

Gruppo 404.it

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

Gruppo 404

30/06/2015

Sommario

SINTESI DEL RAPPORTO	2
INTRODUZIONE.....	4
Istituzione del Gruppo 404	4
Membri del Gruppo 404	5
Organizzazione dei lavori del Gruppo.....	5
Durata e obiettivi del Gruppo 404.....	5
Finanziamento	5
Circa il nome del Gruppo	6
La struttura del Rapporto	6
Ringraziamenti.....	6
PRIMA PARTE.....	7
Visione Strategica	7
Considerazioni specifiche	7
Le iniziative oggi attive in Italia	8
L'Archivio NERA presso CINECA.....	8
Le due installazioni Global LOCKSS Network.....	8
L'Archivio CLOCKSS.....	9
Magazzini Digitali della Biblioteca Nazionale Centrale di Firenze (BNCF)	9
L'Archivio Portico.....	9
Raccomandazioni.....	9
SECONDA PARTE.....	11
Analisi dei fabbisogni.....	11
Indagine presso le iniziative di National Hosting e presso altre importanti iniziative di Archivio su scala internazionale.....	12
Analisi delle risposte.....	13
Raccomandazioni.....	21
CONCLUSIONI	24
ALLEGATI.....	25
APPENDICI	28
GLOSSARIO	48
BIBLIOGRAFIA E STANDARD	52

SINTESI DEL RAPPORTO

Il presente rapporto contiene i risultati del lavoro svolto in dodici mesi, tra luglio 2014 e giugno 2015, dal gruppo di lavoro denominato “Gruppo 404”. Il Gruppo 404 è stato costituito nel luglio 2014 con l’obiettivo di elaborare una visione condivisa sulle iniziative di archivio digitale delle pubblicazioni scientifiche (sia quelle a pagamento sia quelle ad accesso aperto) da intraprendere in Italia a vantaggio della comunità accademica e di ricerca italiana.

Il gruppo, costituito da specialisti, bibliotecari e informatici con consolidate esperienze nell’ambito dell’accesso perpetuo e della conservazione a lungo termine delle risorse informative elettroniche, ha elaborato questo rapporto finale che si compone di due parti: la prima delinea la visione strategica e formula le prime quattro raccomandazioni; la seconda si concentra sull’analisi dei fabbisogni, sulle infrastrutture e sui servizi su base nazionale presenti sia in Italia sia all’estero (Francia, Gran Bretagna, Stati Uniti, Brasile) e si conclude con la formulazione di cinque successive raccomandazioni.

Lo studio è partito dalla necessità di delineare una visione condivisa. Infatti, solo a partire da una visione strategica condivisa e da una partecipazione attiva e collaborativa degli attori coinvolti (università, centri di ricerca, ministeri, biblioteche nazionali) è possibile intraprendere iniziative sostenibili e azioni concrete in grado di soddisfare i fabbisogni della comunità scientifica e di ricerca italiana in termini di accesso perpetuo (PCA, Post Cancellation Access), di conservazione a lungo termine (DP, Digital Preservation) e di deposito legale digitale (DLD, Digital Legal Deposit) delle pubblicazioni scientifiche. Una volta raggiunte la condivisione di vedute e di intenti e una chiara definizione dei bisogni, delle priorità e degli obiettivi, solo allora è possibile formulare delle raccomandazioni sulle azioni da intraprendere. Il lavoro ha quindi analizzato le iniziative presenti in Italia: l’Archivio NERA presso CINECA, le due installazioni di Global LOCKSS Network, l’archivio CLOCKSS, Magazzini Digitali presso la Biblioteca Nazionale Centrale di Firenze, l’archivio Portico. Da questo lavoro sono scaturite le prime raccomandazioni: la prima sottolinea la necessità di ottenere un coinvolgimento sempre più efficace di coloro che a diverso titolo partecipano al governo del sistema accademico e di ricerca italiano e in particolare i Ministeri competenti (MIUR e MIBACT), la CRUI, i singoli Atenei, i principali Centri di ricerca. La seconda si concentra sulla necessità che nel modello di licenza nazionale siano previste delle apposite clausole contrattuali in grado di sostanziare la visione strategica circa l’accesso perpetuo per le risorse elettroniche a pagamento. La terza riguarda le pubblicazioni scientifiche ad accesso aperto, ossia la necessità che anch’esse adottino soluzioni adeguate in tema di conservazione a lungo termine, anche a tutela degli autori. La quarta riguarda la sostenibilità e la riduzione dei costi nell’ambito della DP, del PCA e del DLD sia da parte degli editori, sia delle biblioteche, la cui realizzazione può avvenire se si predispongono accordi e meccanismi automatici che consentano di alimentare il deposito legale, sia per le pubblicazioni ad accesso aperto (OA), sia per le pubblicazioni a pagamento (Toll Access TA) direttamente dagli archivi nei quali gli editori italiani riversano i contenuti ai fini della DP e/o del PCA (per esempio Portico; CLOCKSS; ecc.).

Dall’analisi dettagliata dei fabbisogni applicata al contesto italiano, assegnando un ordine di priorità si è giunti nel corso dei lavori del gruppo alla individuazione dell’accesso perpetuo come il bisogno a cui dare la massima priorità e alla considerazione che sarebbe stata opportuna un’indagine sulle principali iniziative di National Hosting realizzate o in corso di realizzazione a livello internazionale.

Su questi aspetti si concentra la seconda parte del rapporto. Sono stati pertanto organizzati incontri di persona con rappresentanti di alcune di queste iniziative, in particolare con i colleghi francesi di Couperin relativamente al progetto ISTEEX e con il direttore e il responsabile scientifico di LOCKSS. Volendo ampliare il panorama ed entrare in contatto con altre esperienze è stato elaborato un questionario che è stato inviato agli archivi che offrono servizi di National Hosting (ISTEEX in Francia, Scholars Portal in Canada, EJC ed EBC del consorzio OhioLINK per l’Ohio); a interessanti progetti pilota come ad esempio la sperimentazione

SafeNet in Gran Bretagna; agli archivi digitali quali Portico, la rete Cariniana in Brasile, il programma Global LOCKSS Network.

In base alle risposte pervenute e in alcuni casi in base alla corrispondenza intercorsa con i responsabili di questi servizi, ciascuna iniziativa viene presentata nei suoi punti salienti (dalla tecnologia utilizzata, al modello di governance e di finanziamento, ecc.), è accompagnata da una comparazione rispetto alla visione strategica e ai fabbisogni delineati dal Gruppo 404 e si conclude con brevi considerazioni finali che riflettono su ciascuna iniziativa nel suo complesso, tenendo sempre presente quale ricaduta tale soluzione potrebbe avere nel contesto italiano e/o quali possibili forme di collaborazione si potrebbero stabilire con queste realtà.

In seguito all'analisi del contesto italiano e internazionale vengono formulate cinque raccomandazioni che si aggiungono alle quattro della prima parte. Pertanto si raccomanda di verificare la possibile evoluzione tecnologica e organizzativa dell'archivio NERA con la quantificazione dei costi per valutarne la capacità di rispondere alle esigenze qualitative del servizio, l'integrazione con altri servizi gestiti dagli Atenei (Raccomandazione 5); di intraprendere una PNL LOCKSS con sei università che diano sufficienti disponibilità e garanzie coordinandosi con la CRUI in quanto riferimento per le trattative nazionali (Raccomandazione 6); mantenere le relazioni con gli altri archivi nazionali e internazionali al fine di scambiare esperienze e best practice (Raccomandazione 7); acquisire fondi necessari per la gestione del National Hosting anche tramite la partecipazione a bandi (di enti pubblici o privati) nazionali e/o europei, e/o internazionali (Raccomandazione 8); continuare a offrire alle università e agli enti di ricerca italiani un contratto nazionale tramite CRUI-CARE per l'adesione italiana a Portico, in quanto garanzia ulteriore e non sostitutiva del National Hosting e negoziare tramite CRUI-CARE un accordo per una partnership con l'Archivio CLOCKSS, sia per rafforzare la presenza della comunità italiana (biblioteche ed editori) in questo archivio, sia per esperire possibili integrazioni con altri archivi digitali italiani, quali ad esempio il deposito legale digitale della BNCF (Raccomandazione 9).

Lo studio si conclude con l'auspicio che il lavoro svolto da questo gruppo possa essere raccolto da CRUI-CARE e portato avanti tramite la costituzione di un team specialistico *ad hoc* e la nascita di un progetto sperimentale della durata indicativa di ventiquattro mesi (2016-2017). Il progetto e la sua sperimentazione dovranno mettere la comunità accademica italiana nelle condizioni di poter decidere l'assetto organizzativo e di finanziamento finale da adottare per un'iniziativa italiana sostenibile di National Hosting.

INTRODUZIONE

Il presente rapporto contiene i risultati del lavoro svolto in dodici mesi, tra luglio 2014 e giugno 2015, dal gruppo di lavoro denominato “Gruppo 404”.

Gli esiti di questo documento sono principalmente due. Innanzitutto viene delineata una visione strategica condivisa quale presupposto indispensabile per intraprendere quegli interventi operativi necessari in grado di soddisfare i fabbisogni della comunità scientifica e di ricerca italiana in termini di accesso perpetuo (da ora in avanti anche indicato come PCA, Post Cancellation Access), conservazione a lungo termine (da ora in avanti anche indicato come DP, Digital Preservation) e deposito legale digitale (da ora in avanti anche indicato come DLD, Digital Legal Deposit) delle pubblicazioni scientifiche; inoltre vengono formulate delle raccomandazioni sulle azioni da intraprendere per realizzare la visione strategica sulla base delle condizioni date.

La pubblicazione del rapporto realizza tutte le finalità del Gruppo 404 che pertanto, a seguito della sua pubblicazione, è stato sciolto¹.

Il rapporto, che è nelle disponibilità dei membri del Gruppo, verrà messo a disposizione della comunità scientifica e di ricerca italiana e in particolare: della CRUI, che dal 2015 ha assunto pienamente la titolarità delle trattative nazionali per le risorse elettroniche; dei Ministeri competenti, ossia il MIUR poiché gli archivi digitali affidabili sono a servizio della comunità italiana della ricerca e della didattica universitaria e il MIBACT poiché responsabile del deposito legale digitale. Inoltre, il rapporto sarà messo a disposizione degli archivi che hanno partecipato all’indagine relativa alle iniziative di National Hosting.

Ai fini delle definizioni dei termini contenuti nel presente documento si rimanda al Glossario.

I contenuti del presente documento sono soggetti alla licenza Creative Commons CC BY-NC.

Istituzione del Gruppo 404

Il Gruppo 404 è stato costituito il 3 luglio 2014 per iniziativa della Comunità italiana delle Biblioteche aderenti all’archivio CLOCKSS e ha avuto la finalità di elaborare entro giugno 2015 (12 mesi di attività), partendo dalle esigenze concrete delle biblioteche universitarie e di ricerca italiane, una visione condivisa rispetto a quali iniziative e a quali Archivi digitali affidabili e di come essi si relazionino al meglio tra loro, per consentire al sistema della ricerca italiano di garantire in modo adeguato le pubblicazioni scientifiche ai fini dell’accesso perpetuo e della conservazione a lungo termine, sia per i contenuti a pagamento (da ora in avanti anche indicati come TA, Toll Access), sia per i contenuti ad accesso aperto (da ora in avanti anche indicati come OA, Open Access). Inoltre, limitatamente alle pubblicazioni scientifiche italiane, quali soluzioni siano auspicabili perché si creino raccordi con il DLD in modo da rendere più economico e sostenibile il flusso di lavoro e gli impegni economici da parte degli editori italiani.

La fondazione del gruppo è stata ritenuta necessaria per creare un luogo di confronto e di approfondimento capace di superare i limiti dei gruppi riferiti a singoli archivi (il gruppo CLUG DP per l’Archivio NERA del Consorzio CINECA; il gruppo delle biblioteche italiane aderenti a Portico; gli aderenti a Global LOCKSS Network; la comunità italiana di CLOCKSS; l’iniziativa Magazzini Digitali presso la BNCF) in una prospettiva strategica più complessiva.

