

Gruppo 404.it

**L'Italia e gli Archivi digitali affidabili per le
pubblicazioni scientifiche a testo completo**

Gruppo 404

Sommario

Visione strategica	2
Considerazioni specifiche	2
Interventi operativi.....	3
Conclusione	4
Allegato 1.....	5
Raccomandazioni riguardanti le clausole da prevedere nei contratti nazionali ai fini della conservazione a lungo termine, dell'accesso perpetuo, delle attività di elaborazione dati sui contenuti oggetto dei contratti	5
Allegato 2.....	7
Raccomandazioni riguardanti la conservazione a lungo termine delle pubblicazioni scientifiche alle organizzazioni italiane che pubblicano in Open Access	7

Visione strategica

Le pubblicazioni scientifiche a testo completo in edizione elettronica (*ejournals*, *ebooks* e altri documenti simili), di seguito per brevità chiamate pubblicazioni, costituiscono una delle basi documentali fondamentali per la ricerca scientifica. Per un paese che intende fondare il proprio modello di sviluppo economico e sociale sulla ricerca e sull'innovazione è strategico garantirsi il controllo dei metadati e dei contenuti delle pubblicazioni mediante un sistema di archivi digitali affidabili sotto il controllo di istituzioni italiane, in un quadro di cooperazione europea e internazionale, ricorrendo a modelli organizzativi ed economici sostenibili nel tempo.

Questa scelta strategica è già stata assunta e praticata ad esempio da paesi come gli Stati Uniti (dove ad oggi sono state realizzate tra le più rilevanti iniziative di Archivi digitali) e la Francia (con il progetto ISTE¹). Una scelta strategica simile sta maturando anche in Germania (che ha lanciato l'iniziativa NatHosting¹) e in altri paesi tra cui il Regno Unito. In Italia da anni sono attive iniziative importanti che oggi richiedono una rivisitazione, un rilancio e una sistematizzazione in un quadro di riferimento generale, in grado di perseguire al meglio la visione strategica descritta.

Considerazioni specifiche

Rispetto alle modalità di accesso e fruizione delle pubblicazioni, esse possono essere *Open Access* (OA) oppure *Toll Access* (TA)².

Per entrambe le modalità è essenziale assicurare la conservazione a lungo termine delle pubblicazioni.

Per le pubblicazioni TA è in aggiunta necessario garantire gli investimenti³ attraverso infrastrutture tecnologiche sotto il controllo di istituzioni italiane capaci di erogare servizi di accesso *post cancellation* (PCA) ai documenti anche in caso di non rinnovo degli abbonamenti⁴; interruzioni dell'accesso sui server degli editori e altri casi simili⁵.

Inoltre, sia per le pubblicazioni OA che per le pubblicazioni TA, è strategico mettere a disposizione dei ricercatori italiani archivi digitali interoperabili capaci di offrire servizi a valore aggiunto, in particolare relativamente alle attività di *Text Mining* e *Data Mining* (TDM).

Infine, per le pubblicazioni edite in Italia, si pone anche una questione di deposito legale digitale: è necessario mettere a punto soluzioni semplici e poco costose per gli editori accademici italiani capaci di assolvere in modo coordinato le funzioni di conservazione a lungo termine, accesso *post cancellation* (PCA) e deposito legale digitale.

Le iniziative significative già attive che cercano di perseguire la scelta strategica di mantenere in Italia (seppur con differenti livelli di autonomia) il controllo dei dati e dei metadati delle pubblicazioni sono attualmente:

- a) l'Archivio NERA presso CINECA;
- b) due installazioni di LOCKSS Global Network attive in due organizzazioni residenti in Italia;
- c) l'Archivio internazionale e certificato CLOCKSS che ha in Italia uno dei 12 nodi di cui si compone;

¹ Cfr. <http://www.nathosting.de>

² Si fa riferimento in particolare ai contenuti per i quali è stato effettuato un pagamento acquisendo diritti di accesso perpetuo.

³ Non risulta l'esistenza di dati ufficiali a riguardo, ma il Gruppo 404 *stima* che il livello di spesa per risorse elettroniche a testo completo da parte di Università e Centri di ricerca italiani dovrebbe essere compreso tra 80 e 120 milioni di euro all'anno. Si ribadisce che si tratta di una stima che viene riportata per dare un ordine di grandezza del volume di spesa di cui si discute.

⁴ L'esperienza francese del progetto ISTE¹ suggerisce anche che avere la disponibilità di archivi digitali sotto il proprio controllo attraverso i quali garantirsi l'accesso perpetuo accresce la forza contrattuale nelle trattative nazionali avviate con gli editori.

