or similar) or a publisher like Springer (which transfers its data to NERA and joins GLN as well) and for at least one Italian publisher (for example one among Il Mulino, Franco Angeli, Vita e Pensiero, etc.). As to Italian publishers,

this action, always within 12 months, must entail the automatic population of the digital legal deposit of Magazzini Digitali, provided the BNCF ensures the necessary support. As to foreign publishers, an agreement with GLN and CLOCKSS should be achieved in order to exchange plugins, avoid to replicate the work already done by other LOCKSS networks, and create economies of purpose that would result from an international ecosystem using the same LOCKSS technology.

Recommendation No. 7: Maintain relationships with other national and international archives in order to exchange experiences and best practices.

Clarification. The collaboration with the British SafeNet (ex. for Entitlement Registry) and with French ISTEX could be particularly interesting. Collaboration with European partners is also desirable in order to set up interoperable infrastructures, that could provide services to European citizens and consequently access to European Union funds.

Recommendation No. 8: Funding should be sought besides the resources possibly made available by CRUI, seizing any opportunity offered by (public or private) national and / or European and / or international funding projects.

Clarification. For a sustainable and professionally adequate migration from analog to electronic formats distributed via the web, it is important to underline that digital archives should be financed, at least in part, with the funds allocated for the analog world to buy / rent warehouses, purchase of shelving, binding, dusting, etc. The transition from analog to digital world should be achieved reviewing documents preservation expenditure rather than suppressing it, to avoid the risk of professional and managerial wrong choices.

Recommendation No. 9: CRUI-CARE should continue to offer the possibility of a national contract on the best terms for Italian universities and research institutions for the Portico Archive as further guarantee and not in lieu of National Hosting, in order to offer Italian Universities involved the protection of perpetual access rights. CRUI-CARE should negotiate an agreement for a strategic partnership with the CLOCKSS Archives to strengthen the presence of the Italian community (libraries and publishers) in this important archive, and to carry out a possible integration with other Italian digital archives, starting from digital legal deposit (DDL) of the National Library of Florence (BNCF).