¹ Il dominio gruppo404.it, dove è stato pubblicato anche il presente rapporto, non sarà rinnovato; i contenuti del sito saranno ospitati provvisoriamente sui server dell’Università degli Studi di Milano alla URL <<http://users.unimi.it/gruppo404>>; inoltre, il sito gruppo404.it è stato archiviato in The Internet Archive (<<https://archive.org/>>); infine, è stata inoltrata una richiesta di conservazione presso Magazzini Digitali della Biblioteca Nazionale Centrale di Firenze (BNCF). Resta aperta la possibilità che alcuni contenuti vengano ripubblicati su altri siti web interessati ai contenuti del presente lavoro, ovvero che il rapporto sia oggetto di una vera pubblicazione.

Il Gruppo 404 è nato perché i problemi del PCA e della DP rappresentano una sfida difficile e complessa, che può avere successo solo mediante un importante sforzo organizzativo e di cooperazione, capace di mettere a sistema le diverse iniziative nazionali e internazionali e i soggetti che le promuovono.

Membri del Gruppo 404

La partecipazione ai lavori del Gruppo 404 era aperta ai seguenti soggetti: i membri italiani aderenti all'archivio CLOCKSS, i membri di CARE-CRUI che intendevano prendere parte all'iniziativa, personale del consorzio CINECA coinvolto in NERA o in iniziative di pubblicazioni OA, personale della Biblioteca Nazionale Centrale di Firenze coinvolto in Magazzini Digitali, i promotori dell'archivio Portico in Italia, altri specialisti che hanno dimostrato particolare attenzione e impegno sui temi della DP e del PCA. I partecipanti sono specialisti (bibliotecari, informatici, ecc.) coinvolti a diverso titolo nelle iniziative citate.

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

Al Gruppo hanno aderito dalla sua fondazione quindici persone, a cui si è aggiunta una sedicesima nel corso delle attività. Il Gruppo ha nominato al suo interno un coordinatore per facilitare le attività. Hanno fatto parte del Gruppo di lavoro:

- Paolo Baglioni – Istituto Universitario Europeo
- Giovanni Bergamin – Biblioteca Nazionale Centrale di Firenze
- Domenico Castellani – Università degli Studi di Padova (da aprile 2015)
- Carmen Cirulli – Politecnico di Milano
- Anna Grazia Chiodetti – Istituto Nazionale di Geofisica e di Vulcanologia
- Gianluca D'Amato – Università Cattolica del Sacro Cuore – Coordinatore del Gruppo
- Vincenzo De Luise – Università degli Studi di Napoli "Federico II"
- 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

Organizzazione dei lavori del Gruppo

Il Gruppo 404 si è incontrato tre volte in presenza e otto volte via teleconferenza. I documenti del Gruppo sono stati votati e approvati con larghe maggioranze².

Durata e obiettivi del Gruppo 404

Il Gruppo 404 ha avuto durata di dodici mesi e ha approvato un rapporto intermedio (approvato il 30 settembre 2014) e il presente documento finale che include anche gli esiti del rapporto intermedio.

Finanziamento

Per le spese di funzionamento ogni membro del gruppo ha provveduto con spese a carico dell'organizzazione di provenienza. L'Università Cattolica del Sacro Cuore ha messo a disposizione spazi e

² Le votazioni hanno avuto pochi casi di membri che non hanno preso parte alle votazioni; non ci sono mai stati voti contrari.

strutture (incluso il sito web) perché il gruppo potesse lavorare per i dodici mesi decorrenti dalla sua fondazione.

Circa il nome del Gruppo

Il Gruppo 404 ha preso il nome dal codice “404 Not found” (o comunemente noto come “page not found”) che il Consorzio W3C assegna in caso il server non trovi nulla corrispondente alla URI richiesta.

Un gruppo che vuole affrontare e risolvere il problema degli accessi in perpetuo si dà un nome che dichiara un fallimento (pagina non trovata)? Il punto è che la sfida che il PCA, la DP e il DLD pongono alle Biblioteche e ai sistemi documentali è molto impegnativa. Si tratta di una rivoluzione che qualcuno paragona al passaggio dal manoscritto alla stampa a caratteri mobili, altri al passaggio da una cultura orale a una cultura scritta³. L'importanza della sfida da un lato, i dubbi sulle capacità di comprendere e vincere la sfida dall'altro, il tutto combinato con la confortante circostanza per cui sempre più nuove energie e professionalità si stanno dedicando a questa sfida, generano situazioni contraddittorie che, in particolare in noi italiani, suscitano ironia. Dall'ironia si possono avere esiti diversi: evolvere in fatalismo disfattista; suscitare un velleitarismo provinciale; mobilitare orgoglio, passione, competenze e capacità per vincere la sfida. Vedremo quale strada saremo in grado di intraprendere...

La struttura del Rapporto

Il documento si divide in due parti. La prima parte delinea la visione strategica e ha generato le prime quattro raccomandazioni. La seconda parte si concentra sul National Hosting e sulla combinazione del National Hosting con le iniziative di Archivio già in qualche modo attive in Italia e ha generato altre cinque raccomandazioni.

Ringraziamenti

I membri del Gruppo 404 ringraziano le istituzioni di appartenenza per aver dato loro la possibilità di lavorare a questo rapporto; ringraziano inoltre CRUI-CARE per il supporto e per i contatti avuti durante i lavori e in particolare il Team “Convenzione CRUI/Enti e definizione contratto tipo”; ringraziano infine gli archivi che hanno dato il loro supporto alla conclusione dell'indagine riguardante le iniziative di National Hosting.

³ Cfr. ad esempio Maryanne Wolf, *Proust and the Squid. The Story and Science of the Reading Brain*, New York, HarperCollins, 2007, trad. It. *Proust e il calamaro: storia e scienza del cervello che legge*, Milano, Vita e Pensiero, 2009.

PRIMA PARTE

Visione Strategica

Le pubblicazioni scientifiche a testo completo in edizione elettronica (periodici elettronici, libri elettronici 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), il Canada (con Scholars Portal) e la Francia (con il progetto ISTE⁴). Una scelta strategica simile sta maturando anche in Germania (che ha lanciato l'iniziativa NatHosting)⁴ e in altri paesi tra cui il Regno Unito (dove è in corso il progetto SafeNet). In Italia da anni sono attive iniziative importanti che oggi richiedono una profonda rivisitazione, un rilancio e una sistematizzazione in un quadro di riferimento generale, in grado di perseguire al meglio la visione strategica descritta.

Considerazioni specifiche

Come noto, rispetto alle modalità di accesso e fruizione, le pubblicazioni possono essere Open Access (OA) oppure Toll Access (TA)⁵. Per entrambe le modalità è essenziale assicurare la conservazione a lungo termine delle pubblicazioni (DP).

Per le pubblicazioni TA è in aggiunta necessario garantire gli investimenti⁶ attraverso infrastrutture tecnologiche sotto il controllo di istituzioni italiane, capaci di erogare servizi di accesso perpetuo (PCA) ai documenti anche in caso di non rinnovo degli abbonamenti⁷, interruzioni dell'accesso sui server degli editori e altri casi simili⁸.

Inoltre, sia per le pubblicazioni OA sia 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 (da ora in avanti anche indicato come TDM).

Infine, per le pubblicazioni editate 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 perpetuo e deposito legale digitale.

Questi fabbisogni vengono ripresi, analizzati e sistematizzati per priorità nell'analisi dei fabbisogni di cui alla seconda parte del Rapporto.

⁴ Cfr. <<http://www.nathosting.de>>.

⁵ Si fa riferimento in particolare ai contenuti per i quali è stato effettuato un pagamento acquisendo diritti di accesso perpetuo.

⁶ 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.

⁷ L'esperienza francese del progetto ISTE⁴ 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.

⁸ In caso l'editore venda dei contenuti a 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.

Le iniziative oggi attive in Italia

Le iniziative significative già attive in Italia che cercano di perseguire la scelta strategica di mantenere, 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 dodici nodi di cui si compone;
- 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.

L'Archivio NERA presso CINECA

Circa l'evoluzione del Local Hosting in Italia si sono tenute diverse presentazioni e sono stati prodotti diversi documenti, soprattutto da parte dei Consorzi. In particolare si rimanda ai lavori svolti dal gruppo utenti CINECA denominato CLUG Digital Preservation, che ha lavorato da marzo 2013 fino a ottobre 2013. Ai fini del presente rapporto si ritiene utile richiamare solo quegli elementi validi in relazione ai lavori del Gruppo 404.

L'Archivio NERA⁹ presso CINECA è succeduto alle due iniziative di Local Hosting attive in Italia presso i consorzi CASPUR (il servizio si chiamava Emeroteca Virtuale) e CILEA (il servizio si chiamava SDOS), i quali sono stati fusi nel 2013 per incorporazione con il consorzio CINECA (la fusione è iniziata nel 2012 e si è completata nel 2013). Entrambi i servizi usavano il software proprietario di Elsevier e denominato ScienceServer. Il Local Hosting è stato abbinato sin da subito ai contratti consortili o interconsortili come condizione fondamentale per migrare, a partire dal 1999-2000, gli abbonamenti ai periodici dall'edizione a stampa all'edizione solo online, poiché molte istituzioni aderenti chiedevano di poter mantenere il controllo dei metadati e dei contenuti in considerazione del volume della spesa degli investimenti. Il software ScienceServer a partire dalla fine degli anni 2000 non è stato più mantenuto da Elsevier, per cui il CILEA ha iniziato a lavorare alla sua sostituzione. Il 12 dicembre 2012 è stata messa online la nuova piattaforma Software NERA, che ha affiancato le piattaforme SDOS ed Emeroteca Virtuale sino all'autunno 2014¹⁰, quando esse sono state messe offline indirizzando gli utenti sulla sola NERA. La piattaforma NERA è stata sviluppata con software originario open source¹¹.

Allo stato attuale l'archivio NERA non si è rivelato all'altezza di sostituire SDOS ed Emeroteca Virtuale. Ci sono problemi sia relativamente alla fruizione dei contenuti da parte degli utenti, sia per quanto riguarda la completezza dei contenuti in relazione agli archivi precedenti. Il giudizio degli utenti di NERA è unanime circa la necessità di una profonda revisione per poter mettere l'archivio nelle condizioni di offrire un servizio paragonabile a quello di SDOS e di Emeroteca Virtuale.

Le due installazioni Global LOCKSS Network

A partire dal 2006 in Università Cattolica del Sacro Cuore e dal 2007 presso l'Istituto Universitario Europeo di Firenze sono attive due installazioni di Global LOCKSS Network (GLN)¹². Si tratta di server di proprietà

⁹ Cfr. <<http://nera.cineca.it/vufind/>>.

¹⁰ SDOS è stato dismesso in data 1/10/2014.

¹¹ I software usati sono: FedoraCommons per il repository; Apache Solr per il layer di indicizzazione; Vufind e PHP per l'interfaccia web; Maestrale e PHP per l'interfaccia di carico dei dati; JAVA per l'interfaccia di amministrazione.

¹² Cfr. <<http://www.lockss.org/>>.

delle due istituzioni, su cui sono state salvate diverse migliaia di annate di periodici pubblicati da vari editori. Queste due installazioni sono state usate come integrative di SDOS prima e di NERA poi, poiché editori importanti come Elsevier e Wiley non aderiscono a GLN, mentre riversano i dati nel local loading di SDOS-NERA. Pertanto allo stato attuale GLN non è nelle condizioni di poter essere l'unica soluzione di archivio in grado di tutelare gli investimenti in risorse elettroniche delle istituzioni italiane; rappresenta una valida integrazione in considerazione di diversi aspetti e in particolare di due elementi: le istituzioni mantengono il controllo di dati e metadati come delineato nella visione strategica; più di 500 editori hanno scelto di offrire PCA e DP attraverso LOCKSS, tra cui editori come Taylor and Francis, Sage, Oxford University Press.

L'Archivio CLOCKSS

Anche rispetto all'Archivio CLOCKSS¹³ si rimanda ai diversi documenti prodotti e, in particolare, alle presentazioni tenute in occasione degli incontri della comunità italiana di CLOCKSS, organizzate ogni anno a partire dal 2011. Ai fini del presente rapporto si precisa che, allo stato attuale, sono dodici le biblioteche italiane aderenti; si tratta di un archivio che si occupa solo di DP (non offre servizi di PCA). Rispetto alla visione strategica realizza quella cooperazione internazionale e si presta, se vengono conclusi appositi accordi, ad economie di scopo per fare in modo, ad esempio, che gli editori italiani siano conformi alla DP e al DLD con un unico flusso di lavoro e a costi contenuti; si presta inoltre a consentire alle pubblicazioni OA italiane, in particolare a quelle attive presso le Università italiane, di essere sufficientemente coperti rispetto alla DP su scala internazionale a costi contenuti.