⁵ In caso l'editore venda dei contenuti ad un altro editore; in caso l'editore cessi la pubblicazione di un titolo; in caso l'editore decida di non rendere più disponibile l'accesso alle annate di periodico più datate; altri casi simili.

d) Magazzini Digitali della Biblioteca Nazionale Centrale di Firenze (BNCF).

Inoltre, un importante numero di istituzioni universitarie e di ricerca italiane (circa 30) aderisce all'Archivio Portico. Si tratta di uno dei più importanti archivi digitali certificati a livello internazionale ai fini della conservazione a lungo termine e dell'accesso perpetuo delle pubblicazioni accademiche. E' un Archivio sviluppato negli Stati Uniti secondo un modello di servizio cliente-fornitore: ossia, le istituzioni per determinati editori possono ricorrere all'archivio per l'accesso perpetuo ma il controllo resta esclusivamente nelle mani dell'organizzazione statunitense, nel rispetto degli accordi con gli editori.

Interventi operativi

Per poter realizzare la visione strategica sopra accennata è necessario mettere in atto una serie di interventi operativi, primo fra tutti quello di ottenere in misura sempre più efficace il coinvolgimento di coloro che a diverso titolo partecipano al governo del sistema accademico e di ricerca italiano. Sarà essenziale elaborare una proposta su come le diverse iniziative in atto in Italia debbano essere ripensate e di come possano coordinarsi tra di loro. E' importante infine analizzare anche l'opportunità di attivare nuove sperimentazioni, prima fra tutte quella relativa a una evoluzione della Rete LOCKSS Global su base nazionale, che avrebbe anche il vantaggio di accrescere, anche in Italia, le competenze tecniche per la gestione di una tecnologia applicabile su larga scala alla conservazione dei documenti *web-based*.

La circostanza per cui il gruppo CARE, presso la CRUI, sta lavorando ad un modello di licenza nazionale costituisce un'opportunità per il Gruppo 404 di fornire in modo tempestivo un contributo da offrire alle persone impegnate in questa importante attività. In allegato 1 le raccomandazioni riguardanti le clausole da predisporre nei contratti nazionali al fine di tradurre la visione strategica negli accordi con gli editori di pubblicazioni accademiche.

Inoltre, il crescente successo delle pubblicazioni OA anche in Italia ha consigliato il Gruppo 404 di produrre (allegato 2) le raccomandazioni riguardanti la conservazione a lungo termine delle pubblicazioni OA da indirizzare agli editori accademici italiani.

Infine, per rendere più semplice la vita agli editori e meno costose le soluzioni per editori e biblioteche, la necessità di iniziare a predisporre accordi e meccanismi automatici che consentano di alimentare il deposito legale digitale, sia per le pubblicazioni OA che per le pubblicazioni TA, mediante accordi tra la BNCF e i singoli archivi (NERA, CLOCKSS, ecc.) impegna il Gruppo ad individuare soluzioni specifiche e poco costose nei prossimi mesi. Intanto, per le pubblicazioni OA, alcune indicazioni sono già contenute nell'allegato 2.

Conclusione

Il senso di questo primo documento è di contribuire alla formazione di una visione strategica generale, prerequisito fondamentale nella costruzione di un quadro sinergico in cui le iniziative in corso possano essere valorizzate e operare collaborativamente in vista di obiettivi condivisi.

I lavori del Gruppo 404 continueranno sino al 30 giugno 2015 quando termineranno con il rilascio di un documento finale in cui verrà avanzata una proposta complessiva e operativa rispetto alle iniziative da intraprendere, partendo dagli Archivi già attivi in Italia, perché la visione strategica possa trovare piena attuazione.

Questo documento, come il Gruppo 404 che ne è l'estensore, è aperto ai contributi dei diversi interlocutori, nell'intento di concordare una strategia comune e far convergere le diverse iniziative verso obiettivi condivisibili e raggiungibili.