Magazzini Digitali della Biblioteca Nazionale Centrale di Firenze (BNCF)

Per Magazzini Digitali si rimanda alla documentazione presente sul sito web¹⁴. Ai fini del presente rapporto si rimarca l'importanza dell'impegno volto a creare un ecosistema in grado di semplificare i flussi di lavoro e contenere i costi, perché gli editori italiani di pubblicazioni scientifiche siano messi nelle condizioni di popolare il DLD e di dare sufficienti garanzie in tema di DP e di PCA.

L'Archivio Portico

Portico è uno dei più importanti archivi digitali a livello internazionale ai fini della conservazione a lungo termine e dell'accesso perpetuo delle pubblicazioni accademiche. La bibliografia e la documentazione su Portico¹⁵ sono estese. È 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 no profit statunitense ITHAKA, di cui Portico è uno dei tre servizi erogati (gli altri due sono JSTOR e Ithaka S+R). In base alla risposta data da Portico alla indagine (cfr. seconda parte del rapporto) emerge una disponibilità di ITHAKA a lavorare con iniziative di National Hosting.

Raccomandazioni

Raccomandazione n° 1. Per realizzare la visione strategica è necessario ottenere in misura sempre più efficace il coinvolgimento di coloro che a diverso titolo partecipano al governo del sistema accademico e di

¹³ Cfr. <<http://www.clockss.org/>>. Il nodo in Italia è attivo presso l'Università Cattolica del Sacro Cuore di Milano.

¹⁴ Cfr. <<http://www.magazzinidigitali.it/it/home>>.

¹⁵ Cfr. <<http://www.portico.org>>.

ricerca italiano e in particolare i Ministeri competenti (MIUR e MIBACT), la CRUI e i singoli Atenei (soprattutto i più grandi), i principali Centri di ricerca nazionali.

Raccomandazione n° 2. È essenziale che, per le risorse Toll Access per cui sono stati negoziati diritti di accesso perpetuo, la visione strategica si sostanzi in clausole contrattuali; in particolare esse dovranno essere incluse nel modello di licenza nazionale predisposto dal gruppo CARE presso la CRUI, che oggi in Italia segue le trattative nazionali.

Precisazione. In allegato 1 sono riportate le clausole da prevedere nei contratti nazionali al fine di tradurre, in modo sufficientemente adeguato, la visione strategica negli accordi con gli editori di pubblicazioni accademiche¹⁶. Nell'elaborazione delle clausole si è tenuto anche conto di alcuni significativi modelli di contratti nazionali per le risorse elettroniche, di cui vengono prodotti estratti nell'Appendice A.

Raccomandazione n° 3. È necessario che anche le pubblicazioni Open Access adottino soluzioni adeguate in tema di conservazione a lungo termine (DP), anche a tutela degli autori.

Precisazione. In allegato 2 si riportano più nel dettaglio gli elementi che discendono dalla presente raccomandazione.

Raccomandazione n° 4. Con l'obiettivo di rendere meno costose e più sostenibili le attività che editori e biblioteche devono attuare ai fini della conservazione a lungo termine (DP), dell'accesso perpetuo (PCA) e del deposito legale digitale (DLD), si raccomanda di predisporre accordi e meccanismi automatici che consentano di alimentare il deposito legale digitale, sia per le pubblicazioni Open Access sia per le pubblicazioni Toll Access, direttamente da quegli archivi nei quali gli editori italiani riversano i contenuti ai fini della DP e del PCA (es. CLOCKSS, Portico, ecc.).

¹⁶ Questa raccomandazione con il relativo allegato faceva parte del report intermedio pubblicato il 30 settembre 2014. Il report venne condiviso con il team che all'interno di CRUI-CARE stava predisponendo il modello di "Contratto tipo" da sottoporre agli editori con cui intraprendere le trattative nazionali. La prima versione del contratto tipo è stata rilasciata da CARE-CRUI il 22 dicembre 2014 ed è stata pubblicata alla URL <<http://www.cruirisorselettroniche.it/linee-guida/>>. Essa prevede diverse clausole contrattuali coerenti con l'impostazione del rapporto del 30 settembre e con la visione strategica in esso riportata.

SECONDA PARTE

Dopo la pubblicazione del primo rapporto intermedio del 30 settembre 2014, le attività del gruppo si sono concentrate in due direzioni.

Innanzitutto è stata fatta un'analisi dettagliata dei fabbisogni secondo un ordine di priorità, strumento necessario per comprendere quali soluzioni di archivio sono in grado di rispondere non solo alla visione strategica generale ma anche ai fabbisogni concretamente individuati e ordinati per priorità.

Inoltre, poiché durante i lavori del gruppo è emerso che la priorità massima va riconosciuta al tema dell'accesso perpetuo, è stata condotta un'indagine presso le iniziative di National Hosting e presso altre importanti iniziative di Archivio digitale su scala internazionale.

Analisi dei fabbisogni

a) Servizi richiesti (indicati in ordine di priorità decrescente).

- 1) **Funzionalità e affidabilità hardware e software.** L'Archivio¹⁷ deve garantire con le necessarie performance (di stabilità e velocità della connessione, di affidabilità del software, di completezza dei contenuti, di tempestività nel caricamento dei contenuti) l'accesso al testo completo dei documenti per i quali le università italiane hanno acquisito diritti di accesso perpetuo mediante regolari accordi con i titolari dei diritti. L'Archivio deve garantire il rispetto di detti accordi dando le necessarie rassicurazioni agli editori.

Il recupero delle informazioni deve poter avvenire nei seguenti modi.

- i. Mediante browsing sul sito web dell'Archivio.
 1. Con lista A-Z dei titoli, sia essa generale per tutti i contenuti archiviati, sia essa particolare limitandola per editore e/o per tipologia del documento (periodici, monografie).
 2. Rispetto alle unità di documento:
 - a. per ogni periodico sfogliando per annata/volume e quindi per fascicolo e pagina;
 - b. per i libri sfogliando per capitolo quando l'editore ha predisposto la divisione del libro per capitoli.
 - ii. Mediante ricerca per campi indicizzati sul sito web dell'Archivio. A titolo di esempio per i periodici siano ricercabili i campi: Editore, Titolo di periodico, ISSN-p, ISSN-e, ISSN-L, Anno, Volume, fascicolo. Per i libri: Editore, Titolo, Autore, Anno, ISBN-p, ISBN-e.
 - iii. Mediante link resolver. L'Archivio deve poter essere settato come Target per i principali link resolver in commercio sviluppati sulla base dello standard OpenURL.
 - iv. Mediante l'esportazione dei metadati dei contenuti archiviati per consentire l'integrazione con gli OPAC e altri sistemi simili. L'Archivio deve consentire un'uscita dei dati secondo il tracciato KBART. I valori nel campo "title_url" devono essere del tipo Persistent URL.
- 2) **Interoperabilità.** L'Archivio, limitatamente alle pubblicazioni scientifiche edite in Italia, deve prevedere le integrazioni con il Deposito Legale Digitale (DLD) in modo tale che un Content provider italiano che riversi nell'Archivio, seguendo un flusso di lavoro unico, possa essere contestualmente conforme alle previsioni di legge per il DLD presso la Biblioteca Nazionale Centrale di Firenze (BNCF), progetto Magazzini Digitali, senza la necessità di altri eccessivi oneri a suo carico.
 - 3) **Servizi aggiuntivi.** L'Archivio deve permettere di condividere sia i metadati sia il testo completo dei contenuti archiviati con sistemi di Text and Data Mining (TDM).

¹⁷ Per brevità si parla di Archivio, ben sapendo che i fabbisogni possono essere coperti anche mettendo in rete più soluzioni di archivio.

- 4) **Certificazione.** L'Archivio, se non è già certificato secondo gli standard in uso a livello internazionale, deve programmare di raggiungere la certificazione di affidabilità (ISO 16363:2012) entro cinque anni dall'entrata in produzione.

b) Eventi a seguito dei quali deve essere sempre possibile ricorrere all'Archivio.

- 1) Per l'accesso al testo completo dei contenuti archiviati:
 - a. quando l'editore non è in grado di garantire l'accesso;
 - b. quando il server dell'editore ha temporanei problemi di accessibilità;
 - c. per tutti i contenuti (periodici, libri, ecc.) non più in *front list*:
 - i. periodici chiusi;
 - ii. periodici trasferiti ad altro editore;
 - iii. periodici soggetti ai cambiamenti di titolo (accesso ai contenuti pubblicati sotto il titolo non più in uso);
 - iv. altri casi simili.
 - d. quando il sottoscrittore non rinnova l'abbonamento.
- 2) Per le attività di TDM secondo le modalità concordate con gli editori.

c) Utenti autorizzati ad usare gli archivi.

Gli utenti autorizzati dalle istituzioni accademiche e di ricerca italiane che afferiscono alle istituzioni che hanno acquisito i diritti di accesso perpetuo, secondo quanto stabilito dagli accordi con i titolari dei diritti.

Per quanto attiene a Magazzini Digitali, di cui l'Archivio diventerà uno dei contributori, gli utenti autorizzati sono i cittadini italiani in generale, ma limitatamente ai contenuti editi in Italia e soggetti al deposito legale digitale, secondo quanto stabilito dalle leggi dello Stato e dagli accordi con i titolari dei diritti. I contenuti soggetti al deposito legale digitale sono accessibili solo presso determinati luoghi individuati dalla BNCF sulla base della regole stabilite dallo Stato italiano.

Indagine presso le iniziative di National Hosting e presso altre importanti iniziative di Archivio su scala internazionale

Al fine di poter condurre al meglio le proprie attività, il Gruppo, dopo aver concluso l'analisi dei fabbisogni, ha ritenuto necessario avviare un'indagine presso i principali archivi internazionali che sono collegati alle attività di National Hosting.

L'invito a partecipare è stato rivolto alle seguenti iniziative.

- a) Archivi che offrono servizi di National Hosting (ISTEX¹⁸-Panist¹⁹ per la Francia; Scholars Portal²⁰ per il Canada; EJC²¹ e EBC²² per il consorzio USA OhioLINK);
- b) i più recenti e innovativi progetti pilota di National Hosting (la sperimentazione SafeNet²³ in UK con l'applicazione di LOCKSS²⁴ per la realizzazione del National Hosting);
- c) gli archivi che, pur non essendo National Hosting in senso proprio, prestano dei servizi che possono essere in competizione con il National Hosting (archivio Portico²⁵ in USA e archivio e-Depot²⁶ in Olanda).

¹⁸ Cfr. <<http://www.istex.fr/category/plateforme/>>

¹⁹ Cfr. <<http://panist.fr/view/nonConnexion.php>>

²⁰ Cfr. <<http://www.scholarsportal.info/>>

²¹ Cfr. <<http://journals.ohiolink.edu/>>

²² Cfr. <<http://ebooks.ohiolink.edu/ebc-home/>>

²³ Cfr. <<http://www.jisc.ac.uk/rd/projects/safenet-distributed-digital-preservation-network>>

²⁴ Cfr. <<http://www.lockss.org/>>

²⁵ Cfr. <<http://www.portico.org>>

²⁶ Cfr. <<https://www.kb.nl/en/organisation/research-expertise/long-term-usability-of-digital-resources>>

Rispetto al progetto francese ISTEX, si precisa che il 3 luglio 2014 si è avuta la possibilità di conoscere e di approfondire l'importante progetto francese grazie all'intervento di Grégory Colcanap, Coordinatore del Consorzio Couperin, durante un incontro che ha avuto luogo a Milano.

Infine, in considerazione dell'interesse suscitato dall'uso della tecnologia LOCKSS ai fini del National Hosting, l'indagine è stata estesa a un'altra interessante esperienza, ossia la brasiliana Rede Cariniana²⁷. Quindi in data 22 gennaio 2015 si è organizzato un incontro a Milano con Victoria Reich e David S.H. Rosenthal, rispettivamente il Direttore e il Responsabile scientifico del programma LOCKSS presso la biblioteca dell'Università di Stanford.