Milano, 30 settembre 2014

Gruppo 404

Email: info@gruppo404.it

Sito web: www.gruppo404.it

Tel. +39 (0)2 7234 2339

Membri del Gruppo 404:

- Paolo Baglioni – Istituto Universitario Europeo
- Giovanni Bergamin – Biblioteca Nazionale Centrale di Firenze
- Carmen Cirulli – Politecnico di Milano
- Anna Grazia Chiodetti – INGV
- Gianluca D'Amato – Università Cattolica del Sacro Cuore – Coordinatore del Gruppo
- Vincenzo De Luise – Università degli Studi Federico II di Napoli
- Paola Gargiulo – CINECA
- Tommaso Giordano – Istituto Universitario Europeo
- Andrea Marchitelli – CINECA
- Anna Ortigari – Università degli Studi di Bologna
- Massimo Reboldi – Università degli Studi di Brescia
- Eleonora Rodi – CINECA
- Andrea Sada – Università Cattolica del Sacro Cuore
- Rossella Targa – Università degli Studi di Verona
- Federica Zanardini – Università Statale di Milano

La partecipazione al gruppo è di tipo personale e non impegna le singole istituzioni di appartenenza.

Allegato 1

Raccomandazioni riguardanti le clausole da prevedere nei contratti nazionali ai fini della conservazione a lungo termine, dell'accesso perpetuo, delle attività di elaborazione dati sui contenuti oggetto dei contratti

Il gruppo CARE, presso la CRUI, sta predisponendo un testo base per i contratti nazionali da sottoporre agli editori di pubblicazioni scientifiche. E' utile rimarcare ancora una volta la grande utilità dell'iniziativa anche ai fini dell'accesso perpetuo e della conservazione a lungo termine. Ai fini delle tematiche di cui al presente documento e anche dopo aver visionato i testi delle licenze nazionali francese (progetto ISTEEX), tedesca (modello di licenza predisposto dal gruppo di lavoro sulle licenze nazionali all'interno dell'iniziativa "Digital Information" promossa dell'Alliance of Science Organizations in Germany) e britannica (NESLI2 di JISC Collection) si suggerisce di prevedere, nel redigendo schema di contratto nazionale, le seguenti clausole.

Si precisa che, non conoscendo la strutturazione giuridica della licenza nazionale, la formulazione di seguito suggerita necessita degli opportuni adattamenti e integrazioni per rendere le previsioni congrue nell'economia generale del testo contrattuale. Per semplicità, con il termine Licenziatario si intende sia il firmatario della licenza (CRUI) sia i titolare dei diritti (le Istituzioni aderenti al contratto).

Conservazione digitale a lungo termine e accesso perpetuo, attività di text e data mining

Conservazione a lungo termine

Ai fini della conservazione digitale di lungo termine (Digital Preservation o Long Term Preservation) il Licenziante si impegna a riversare con regolarità e completezza i contenuti oggetto del presente contratto in almeno due tra gli archivi digitali certificati CLOCKSS, Portico e la rete LOCKSS GLOBAL Network.

Accesso perpetuo

Ai fini dell'accesso perpetuo (o perpetual access o post termination accesso o post cancelation access) il Licenziante si impegna a riversare con regolarità e completezza i contenuti (dati e metadati) oggetto del presente contratto all'Archivio NERA del Consorzio CINECA rispettando le specifiche tecniche e i tempi di cui all'Allegato tecnico per la consegna dei dati e dei metadati. In aggiunta o in alternativa il Licenziante si impegna a predisporre i propri contenuti oggetto del presente contratto perché siano salvati con regolarità e completezza nella rete LOCKSS Italia quando questa dovesse diventare operativa secondo le specifiche tecniche che verranno concordate.

Il riversamento dei dati avverrà senza costi aggiuntivi.

Una volta cessato il presente accordo, ai fini dell'accesso perpetuo, il Licenziatario potrà scegliere a sua discrezione una o più delle seguenti opzioni:

- a. accesso sui server dell'editore senza costi aggiuntivi [ovvero pagando una quota minima concordata tra le parti];*
- b. accesso sui server dell'Archivio NERA (o sui server della rete LOCKSS Italia in caso di sua costituzione);*
- c. accesso tramite Portico o LOCKSS GLOBAL per gli editori che vi aderiscono;*
- d. accesso alla copia di archivio dei contenuti oggetto del presente contratto che il Licenziante si impegna ad inviare al Licenziatario in un formato elettronico concordato tra le parti.*

E' consentito l'accesso sugli archivi di cui alle lettere b), c), d) anche quando, in vigenza di contratto, i server dell'editore dovessero essere momentaneamente non disponibili, ovvero per quelle pubblicazioni per le quali il Licenziatario ha diritti di accesso perpetuo ma non risultano più correnti (front list) per i più diversi motivi

(titolo cessato; titolo cambiato; titolo pubblicato da altro editore a partire da una certa annata; titolo non più parte della collezione; altri casi simili). Inoltre è consentito l'accesso sugli archivi di cui alle lettere b), c), d), anche in vigore di contratto, a tutti i contenuti oggetto del presente contratto ad eccezione dei contenuti relativi all'anno corrente (a titolo esemplificativo, per i periodici, nel corso del 2014 attraverso gli archivi si può accedere ai contenuti fino al 2013 incluso e non anche all'annata 2014).

Ai fini dell'accesso perpetuo, per le copie di cui ai punti b), c) e d), al Licenziatario è consentito migrare i formati dei contenuti oggetto del presente accordo in maniera da garantire l'accessibilità dei contenuti nel tempo.

Il Licenziatario si impegna anche per l'accesso perpetuo al rispetto delle clausole contrattuali relative a utenti autorizzati e usi consentiti.

Attività di text mining e data mining e relativi diritti di proprietà

Gli utenti autorizzati possono effettuare attività di text mining e di data mining (TDM) sui contenuti oggetto del presente accordo utilizzando gli archivi digitali consentiti dalla presente licenza. I diritti di proprietà sui risultati delle attività di TDM restano a vantaggio dell'Istituzione a cui gli utenti autorizzati che hanno effettuato queste attività appartengono.

In un orizzonte di medio-lungo periodo può emergere la necessità che il Licenziatario abbia interesse o si trovi nella condizioni di dover individuare, con l'accordo del Licenziante, un archivio diverso rispetto a NERA o a LOCKSS Italia a cui riferirsi per la tutela dell'accesso perpetuo e per le attività di TDM. Il Licenziante riconosce al Licenziatario la facoltà di utilizzare archivi diversi da quelli citati. Il cambiamento sarà oggetto di modifica al presente contratto.

Il mancato rispetto di quanto dettagliato nei paragrafi relativi alla conservazione a lungo termine, all'accesso perpetuo e alle attività di TDM è causa di risoluzione anticipata del contratto.

[RISPETTO AL PRESENTE ALLEGATO 1 LE PERSONE DEL GRUPPO 404 CHE AFFERISCONO AL CONSORZIO CINECA SI SONO ASTENUTE RITENENDO LA MATERIA DI ESCLUSIVA COMPETENZA DEGLI ENTI CHE SOTTOSCRIVONO I CONTRATTI]

Milano, 30 settembre 2014

Gruppo 404

Allegato 2

Raccomandazioni riguardanti la conservazione a lungo termine delle pubblicazioni scientifiche alle organizzazioni italiane che pubblicano in Open Access

Anche le riviste scientifiche ad accesso aperto necessitano di essere rese accessibili nel tempo, tale principio viene ribadito nell'articolo 11 delle "Raccomandazioni della Commissione Europea agli Stati Membri del luglio 2012 sull'accesso e la conservazione dell'informazione scientifica". Pertanto si raccomanda agli editori che pubblicano riviste e monografie ad accesso aperto in linea con le esortazioni dalla Commissione Europea di aderire ad archivi digitali affidabili, le cui politiche di accesso devono necessariamente prevedere la fruibilità ad accesso aperto di tali contenuti depositati nei propri archivi.

Si raccomanda inoltre di preferire archivi in grado di consentire all'editore di rendersi conforme non solo alla conservazione a lungo termine, ma anche al deposito legale digitale presso Magazzini digitali, di cui alla legge n. 106 del 15.4.2004⁶.

A solo titolo esemplificativo una via praticabile potrebbe essere quella di aderire all'Archivio internazionale e certificato CLOCKSS e che questo archivio riversi i dati in Magazzini Digitali in forza di un accordo tra CLOCKSS e la BNCF. Sempre a titoli esemplificativo, per gli editori OA che usano istanze OJS, potrebbe essere utile autorizzare la costituenda rete PKP LOCKSS⁷ a conservare i contenuti pubblicati rafforzando nel contempo la collaborazione già esistente tra gli editori OA italiani che usano un'istanza OJS e la BNCF perché i contenuti vengano riversati in modo automatico in Magazzini Digitali.

Milano, 30 settembre 2014

Gruppo 404

⁶ "Norme relative al deposito legale di documenti di interessi culturali destinati all'uso pubblico", pubblicata sulla Gazzetta Ufficiale n. 98 del 27 aprile 2004 Serie Generale; "Regolamento recante norme in materia di deposito legale dei documenti di interesse culturale destinati all'uso pubblico", emanato con D.P.R. n. 252 del 3 maggio 2006 - G.U. n. 191 dell'8 agosto 2006.

⁷ Cfr. <https://pkp.sfu.ca/pkp-lockss/>