In Appendice B si riportano i seguenti documenti relativi all'indagine:

- a) la lettera inviata agli archivi;
- b) il relativo questionario, che si compone di trentacinque domande divise in dieci sezioni.

Analisi delle risposte

Le risposte ricevute dagli archivi talvolta contenevano l'autorizzazione a pubblicarle nel rapporto e talvolta no poiché erano coperte da clausole di riservatezza. Per questo motivo, per non creare asimmetrie, si è deciso di non pubblicare alcuna risposta, ma di riassumerle tutte nelle presentazioni che seguono.

Alcune iniziative di archivio (come SafeNet e OhioLINK) non hanno compilato il questionario ma hanno fornito informazioni in altro modo, come specificato di seguito. Tutti gli archivi hanno fatto pervenire una risposta, a eccezione di e-Depot.

Le risposte ricevute dagli Archivi hanno consentito al Gruppo di acquisire un importante e insostituibile patrimonio di informazioni. Ai fini dei commenti che seguono, oltre alle informazioni contenute nelle risposte all'indagine, si è tenuto conto anche di altre informazioni acquisite mediante scambio di comunicazioni di posta elettronica, telefonate, analisi della documentazione pubblica resa disponibile sui singoli siti web, bibliografia.

L'analisi delle risposte ricevute è stata effettuata usando come "vagli" i guadagni contenuti nei due documenti di cui alle pagine precedenti:

- a) la visione strategica e
- b) l'analisi dei fabbisogni.

OhioLINK (USA)

Presentazione. L'Ohio Library and Information Network, OhioLINK²⁸, è un consorzio tra 89 biblioteche universitarie dello Stato dell'Ohio più la biblioteca di Stato dell'Ohio. Il consorzio da molti anni copia in archivi locali periodici e libri elettronici inclusi nei contratti negoziati con gli editori; i metadati e i contenuti vengono resi accessibili agli utenti autorizzati (light archive). Dei contenuti archiviati viene fatto un backup, ma le pratiche e le procedure oggi in uso sono considerate da OhioLINK non sufficienti ai fini dell'archiviazione vera e propria e per questo è in corso una revisione delle attività. Infatti, i due servizi in uso per l'archiviazione in locale, ossia Electronic Journal Center (EJC) ed Electronic Books Center (EBC) continuano a usare, sebbene con alcune variazioni e personalizzazioni, il software ScienceServer, ma proprio nel mese di marzo 2015 è stata pubblicata una Request for Information (RFI) per trovare nuove soluzioni software più rispondenti alle esigenze di archiviazione digitale. Non sono state fornite informazioni sui costi e sullo staff. L'Archivio è gestito dagli organi di governo del consorzio che è un consorzio di biblioteche.

Risposta alla survey. Non è stato compilato il questionario ma è intercorso uno scambio di comunicazioni di posta elettronica con Amy Pawlowski, Deputy Director di OhioLINK.

²⁷ Cfr. <<http://cariniana.ibict.br/>>

²⁸ Cfr. <www.ohiolink.edu>

Visione strategica. È interessante notare come OhioLINK, sin da quando ha iniziato a occuparsi di contratti consortili per lo Stato dell'Ohio per le risorse elettroniche, ha assunto una visione strategica rispetto al PCA come quella delineata nel presente rapporto, ossia riconoscendo la necessità assoluta di mantenere il controllo di dati e metadati per rendere i diritti di accesso negoziati effettivamente esercitabili.

Analisi dei fabbisogni. OhioLINK ha condiviso col gruppo un elenco di specifiche funzionali a cui la nuova soluzione di software deve rispondere. Le specifiche funzionali riflettono in diversi punti i fabbisogni riportati nell'analisi di cui alle pagine precedenti.

Considerazioni finali. Sia per la visione strategica sia per l'analisi dei fabbisogni i risultati del Gruppo 404 sono molto vicini agli assunti che muovono l'azione di OhioLINK in tema di Archivi digitali. È consigliabile rimanere in contatto con questa importante realtà e in particolare seguire gli sviluppi e le scelte che matureranno in seguito alla RFI, anche al fine della eventuale reingegnerizzazione di NERA.

Rede Cariniana (Brasile)

Presentazione. La rete Cariniana è un progetto di 4 anni (marzo 2014-dicembre 2018) che ha come obiettivo la conservazione a lungo termine delle pubblicazioni scientifiche brasiliane. Il focus principale riguarda i periodici Open Access pubblicati in Brasile. In una seconda fase, che sta per iniziare, si procederà con la conservazione delle tesi di dottorato. La terza fase riguarderà i dati scientifici da depositare nella rete. La fase sperimentale è durata un anno (2013). La rete è finanziata per intero dal Governo brasiliano e in particolare dal MCTI (Ministério da Ciência, Tecnologia e Inovação). Il MCTI dà risorse a IBICT (Instituto Brasileiro de Informação em Ciência e Tecnologia). L'ammontare del finanziamento è un dato riservato. La rete è composta ad oggi da 9 nodi: oltre a IBICT, altre 8 istituzioni. Le istituzioni che vogliono avere un nodo devono sottoscrivere il contratto predisposto da IBICT²⁹.

La rete usa il software open source LOCKSS e ha un contratto di consulenza e di assistenza con il personale di LOCKSS presso l'Università di Stanford. Lo staff di Cariniana si compone di 22 persone: 6 di IBICT (tre bibliotecari, uno sviluppatore, due sistemisti) e 16 provenienti dalle istituzioni aderenti, le quali devono mettere a disposizione ciascuna un bibliotecario e un sistemista. Delle 22 persone solo 1 è full time; 10 sono part time; 11 sono borsisti.

Le dimensioni dell'archivio sono ancora piccole rispetto allo storage: solo 157 GB per 601 ejournals OA, per 10.152 Archival Units (corrispondenti all'incirca alle annate di periodico) archiviate. Le dimensioni sono basse perché la rete archivia in prevalenza pdf e qualche foto di archivio istituzionale. Tutti i periodici elettronici archiviati usano OJS come piattaforma.

Le istituzioni servite dalla rete sono 114 tra università e centri di ricerca brasiliani. Il governo della rete è condiviso da MCTI e IBICT.

Risposta alla survey. Il questionario è stato compilato da Miguel Ángel Márdero Arellano, coordinator of the Brazilian Network of Digital Preservation Services.

Visione strategica. La visione strategica sottesa alla Rede Cariniana attiene in prima battuta alla conservazione a lungo termine (DP) delle pubblicazioni brasiliane, per cui trova corrispondenza con la visione strategica di cui al presente rapporto più per quanto attiene al deposito legale digitale e alla DP che non all'accesso perpetuo. Inoltre Cariniana usa LOCKSS e ha un accordo con Stanford per l'assistenza e lo sviluppo dei plugin. I brasiliani non scrivono i plugin in autonomia e per questo aspetto ci sono elementi di divergenza rispetto alla visione strategica delineata, che punta anche a far crescere le capacità autonome in Italia nel fare DP.

Analisi dei fabbisogni. Le 114 istituzioni sono servite in senso generale poiché l'archivio è un dark archive e non fornisce un accesso alternativo e concorrente: serve solo in caso di necessità. Ai fini dell'Italia questo è

²⁹ Molta documentazione interessante, ma in portoghese, si trova alla URL
<<http://cariniana.ibict.br/index.php/publicacoes/itemlist/category/8-documentacao>>.

un punto di vista leggermente differente, perché nel lavoro di questi mesi è emerso all'interno del Gruppo che la priorità assoluta viene data dall'analisi dei fabbisogni al PCA; inoltre, ai fini dell'accesso perpetuo, è importante dare da subito servizi perché le soluzioni di archivio, sebbene debbano lavorare in un'ottica di medio-lungo periodo, se vogliono trovare risorse devono essere in grado di dare anche servizi nel breve periodo, in modo tale da aumentare la consapevolezza dell'importanza strategica tra gli utenti e tra gli operatori dei servizi resi dagli archivi digitali. Collimano i fabbisogni invece rispetto a un approccio cooperativo che si faccia carico di fare DP e DLD delle pubblicazioni nazionali.

Considerazioni finali. Il progetto ha diversi elementi di interesse ai fini della DP e del DLD. Gli sviluppi riguardanti oltre i periodici OA, ossia per quanto attiene la conservazione delle tesi di dottorato e il deposito e la conservazione dei dati usati come base per le ricerche scientifiche, sono molto interessanti anche in considerazione dei nuovi bisogni di conservazione digitale che stanno emergendo in Italia e in Europa, in particolare per la conservazione dei dati grezzi usati a supporto delle ricerche.

Global LOCKSS Network GLN (USA)

Presentazione. Il programma LOCKSS, iniziato nel 1998, è una divisione di Stanford University Libraries. Esso è un sistema di conservazione digitale che usa un software open source, chiamato appunto LOCKSS, che ha ricevuto negli anni molti riconoscimenti e impieghi. Il software si basa sul principio della moltiplicazione delle copie e sul sistema di conservazione distribuito: infatti per poter funzionare esso ha bisogno di almeno sei server, la cui distribuzione deve seguire le migliori pratiche in uso³⁰. La rete si basa su "nodi" replicati tramite un protocollo peer to peer. Il software ha avuto un valutazione molto alta (punteggio di 5 su un massimo di 5) in occasione della certificazione dell'Archivio CLOCKSS da parte di CRL³¹.

Il progetto LOCKSS prevede la possibilità di costruire due tipi di rete di conservazione: la Global LOCKSS Network (GLN) e la Private LOCKSS Network (PLN).

Attraverso GLN le biblioteche possono conservare le collezioni di contenuti OA o TA di cui hanno un abbonamento. In tale contesto è possibile conservare collezioni di editori che partecipano alla rete. GLN è un Dim Archive.

Attraverso Private LOCKSS Network la biblioteche possono conservare le "collezioni speciali", ovvero dati digitali prodotti dalla biblioteca stessa. Tuttavia, dopo il successo della PLN CLOCKSS (che è un archivio indipendente da Stanford e che archivia i contenuti dei più importanti editori scientifici internazionali), la novità riguarda il possibile utilizzo di PLN ai fini del National Hosting e questo sviluppo potrebbe essere funzionale ai bisogni italiani.

Con il software LOCKSS è possibile salvaguardare contenuti di diversi tipi: libri e periodici elettronici; documenti governativi; repository istituzionali; database statici; tesi di dottorato; immagini; video; ecc.

L'infrastruttura LOCKSS permette una conservazione a lungo termine certificata, sostenibile e affidabile di conservazione dei propri materiali digitali. Presso Stanford sono impiegate 15 persone che lavorano sia per GLN sia per le PLN con cui ci sono accordi di supporto e di consulenza. La divisione di LOCKSS presso Stanford ha i conti in equilibrio, con un budget che è stato rivelato ma rimane dato confidenziale. I proventi derivano dalle quote versate dalle biblioteche che aderiscono a GLN; dai contratti con alcune PLN; da contributi governativi; da contributi di fondazioni private. Rispetto alla governance, ogni rete LOCKSS determina il sistema di governo che ritiene più adatto.

Risposta alla survey. Il questionario è stato compilato da Victoria Reich, Executive Director LOCKSS Program, Stanford University Libraries.

³⁰ *A Guide to Distributed Digital Preservation*, edited by Katherine Skinner and Matt Schultz, Atlanta, Educopia Institute, 2010; cfr. <<http://www.metaarchive.org/GDDP>>

³¹ Center for Research Libraries, *Certification Report on CLOCKSS*, luglio 2014; cfr. <http://www.crl.edu/sites/default/files/attachments/misc/CLOCKSS_Report_2014.pdf>

Visione strategica. L'approccio di LOCKSS è molto aderente alla visione strategica delineata dal Gruppo 404, poiché il programma ha l'obiettivo di creare una rete di biblioteche che partecipano attivamente alla salvaguardia del materiale digitale. Il modello promosso è di tipo associativo: punta a coinvolgere il maggior numero di enti con il principio che la conservazione deve essere uno sforzo collettivo, il cui mandato non deve essere affidato a soggetti singoli. La distribuzione dei contenuti e la condivisione dei costi e delle competenze, sia su base nazionale sia su base di cooperazione internazionale, oltre a creare una cultura distribuita della conservazione, dà migliori garanzie per la conservazione a lungo termine.

Analisi dei fabbisogni. Le istituzioni che usano GLN hanno un proprio server locale, che è parte della rete col quale possono fornire un accesso alternativo ai contenuti salvati. Infatti tutti gli oltre 500 editori che aderiscono a GLN accettano che GLN sia usato dalle istituzioni ai fini del PCA. È ad esempio previsto, per alcuni editori, che l'accesso post-termination avvenga su rete GLN. I punti di forza del sistema consistono nella distribuzione dei costi e delle competenze, i punti di debolezza riguardano la mancata adesione di alcuni editori importanti. Tale limite potrebbe essere superato dalla costruzione di una PLN per effettuare il National Hosting.

Considerazioni finali. La rete Global, a fronte di uno sforzo delle istituzioni per acquisire le competenze necessarie, supera il problema della concentrazione su un unico soggetto delle istanze di conservazione, modello che in passato ha incontrato problemi di sostenibilità; in particolare, la sperimentazione di una PLN potrebbe rispondere ai fabbisogni e alla visione strategica delineati nel presente rapporto.

SafeNet (UK)

Presentazione. È un progetto pilota di due anni (agosto 2014-luglio 2016) finanziato da JISC, che ha come obiettivo quello di costruire una infrastruttura nazionale (National Hosting) in grado di garantire l'accesso perpetuo ai periodici elettronici TA. Il progetto lavora perché il servizio entri in funzione dal primo agosto 2016 ed è la naturale continuazione di tre precedenti esperienze nate in JISC, ossia LOCKSS UK, PECAN e GLN in UK. Il progetto ha scelto di usare la tecnologia LOCKSS, cioè di usare una PLN, ai fini del National Hosting. Il gruppo di progetto si compone di 8 persone, di cui 6 provenienti da EDINA (che è una divisione dell'Università di Edimburgo, individuata da JISC come "National Data Center" e con cui JISC ha un contratto di fornitura) e 2 provenienti da JISC Collections. Il progetto è affiancato da un gruppo di bibliotecari chiamato a dare consulenza e assistenza al progetto. Partner del progetto è il team LOCKSS presso l'Università di Stanford, che offre a SafeNet consulenza e assistenza. I plugin sono creati da Stanford e da EDINA, che ha sviluppato negli anni scorsi le necessarie competenze per gestire in autonomia, sia dal punto di vista sistemistico sia di programmazione, una rete LOCKSS. Tuttavia, mentre Stanford rilascia sempre i sorgenti, per ora EDINA non li rilascia ma pensa di farlo in futuro. SafeNet invece sta sviluppando in autonomia l'Entitlement Registry, il registro dei diritti, che è un elemento software molto importante ai fini dell'uso della PLN per il National Hosting, perché è ciò che consente a una istituzione e ai suoi utenti autorizzati l'accesso sul server della PLN a un determinato contenuto, in quanto se ne ha diritto sulla base dei contratti sottoscritti. Il registro dei diritti che SafeNet sta sviluppando ha un tracciato più semplice di The Keepers Registry, viene popolato dalla banca dati ISSN a cui si aggiungono le informazioni sui diritti di accesso fornite dalle Biblioteche. Il progetto è concentrato sui soli periodici TA ma in futuro è facile prevedere che, se funzionerà, verrà esteso ad altri contenuti sia TA sia OA. Tuttavia la priorità resta quella dei periodici elettronici TA.

Risposta alla survey. Non è stato compilato il questionario ma è intercorso uno scambio di comunicazioni di posta elettronica con Robbie Ireland e Adam Rusbridge e si è tenuta una teleconferenza l'8 aprile 2015 con Peter Burnhill e Robbie Ireland.

Visione strategica. Anche in UK sta emergendo con forza l'importanza di controllare dati e metadati dei contenuti TA ai fini dell'accesso perpetuo, tanto che il progetto è stato intrapreso poiché la tutela dell'accesso perpetuo in particolare è sentito come prioritario dagli associati JISC. È anche rilevante

sottolineare come in UK ritengano importante assumere un approccio distribuito, con l'uso di un software open source e la formazione del personale, in modo da aumentare il proprio livello di autonomia ai fini del PCA e della DP.

Analisi dei fabbisogni. Per JISC il fabbisogno primario resta quello del PCA per i periodici elettronici TA. La DP e il DLD sono considerati bisogni importanti ma con meno priorità (forse anche perché si tratta di sfide coperte da altri attori, ad esempio per la DP le iniziative CLOCKSS e Portico; per il DLD la British Library). Il TDM non è incluso nel progetto.

Considerazioni finali. Il progetto SafeNet è molto importante ai fini dei fabbisogni italiani. I britannici sono interessati e disposti a cooperare con le altre iniziative di National Hosting, purché si tratti di realtà che si trovino a un livello di azione che consenta una proficua collaborazione. A questo fine è molto utile la collaborazione relativa all'Entitlement Registry. In particolare gli ultimi sviluppi del progetto hanno fatto emergere la preoccupazione riguardante gli editori medio-piccoli stranieri e questa preoccupazione sottolinea come quanto descritto nella visione strategica: la necessità di 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, è un punto di arrivo, analogamente ad altre esperienze estere³².

Scholars Portal (Canada)

Presentazione. Scholars Portal è stato creato nel 2001 dal consorzio di biblioteche universitarie della provincia canadese dell'Ontario, OCUL (Ontario Council of University Libraries), come archivio locale a vantaggio delle 21 Istituzioni che vi aderiscono. L'archivio usa un software proprietario, MarkLogic Server³³ e nel febbraio 2013, relativamente all'archiviazione dei periodici elettronici, ha ottenuto la certificazione da parte di CRL³⁴. Oltre ai periodici elettronici, Scholars Portal archivia anche: libri elettronici; dataset numerici; set di dati geospaziali. OCUL intende sottoporre a prossimo processo di certificazione la parte relativa ai libri elettronici. L'Archivio conserva sia periodici OA sia TA e, limitatamente ai periodici, ha raggiunto i 50 TB di storage con una crescita annua di 8 TB di cui 6 di dati e metadati, 2 TB di software. In totale l'archivio impiega 3 persone full time (di cui un bibliotecario e due informatici) e 6 persone part time per un totale di 3,9 FT: tutto personale interno di OCUL. A livello di governance, l'archivio è un servizio di OCUL il quale, ai fini di Scholars Portal, ha anche un accordo con l'University of Toronto Libraries. OCUL è una organizzazione affiliata al Council of Ontario Universities (COU)³⁵. Infatti il suo organo di governo è nominato da COU, e ne fanno parte i direttori delle biblioteche universitarie per il loro ruolo. L'archivio è in equilibrio economico dal 2006, quando i membri di OCUL hanno assunto la piena responsabilità dei costi di funzionamento di Scholars Portal, che sono costi diretti (staff, hardware, software) e indiretti (spese generali di funzionamento). OCUL ha anche reso noti i costi annuali di funzionamento di Scholars Portal. Per la maggior parte dei contenuti Scholars Portal funge da light archive, in quanto mette a disposizione i contenuti a testo completo con accesso alternativo e concorrente a vantaggio delle istituzioni e dei loro utenti autorizzati. Solo in alcuni casi i contratti di archiviazione che Scholars Portal ha con gli editori sono per il solo dark archive, ossia l'accesso può essere aperto agli utenti autorizzati solo in determinati casi. Con alcuni editori (come Project MUSE e JSTOR) non sono riusciti a ottenere i diritti di archiviazione in Scholars Portal. OCUL ha tenuto a precisare un punto molto importante. Scholars Portal non è ancora un Archivio nazionale per tutto il Canada ma è un servizio a vantaggio solo delle 21 istituzioni della provincia

³² Anche l'iniziativa tedesca che sta indagando quali progetti intraprendere per il National Hosting in Germania ha recentemente individuato nella combinazione di Portico con una PLN LOCKSS una possibile soluzione; cfr. <<https://opus4.kobv.de/opus4-bib-info/frontdoor/index/index/docId/1790>>.

³³ Cfr. <<http://www.marklogic.com/>>

³⁴ Center for Research Libraries. *Report on Scholars Portal Audit*, Febbraio 2013

<http://www.crl.edu/sites/default/files/attachments/pages/ScholarsPortal_Report_2013_%C6%92.pdf>

³⁵ Cfr. <<http://www.cou.on.ca/>>

dell'Ontario, che rappresentano circa il 40% della popolazione universitaria canadese, anche se Scholars Portal archivia la stragrande maggioranza dei pacchetti di periodici elettronici negoziati dall'ente che in Canada conduce le trattative nazionali, ossia il Canadian Research Knowledge Network (CRKN)³⁶. Infatti, è in corso una discussione per cui Scholars Portal diventi un archivio nazionale dando seguito alle conclusioni del gruppo di lavoro presso CRKN (una specie di Gruppo 404 sorto all'interno di CRKN), che aveva raccomandato nel febbraio 2012³⁷ che Scholars Portal diventasse l'archivio nazionale per tutti i 75 membri del CRKN.

Risposta alla survey. Il questionario è stato compilato da Kathy Scardellato, Ontario Council of University Libraries, Executive Director.

Visione strategica. La visione strategica che è alla base di OCUL e che ha portato alla creazione di Scholars Portal è nella sostanza la stessa delineata nel presente rapporto.

Analisi dei fabbisogni. Anche i fabbisogni dei membri di OCUL ai quali si è inteso far fronte con Scholars Portal sono gli stessi: l'accesso perpetuo come priorità, seguito da altri bisogni parimenti condivisibili. È anche considerato essenziale poter effettuare le attività di TDM sui contenuti dell'Archivio, e a questo fine si stanno concludendo accordi separati con gli editori per fornire garanzie in tal senso agli utenti di OCUL.

Considerazioni finali. L'Archivio Scholars Portal è un esempio riuscito di Local Loading che sta evolvendo in National Hosting. Esso dà servizi quotidianamente agli utenti autorizzati delle istituzioni che lo hanno fondato e sono servizi molto usati e apprezzati. È importante mantenere i rapporti con OCUL e in prospettiva con CRKN, per seguire le evoluzioni dell'archivio e per scambiarsi esperienze, opinioni e informazioni.

Portico (USA)

Presentazione. Portico, insieme con CLOCKSS, è uno degli archivi più riconosciuti a livello internazionale per la DP; a differenza di CLOCKSS, per gli editori che lo richiedono³⁸, Portico offre anche servizi di accesso perpetuo. Pertanto Portico è un dark archive per alcuni editori e un dim archive per la maggior parte degli editori aderenti. Portico è uno dei tre servizi forniti dall'organizzazione ITHAKA³⁹, ente senza scopo di lucro di diritto USA. L'Archivio, nato nel 2005, usa una varietà di software, tuttavia esso si basa principalmente su due software proprietari (non si è autorizzati a rivelare i nomi per motivi di riservatezza); in aggiunta fa ricorso ad alcuni software open source. L'Archivio conserva principalmente: periodici elettronici; libri elettronici; digitalizzazioni di collezioni analogiche curate e distribuite da editori commerciali. Portico usa una strategia di conservazione basata sulla normalizzazione con migrazione dei formati, conservando però sempre anche i file originali. L'Archivio è impegnato sia nella conservazione dei contenuti TA sia dei contenuti OA. Portico conserva quattro copie dei contenuti archiviati di cui una in *cloud* e una in un *external hard drive* presso la Biblioteca Nazionale Olandese. Rispetto al personale, Portico è gestito con personale interno, con personale a contratto, mediante accordi di fornitura con fornitori specializzati. Sebbene il complesso modello organizzativo di ITHAKA, il ricorso a collaboratori esterni e ad accordi con fornitori non abbia reso semplice per Portico fornire una risposta sullo staff, si può tuttavia affermare che l'Archivio è gestito in totale da circa 21,5 persone FT direttamente dedicate alle sue attività. Il servizio è in equilibrio economico (sono state fornite anche informazioni circa il costo annuo di funzionamento ma si tratta di informazioni coperte da clausole di confidenzialità), fondato sulle quote ricevute dalle biblioteche e dagli editori aderenti. Rispetto alla governance, essendo un servizio di ITHAKA, Portico è governato dal Consiglio

³⁶ Cfr. <<http://crkn.ca>>

³⁷ Cfr. <http://crkn.ca/sites/default/files/file/PATG%20report_OCUL%20comments.pdf>

³⁸ Infatti, non tutti gli editori che aderiscono a Portico accettano di dare servizi anche di PCA, ad esempio il caso di Cambridge University Press.

³⁹ Cfr. <<http://ithaka.org/>>

di amministrazione di ITHAKA, che è affiancato da un organismo che ha compiti consultivi (Advisory Committee): quest'ultimo è composto da rappresentanti degli enti aderenti. Portico ha dichiarato che è disposto a collaborare con eventuali iniziative di National Hosting come pure a iniziative di repliche locali dell'Archivio, sebbene queste collaborazioni dipendano anche dal rispetto di requisiti tecnici, di policy e di aspetti legali relativi al paese ospitante. Fino ad oggi Portico ha intrapreso iniziative di collaborazione con altre importanti istituzioni straniere, ma più sul versante del DLD e mai sul versante del PCA. A questo fine collabora con la British Library e a fine maggio 2015 è stata annunciata una collaborazione con la Biblioteca Nazionale Olandese per l'Archivio e-Depot⁴⁰. Ai fini del TDM, Portico non offre servizi di questo tipo ma ha istituito un'area di studio al suo interno per verificare se includere questo tipo di attività tra i servizi da offrire in futuro. Considerata l'attenzione del Gruppo 404 rispetto alla certificazione, è importante riferire che Portico è stato il primo archivio ad essere certificato da CRL nel 2010⁴¹.

Risposta alla survey. Il questionario è stato compilato da Kate Wittenberg, Managing Director of Portico.

Visione strategica. Le organizzazioni che hanno fondato ITHAKA, la quale poi ha lanciato Portico come nuovo servizio, hanno condiviso una visione strategica simile a quella elaborata nel presente rapporto. Hanno tuttavia ritenuto di dover fondare un'organizzazione separata, in grado di gestire servizi avanzati di documentazione in era digitale a servizio della ricerca e della didattica e finanziando questi servizi attraverso la raccolta di quote da parte di biblioteche ed editori che ne volevano usufruire, per cui raccogliendo quote oltre i fondatori di ITHAKA. È un esempio di come una innovazione organizzativa ha consentito innovazioni successive e incrementali in termini di prodotti e di servizi.

Analisi dei fabbisogni. L'analisi che è stata fatta da ITHAKA quando ha deciso di lanciare l'Archivio Portico ha portato a individuare due fabbisogni centrali, ossia la DP e la PCA, e in particolare di garantire alle istituzioni l'accesso perpetuo per i contenuti per i quali hanno acquistato, di solito a caro prezzo, dei diritti di accesso.

Considerazioni finali. Portico resta una realtà riuscita e di riferimento nell'ambito dei fabbisogni di DP e di PCA. Il modello tradizionale di business di Portico non è compatibile con la visione strategica delineata dal Gruppo 404 rispetto al controllo dei metadati e dei contenuti delle pubblicazioni scientifiche da parte delle istituzioni italiane titolari dei diritti. È consigliabile mantenere un contratto nazionale con Portico alle migliori condizioni possibili per le istituzioni italiane come opportunità integrativa ma assolutamente non sostitutiva al National Hosting. A questo fine è importante sottolineare come Portico, rispondendo al questionario, ha dichiarato una disponibilità a collaborare per le iniziative di National Hosting e questa dichiarazione è interessante e sicuramente inedita, poiché Portico si è presentato sino ad oggi con un modello di business unico e semplice: le biblioteche interessate pagano Portico come un outsourcer per ricevere servizi. Questa disponibilità potrebbe essere indagata ad esempio ai fini di una reingegnerizzazione di NERA.

ISTEX (FRANCIA)

Presentazione. Il progetto ISTEX è iniziato ufficialmente nell'aprile 2012. Obiettivo è quello di fornire accesso online alle collezioni d'archivio di letteratura scientifica di tutte le discipline all'intera comunità di ricerca e di didattica superiore francesi. Base del progetto è la costruzione di una politica nazionale per l'acquisizione dei documenti che copra ogni tipo di materiale: archivi di riviste, database, libri elettronici. Il progetto ISTEX fa parte del programma del Ministero francese dell'Istruzione (MESR) "Investissements d'Avenir", che ha l'obiettivo di rafforzare a livello mondiale il ruolo della Francia nell'ambito della ricerca e

⁴⁰ Cfr. <http://www.portico.org/digital-preservation/news-events/news/national-library-of-the-netherlands-portico-partnership-at-the-forefront-of-digital-preservation-and-international-collaboration?utm_source=internal&utm_medium=email&utm_campaign=portico_kbnewsrelease_052815>

⁴¹ Cfr. Center for Research Libraries, *Report on Portico Audit Findings*, 10 gennaio 2010
<<http://www.crl.edu/sites/default/files/attachments/pages/CRL%20Report%20on%20Portico%20Audit%202010.pdf>>

della didattica. Il progetto ha quattro partecipanti: CNRS (Centre National de la Recherche Scientifique), ABES (l'Agence Bibliographique de l'Enseignement Supérieur), Couperin (Consortium Universitaire de Publications Numériques) e Università della Lorena per conto della CPU (Conférence des Présidents d'Université).

Il progetto prevede due fasi:

- 1) l'acquisizione dei contenuti tramite la stipula di licenze nazionali;
- 2) la creazione di una piattaforma per ospitare tutti i contenuti e i metadati delle risorse acquisite e per fornire l'accesso (inizialmente assicurato sulle piattaforme degli editori) alla comunità scientifica e accademica francese.

Il progetto è stato finanziato per intero dal Governo francese con 60 milioni di euro, di cui 55 destinati all'acquisizione delle risorse elettroniche e 5 destinati alla creazione della piattaforma vera e propria. La governance del progetto ha due livelli. Un Comitato esecutivo, composto da rappresentanti dei quattro enti coinvolti nel progetto e da un rappresentante del Ministero; un Comitato direttivo, composto da rappresentanti delle università, delle istituzioni di ricerca, delle Grandes Écoles, del Ministero. Questo doppio livello è stato scelto per garantire sia la presenza di specialisti della documentazione, sia quella di persone con responsabilità politiche di alto livello nelle università, nei centri di ricerca, ecc. Tuttavia il ruolo centrale è rivestito dal Ministero, che nomina i rappresentanti dei Comitati. Lo sviluppo del software è affidato ad INIST, che ricorre anche a software open source, ad esempio per il motore di ricerca usa ElasticSearch. Punti di forza del progetto sono: l'uguaglianza territoriale dei diritti di accesso, la multidisciplinarietà, l'accesso unificato a tutti i contenuti e la possibilità di stipula di contratti a condizioni vantaggiose (rispetto alla contrattazione ente per ente).

Risposta alla survey. Il questionario è stato compilato da André Dazy, Couperin, Coordinator of the department services and forecasting studies

Visione strategica. Interessante, dal punto di vista della conservazione nazionale, è l'unione indissolubile dei due aspetti del progetto: l'acquisizione dei contenuti per i quali il Governo ha messo a disposizione risorse importanti è possibile solo se l'editore accetta che i dati e i metadati dei contenuti acquistati siano conservati nella piattaforma nazionale. La piattaforma di archiviazione nazionale è centrale e particolarmente importante è il supporto del Ministero, ente finanziatore del progetto, che rafforza la connotazione di progetto nazionale e che ritiene strategica al più alto livello la creazione e la manutenzione di un archivio digitale nazionale a servizio della comunità accademica e di ricerca francese.

Analisi dei fabbisogni. Alla definizione dei fabbisogni e dei requisiti è stata data una particolare rilevanza attraverso un sondaggio nazionale diretto a tutti i ricercatori francesi, in modo tale da acquisire una conoscenza precisa delle esigenze della comunità a cui il progetto è finalizzato. Per il progetto francese il fatto che l'unico modo di garantirsi efficacemente i diritti di accesso perpetuo negoziati sia quello di predisporre un National Hosting è un guadagno ormai consolidato da tempo (e già praticato attraverso la piattaforma PANIST, finalizzata al PCA). Rispetto all'analisi dei fabbisogni di cui al presente rapporto, i francesi hanno individuato molti altri fabbisogni e servizi a valore aggiunto che la piattaforma deve soddisfare. Non solo deve consentire il TDM, ma la piattaforma documentale deve classificare i contenuti all'interno di ontologie semantiche; deve consentire ai ricercatori di sapere chi sta lavorando su un determinato argomento e con chi; ecc. Infine, la piattaforma deve essere interoperabile con gli altri progetti in corso in Francia⁴².

Considerazioni finali. Il progetto ISTEEX è forse il progetto più ambizioso di quelli interessati dalla presente indagine. Esso riveste particolare interesse sia per come sono stati impostati e affrontati gli aspetti organizzativi e contrattuali, sia per il modello di finanziamento. Infatti dei progetti analizzati, solo il progetto francese e quello brasiliano sono finanziati interamente dai rispettivi Governi. Interessante anche

⁴² Cfr. André Dazy, *ISTEX: a powerful project for scientific and technical electronic resources archives*, in Insights, November 2014, pp. 269-273.

il concetto di accesso paritario per tutta la comunità scientifica e didattica francese, nell'ottica di uno sviluppo culturale collettivo dell'intera nazione. I francesi ritengono interessante un'iniziativa di archivi digitali a livello europeo, anche se le questioni giuridiche, per certi versi differenti da paese a paese, possono rappresentare un ostacolo importante.

Raccomandazioni

Se il controllo dei contenuti e dei metadati delle pubblicazioni scientifiche è strategico, in particolare per le risorse TA per cui si è acquistato un diritto di accesso perpetuo, allora una strategia nazionale sulla conservazione e un'iniziativa di National Hosting sono necessarie.

La soluzione di archivio NERA presenta oggi due limiti molto importanti.

- a) Non funziona in modo adeguato e richiede, perciò, sostanziali adeguamenti sia nelle scelte tecnologiche adottate sia nell'organizzazione della produzione.
- b) Il modello di governance adottato, cioè il fatto che il consorzio CINECA abbia il totale controllo dell'archivio e che non ci siano organismi d'indirizzo, di controllo, di analisi sulla sostenibilità che facciano riferimento ai titolari dei diritti di accesso perpetuo, diritti acquisiti a costi spesso elevati nel corso di oltre quattordici anni di contratti consortili.

Il destinatario delle raccomandazioni contenute in questo documento è l'intero sistema accademico nazionale; anche il MIBACT è destinatario delle raccomandazioni per quanto inerisce al Deposito Legale Digitale. Il fatto che la negoziazione nazionale è stata centralizzata da CRUI rende CRUI il primo luogo associativo a cui rivolgersi perché queste raccomandazioni trovino seguito. Tuttavia l'importanza di testare in tempi brevi le soluzioni proposte dal Gruppo 404 non esclude che qualche altro attore (principalmente atenei), o aggregato di attori, non possa essere disponibile a fare partire anche da subito un progetto di sperimentazione in merito, anche in raccordo con CRUI.

In questa configurazione anche CINECA potrebbe mantenere il suo ruolo di importante attore per il National Hosting purché:

- 1) mantenga un ruolo di partner tecnologico e operi sulla base di specifiche tecnologiche ed esigenze definite dalle Università;
- 2) si aumentino e si utilizzino le competenze tecnologiche all'interno degli atenei ai fini degli archivi digitali affidabili, poiché l'importanza di detti archivi nell'immediato futuro aumenterà di molto la sua centralità (ad esempio la conservazione degli Institutional Repositories; la conservazione dei dati grezzi della ricerca; la conservazione degli snapshot a una certa data delle banche dati citazionali; ecc.) tale per cui è necessario rendere determinate competenze maggiormente pervasive all'interno dei sistemi bibliotecari e documentali delle università e dei centri di ricerca italiani;
- 3) si definisca un modello organizzativo e una governance distribuita all'interno dei quali le decisioni spettano in definitiva ai titolari dei diritti, che sono principalmente le Università.

Il Gruppo 404 ritiene, in linea generale, che le due ipotesi più percorribili siano le seguenti:

- A) creare un'organizzazione *ad hoc* tra le biblioteche italiane universitarie e di ricerca, simile a OCUL che è consorzio tra biblioteche universitarie. Questa sarebbe un'importante innovazione organizzativa, in grado di incubare e generare nuove infrastrutture tecnologiche e organizzative a servizio della ricerca e della didattica.
- B) lavorare a livello di CRUI con un coinvolgimento diretto di specialisti, provenienti dalle diverse università (come in essere col gruppo CARE per le trattative nazionali) e lasciando aperta ogni evoluzione organizzativa alle nuove necessità che si andranno a consolidare. Sulla base di queste considerazioni si raccomanda quanto segue.

Raccomandazione n° 5. Verificare e concordare la possibile evoluzione tecnologica dell'Archivio NERA con la quantificazione dei relativi costi per valutarne la capacità di rispondere adeguatamente alle effettive

esigenze qualitative del servizio, l'integrazione con altri servizi gestiti dagli Atenei e la sostenibilità economica.

Precisazione. Dal punto di vista tecnologico le disponibilità di Scholars Portal, Portico, ISTEEX, Global LOCKSS Network, acquisite durante i lavori del Gruppo 404, possono essere la base da cui ripartire per una reingegnerizzazione dell'archivio. La collaborazione con Scholars Portal potrebbe essere particolarmente significativa in quanto potrebbe aiutare non solo per le scelte tecnologiche e per l'organizzazione della produzione, ma anche per il recupero dei contenuti, in quanto Scholars Portal ha avuto una genesi simile a quella di NERA, ossia è una migrazione di ScienceServer.

Tempi. Produrre una proposta entro 15 mesi dall'inizio della verifica.

Raccomandazione n° 6. Intraprendere l'esperienza di una PLN LOCKSS con sei università italiane che diano sufficienti disponibilità e garanzie. Questo è molto auspicabile che avvenga in coordinamento con CRUI in quanto riferimento per le trattative nazionali.

Precisazione. Ciascuna delle sei università, da individuare preferibilmente tra gli atenei italiani che già aderiscono all'archivio CLOCKSS che usa la medesima tecnologia e che sono già in contatto con lo staff del programma LOCKSS dell'Università di Stanford, metterà a disposizione:

- a) l'hardware da inserire nella propria rete;
- b) il personale (un bibliotecario, un sistemista, uno sviluppatore, tutti part time).

Se interessata, CRUI metterà a disposizione un budget da cui attingere per le necessità legate all'iniziativa.

Tempi. Realizzare entro dodici mesi dall'inizio delle attività l'accesso perpetuo alternativo rispetto all'accesso su server dell'editore per almeno un editore estero importante che non riversa i dati in NERA e li riversa in GLN (ad esempio uno tra Taylor and Francis, Sage, OUP e simili) ovvero per un editore come Springer (che riversa in NERA e aderisce anche a GLN) e per almeno un editore italiano (ad esempio uno tra Il Mulino, Franco Angeli, Vita e Pensiero, ecc.). Nel caso dell'editore italiano quest'azione, sempre entro i dodici mesi, deve sperimentare il popolamento in automatico del deposito legale digitale di Magazzini Digitali, purché la BNCF garantisca il necessario supporto. Nel caso degli editori stranieri si negozi con GLN e CLOCKSS un accordo per lo scambio dei plugin in modo da non dover replicare il lavoro già fatto da altre reti LOCKSS, creando economie di scopo che deriverebbero da un ecosistema a livello internazionale che usa la medesima tecnologia LOCKSS.

Raccomandazione n° 7. Mantenere le relazioni con gli altri archivi nazionali e internazionali in modo da realizzare uno scambio di esperienze e di best practice.

Precisazione. Particolarmente interessante potrebbe risultare la collaborazione con i britannici di SafeNet (ad esempio per l'Entitlement Registry) e con i francesi di ISTEEX. Tendere a una collaborazione con i partner europei è auspicabile anche per configurare infrastrutture interoperabili in grado di dare servizi ai cittadini europei con conseguente accesso a fondi dell'Unione Europea.

Raccomandazione n° 8. Ai fini del finanziamento, cercare fondi anche oltre le risorse messe a disposizione eventualmente da CRUI-CARE, cogliendo eventuali opportunità offerte da bandi (pubblici o privati) nazionali e /o europei, e/o internazionali.

Precisazione. Con la migrazione dai formati analogici ai formati elettronici distribuiti via web, ai fini di una migrazione sostenibile e professionalmente adeguata, è importante sottolineare come al finanziamento degli archivi digitali devono essere indirizzate, almeno in parte, le spese previste nel mondo analogico per acquisto-affitto di magazzini; acquisto della scaffalatura; rilegatura; spolveratura, ecc. La spesa per la

conservazione dei documenti nel passaggio dall'analogico al digitale deve essere riqualificata ma non deve essere soppressa, se non al prezzo di scelte errate sotto il profilo professionale e gestionale.

Raccomandazione n° 9. Continuare (CRUI-CARE) a offrire la possibilità di un contratto nazionale alle migliori condizioni per le università e gli enti di ricerca italiani per l'Archivio Portico, in quanto garanzia ulteriore e non sostitutiva del National Hosting ai fini della tutela dei diritti di accesso perpetuo da offrire alle Università italiane interessate. Negoziare (CRUI-CARE) un accordo per una partnership strategica anche con l'Archivio CLOCKSS, sia per rafforzare la presenza della comunità italiana (biblioteche ed editori) in questo importante archivio, sia per esperire possibili integrazioni con altri archivi digitali italiani, ad iniziare dal deposito legale digitale della Biblioteca Nazionale Centrale di Firenze (BNCF).

CONCLUSIONI

La soluzione migliore per dare corso alla visione strategica delineata e alle raccomandazioni che il lavoro di questi mesi ha suggerito si ritiene possa essere quella per cui CRUI-CARE, dando un seguito ai lavori del Gruppo 404, costituisca un team specialistico *ad hoc* e dia vita a un progetto sperimentale della durata indicativa di ventiquattro mesi (2016-2017), con un budget minimo da assegnare al team che potrebbe essere attinto, se ve ne siano i margini, da risorse ricavate dai budget 2015-2017 dell'Area risorse elettroniche della Fondazione CRUI, ovvero attraverso l'impiego di altre risorse che CRUI-CARE riuscirà a mobilitare, ovvero partecipando a bandi di organizzazioni private (ad esempio Fondazioni bancarie) e pubbliche, ovvero dalla combinazione di diverse fonti di finanziamento.

La fase sperimentale, che può assumere i tempi indicati nel paragrafo precedente, deve portare a una relazione finale sull'attività svolta e sui servizi resi durante la fase di progetto e dovrà essere corredata da evidenze circa il livello di soddisfazione e di interesse da parte delle università che hanno usato i servizi, in modo tale da mettere CRUI-CARE nelle condizioni di decidere l'assetto organizzativo e di finanziamento finale da dare al National Hosting in cui anche i singoli Atenei, soprattutto i più grandi e attrezzati, siano chiamati a fare la loro parte in un quadro di cooperazione nazionale.

Se si riuscirà a realizzare soluzioni di National Hosting come quelle delineate, ossia in grado anche di contribuire al deposito legale digitale, il finanziamento del National Hosting non potrà che basarsi almeno su tre pilastri: il contributo di CRUI, prevedendo risorse dal budget dell'Area Risorse elettroniche a partire dal 2018; il contributo delle singole istituzioni, anche non associate CRUI, che scelgono di avvalersi del servizio; il contributo governativo per i servizi relativi al deposito legale digitale (MIBACT) e per il sostegno che il National Hosting, come parte fondamentale del sistema documentale, offre alla comunità accademica e scientifica Italia (MIUR) nello spirito delle stesse previsioni di legge⁴³.

⁴³ Art. 4 comma 3 del decreto legge n° 91 dell'8 agosto 2013, convertito con modificazioni dalla Legge 7 ottobre 2013, n. 112: "Al fine di ottimizzare le risorse disponibili e di facilitare il reperimento e l'uso dell'informazione culturale e scientifica, il Ministero dei beni e delle attività culturali e del turismo ed il Ministero dell'istruzione, dell'università e della ricerca adottano strategie coordinate per la piena integrazione, interoperabilità e non duplicazione delle banche dati rispettivamente gestite, quali quelle riguardanti l'anagrafe nazionale della ricerca, il deposito legale dei documenti digitali e la documentazione bibliografica".

ALLEGATI

Allegato 1. 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 contratti⁴⁴

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 Global LOCKSS Network.

Accesso perpetuo

Ai fini dell'accesso perpetuo (o perpetual access o post termination access o post cancellation 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 Global LOCKSS Network 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.

È 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 vigenza 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

⁴⁴ 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.

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.

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.

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 15.4.2004 ("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).

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 LOCKSS 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.

APPENDICI

Appendice A - Estratti dai modelli di licenza nazionale

Questa appendice contiene gli estratti delle clausole relative alla Conservazione digitale a lungo termine e accesso perpetuo e le clausole relative alle attività di text e data mining contenute nella versione più recente dei modelli di licenza in uso presso alcune iniziative che si occupano di trattative nazionali. Sono stati presi in esame le licenze di:

- 1) NESLi2- Jisc
- 2) ISTEEX-Panist
- 3) Working Group National Licensing
- 4) CRKN/RCDR

NESLi2-Jisc

Le clausole sono estratte dal modello inglese di licenza NESLi2 for e-journals, *Jisc Collections Model Journal Licence 2015*⁴⁵, di cui esistono modelli in uso dal 2004. La versione dell'estratto è quella proposta dal 2013.
(...)

SCHEDULE 4: NESLI2 LICENCE FOR JOURNALS - ESTRATTO

(...)

BACKGROUND

(...)

3. This Licence is based on the PA/Jisc and NESLI model licence for journals

3. PERMITTED USES

(...)

- 3.1.6 allow Authorised Users for Educational Purposes only to:

(...)

3.1.6.7 save and/or deposit in perpetuity parts of the Licensed Material in electronic repositories operated by the Institution and/or by an Authorised User on a Secure Network. Access to and use of such repositories shall be limited to Authorised Users; and

3.1.6.8 use the Licensed Material to perform and engage in text mining/ data mining activities for academic research and other Educational Purposes and allow Authorised Users to mount,

⁴⁵ Cfr. <<https://www.jisc-collections.ac.uk/nesli2>>. L'archivio dei modelli di licenza è consultabile alla URL <<https://www.jisc-collections.ac.uk/journals/Journals-Model-Licence/>>. Il *Jisc Collections Model Journal Licence* è reperibile alla URL <https://www.jisc-collections.ac.uk/Documents/Model_Licences/2015/Jisc_Collections_Model_Journals_Licence_2015.docx> ed è parte del Jisc Collections Contract, reperibile alla URL <[https://www.jisc-collections.ac.uk/Documents/Model_Licences/Model_NESLi2_Contract_2013_\(JOOS_and_AAU\).doc](https://www.jisc-collections.ac.uk/Documents/Model_Licences/Model_NESLi2_Contract_2013_(JOOS_and_AAU).doc)>. Una guida ai contenuti e tipologie del *JISC Model License* è consultabile alla url <<http://www.jisc-collections.ac.uk/Help-and-information/How-Model-Licences-work/Guide-to-Model-Licence/>>.

load and integrate the results on a Secure Network and use the results in accordance with this Licence.

(...)

8. TERM AND TERMINATION

(...)

8.5 Upon termination of this Licence (except for a material breach by the Institution of its obligations under this Licence), the Publisher will provide (at the option of the Institution) the Institution and its Authorised Users with continuous access to and use of the full text of the Licensed Material which was published and paid for during the term of this Licence and preceding licences (where applicable) between the Publisher and the Institution, without charge, either by one or more of the following options:

- i) continuing online access to archival copies of the same Licensed Material on the Publisher's server;
- ii) supplying archival copies of the same Licensed Material to the Institution in an electronic medium mutually agreed between the parties;
- iii) supplying archival copies of the same Licensed Material to a central archiving facility operated on behalf of the UK HE/FE community or other archival facility;
- iv) granting access to the same Licensed Material through one of the e-journals archiving solutions as listed in Schedule 3.

For the avoidance of doubt access and use of archival copies shall be subject to the terms and conditions as set out in Clauses 3 and 4 of this Licence.

8.6 The Institution is permitted to:

- i) mount the archival copies of the Licensed Material supplied by the Publisher in accordance with Clauses 8.5 (ii) and 8.5 (iii)
- ii) communicate, make available and provide access to such archival copies via a Secure Network to Authorised Users in accordance with the terms of this Licence; and
- iii) make copies or re-format the Licensed Material contained in the archival copies in any way to ensure their future preservation and accessibility in accordance with this Licence.

8.7 In the event that ownership of a part or parts of the Licensed Material is sold by the Publisher or otherwise transferred to another publisher, the Publisher will use all reasonable efforts to retain a non-exclusive copy of the volumes published during the Subscription Period and make them available without charge to the Institution:

- i) through the Publisher's server; or
- ii) by supplying such material without charge to the Institution in accordance with the procedure described in Clause 8.5.

8.8 In the event that the Publisher ceases to publish a part or parts of the Licensed Material (including back issues of a title as part of the Licensed Material), the Publisher will:

- i) maintain a digital archive of such Licensed Material;
- ii) make the digital archive available to the Institution without charge through the Publisher's server, via a third party server (including e-journals archiving initiatives as listed in Schedule 3 or by supplying the digital archive to the Institution without charge in accordance with

the procedure described in Clause 8.5.

- 8.9 The archival copies supplied in accordance with Clause 8.5 (ii) and (iii) will contain all textual content of the Licensed Material but will not contain all links and other features and functionality associated with the online version available via the Publisher's server. Access and use of such copies will be free of charge.

(...)

- 8.11 Upon termination of this Licence, copies of parts of the Licensed Material made by the Institution or Authorised Users may be retained. Such copies may be used after termination of this Licence subject to the terms of Clauses 3 and 4 of this Licence, which terms shall survive any termination of this Licence.

9. ACKNOWLEDGEMENT AND PROTECTION OF INTELLECTUAL PROPERTY RIGHTS

(...)

- 9.2 For the avoidance of doubt, the Publisher hereby acknowledges that any database rights created by the Institution as a result of local mounting of the Licensed Material as referred to in Clause 8.6 shall be the property of the Institution.

- 9.3 For the avoidance of doubt, the Publisher hereby acknowledges that any database rights created by Authorised Users as a result of textmining/datamining of the Licensed Material as referred to in Clause 3.1.6.8 shall be the property of the Institution.

(...)

SCHEDULE 3– INDUSTRY STANDARDS

The Publisher agrees to implement the following industry standards to enhance access and use of the Licensed Material.

The Publisher will:

(...)

3. Archive the Licensed Material to ensure that it is preserved for future scholarship in at least one of the following archiving solutions: Portico, Clockss or Lockss, and inform the Institution in which of the archiving solutions the Licensed Material may be found;

[DETAIL HERE THE ARCHIVING SOLUTIONS WHERE THE LICENSED MATERIAL HAS BEEN DEPOSITED]

(...)

6. Use all reasonable efforts to keep to the Code of Practice of Project Transfer (<www.projecttransfer.org>) to ensure that journal content remains easily accessible by the Institution and its Authorised Users when there is a transfer of material between parties, and to ensure that the transfer process occurs with minimum disruption; and
7. Provide link-resolver vendors and other library systems suppliers quarterly with full details of the Licensed Material in accordance with the most current KBART standard (<<http://www.uksg.org/kbart/s5/guidelines>>); and also with related data of (i) the first and final year, volume, and issue and (ii) the algorithm or syntax for constructing an article-level link from an article's metadata within the Licensed Material

Le clausole sono estratte dal contratto tipo del progetto francese ISTEX, dell'Agence Bibliographique de l'Enseignement Supérieur, versioni del 15.10.2012 e 15.1.2015. Sono definite le condizioni previste nel progetto per la tutela dei PCA e i requisiti di fornitura dei dati e metadati per l'accesso sulla piattaforma ISTEX dei contenuti acquistati⁴⁶.

AGENCE BIBLIOGRAPHIQUE DE L'ENSEIGNEMENT SUPERIEUR CONTRACT LICENCE TYPE - ESTRATTO

(...)

Article 1. Préambule

(...)

4. Le projet ISTEX est un programme d'acquisition d'archives scientifiques dont le but est de créer une plateforme numérique aux meilleurs standards internationaux et accessible à partir de toutes les bibliothèques et établissements d'enseignements supérieurs français, ci-après dénommée « la Plateforme ISTEX ». La Plateforme ISTEX a notamment pour objectifs de proposer, outre l'accès aux données et métadonnées disponibles, des services à valeur ajoutée basés sur le traitement des données en texte intégral (par exemple et de manière non limitative, interrogation en texte intégral sur les objets numériques indexés dans leur totalité, production de synthèse documentaire par analyse de sous-corpus individualisés pour l'occasion et auxquels sont appliquées des méthodes de texte-mining...). L'opérateur chargé de la mise en place de la Plateforme ISTEX dans le cadre de ce projet est le Centre National de Recherche Scientifique (CNRS).

(...)

Article 5. Durée – Entrée en vigueur

(...)

5.2. Durée du droit d'accès à la base de données

14. L'Editeur concède un Droit d'accès distant à sa Base de données, à titre gratuit, pour une durée de cinq (5) ans.

(...)

Article 7. Données

7.1. Modalités d'accès

(...)

30. Les Métadonnées seront livrées en XML et une DTD ou un schéma de données y seront associés. Elles devront inclure l'intégralité des informations bibliographiques disponibles. L'Editeur s'engage à fournir toutes les éventuelles mises à jour des Métadonnées sur une base mensuelle le cas échéant.

31. Les données (texte intégral) devront être livrées dans les formats utilisés par l'éditeur pour sa propre plate-forme. Elles seront fournies en XML et une DTD ou un schéma de données y seront associés. Si

⁴⁶Cfr. <<http://www.istex.fr/submitted-offer-2/>> e
<http://www.istex.fr/wp-content/uploads/2015/02/2015_Licence-type-ISTEX.pdf>.

l'éditeur ne dispose pas de données en format XML pour sa propre plate-forme, les données pourront être fournies en PDF mode texte uniquement.

33. L'Editeur reconnaît que les Données et Métadonnées livrées sont dans un format manipulable permettant l'exercice des Droits Conçédés décrits à l'article 9.2 de la présente Licence. Les caractères doivent utiliser le codage UTF8.

34. Données et Métadonnées seront livrées via un serveur FTP ou sur un support physique. L'Editeur s'engage à fournir la Documentation afférente aux Données et Métadonnées.

7.2. Hébergement des données

35. Dans le cadre du projet ISTEX, les Données seront hébergées sur la Plateforme ISTEX et accessibles sur cette Plateforme via internet par les Utilisateurs autorisés.

36. L'hébergement des Données sera assuré par le CNRS.

7.3. Normes et protocoles

37. Afin de faciliter l'échange de Données et notamment afin de faciliter la recherche des informations dans la Base de données l'Editeur s'engage à respecter les normes et protocoles référencés en annexe « Modalités d'accès ».

7.4. Conservation

38. Dans le cadre de la présente licence, l'Editeur s'engage à participer à un programme de conservation des Données à l'échelon international de type Portico ou LOCKSS (Lots of Copies Keep Stuff Safe) et à charger des copies de sauvegarde dans ces archives permanentes.

39. Les Données devront faire l'objet d'un archivage pérenne par l'Editeur en utilisant les meilleures techniques en vigueur.

40. L'ABES est autorisée à établir deux (2) copies de sécurité sous format analogique ou numérique et à les faire héberger par les institutions de son choix, sous réserve d'en informer préalablement l'Editeur, et conformément aux termes définis dans le présent contrat. (...)

Article 9. Droit de propriété

(...)

48. Les droits spécifiques concédés dans le cadre de la mise en place de la plateforme ISTEX sont les suivants :

- concéder au CNRS le droit de charger les Données et les Métadonnées, communiquer les Données et les Métadonnées via un réseau sécurisé et plus spécifiquement via la Plateforme ISTEX, rendre les Données et les Métadonnées disponibles, fournir l'accès et permettre l'utilisation des Données et des Métadonnées, en conformité avec les termes et conditions du présent contrat ;

- concéder au CNRS et aux Bénéficiaires, le droit d'exposer les Données par l'intermédiaire d'un outil de découverte ou d'autres fournisseurs de services pour les bibliothèques dans la mesure où l'Editeur a conclu un accord avec les fournisseurs en question pour l'utilisation des Données;

- concéder aux Bénéficiaires et aux Utilisateurs autorisés le droit d'utiliser les Données hébergées sur la plateforme de l'Editeur ou sur la Plateforme ISTEEX pour se livrer à des opérations de fouilles de texte (*text-mining* et *data-mining*), conformément aux missions des Bénéficiaires.

;

(...)

Article 13. Responsabilité

(...)

13.2.2. Dispositions particulières applicables une fois mise en place la Plateforme ISTEEX

65. Une fois la Plateforme ISTEEX opérationnelle et les Données hébergées sur cette plateforme, l'ABES s'assurera que l'opérateur chargé de la mise en place de la Plateforme ISTEEX, le CNRS, se dotera d'outils techniques permettant de détecter des utilisations des Données contraires aux termes de la présente licence, notamment par l'intermédiaire des adresses IP.

(...)

Article 28. Annexes

100. La présente licence comprend les annexes suivantes :

(...)

Annexe Description de la Documentation

(...)

101. Chaque livraison doit être accompagnée d'un document précisant de façon la plus précise et la plus exhaustive possible :

- Une référence au contrat signé, des contacts (notamment un correspondant technique) chez l'éditeur et une datation des données.

- Le type de documents contenu dans la livraison (Titre de périodiques, E-book, Book Series, Dictionnaire, ...)

- Pour chaque type de document, il faudra préciser le nombre total d'éléments et de sous éléments. Par exemple pour un titre de périodique, la couverture (date de début : date de fin), le nombre de volumes, de numéros, et d'articles et de PDF associés.

- Si pour un type de document, la DTD a changé de version au cours du temps, il faudra une table associant pour chaque période, la version de DTD qui correspond aux métadonnées.

- Une description du format des données, c'est à dire l'organisation des données sur le disque (ex ISSN/Année/volume/Numero/... ou Année/ISSN/volume/Numero/...).

- Si une nomenclature spécifique est utilisée, une description de celle-ci sera demandée (ex VOL=Year.VolumeID)

102. Avant la signature du contrat, l'éditeur s'engage à fournir les éléments cités ci-dessus ainsi qu'un jeu de test comprenant au moins 1% de chaque type de données conformes aux descriptions qu'il en fera et au schéma de données associé.

(...)

104. Normes et protocoles

- L'Editeur s'engage à respecter au moins l'un des trois protocoles suivants:
 - protocole OAI;
 - protocole Z39-50;
 - protocoles SRU et SRW.
- (...)

105.Configuration

[à compléter]

106.Mises à jour de la base de données

-périodicité des mises à jour

Working Group National Licensing

In Germania, il *Working Group National Licensing*⁴⁷, gruppo di lavoro nazionale per l'elaborazione di un testo di contratto nazionale per le risorse elettroniche, ha predisposto un modello di licenza sulla base delle linee guida preparate da DFG (*Deutsche Forschungsgemeinschaft - German Research Foundation*) e la *Helmholtz Association* per gli acquisti finanziati in tutto o in parte da DFG.

Di seguito si produce un estratto dal modello di licenza, versione del 15.1.2013 del "*Model Licence Agreement Stand*" e un estratto dalle Linee guida prodotte dal gruppo di lavoro rispetto ai temi collegati all'accesso perpetuo⁴⁸.

MODEL LICENCE AGREEMENT

Stand 15.01.2013 - ESTRATTO

(...)

§ 2 Licence grant

1. The Publisher hereby grants to the Licensee, subject to and in accordance with the terms of this Licence Agreement, a non-exclusive perpetual licence to access and use the Licensed Material and to allow Institutions and Authorised Users to access and use the Licensed Material hosted on the Publisher's Platform via Secure Network.

(...)

2. The Publisher hereby grants to the Licensee, subject to and in accordance with the terms of this Licence Agreement, a non-exclusive perpetual (licence) right to host the Licensed Material on Licensee's Local Hosting Server or on the server of a third party, allow the Institutions to communicate the Licensed Material via a Secure Network to Authorised Users, to allow Authorised Users to access and use the Licensed Material via Secure Authentication.

⁴⁷Per maggiori informazioni sul gruppo di lavoro cfr. <<http://www.allianzinitiative.de/en/core-activities/national-licensing/working-group-documents.html>>.

⁴⁸Per maggiori informazioni sul gruppo di lavoro cfr. <<http://www.allianzinitiative.de/en/core-activities/national-licensing/working-group-documents.html>>. Il testo completo del modello di licenza, "Model Licence Agreement Stand 15.01.2013", è disponibile alla URL:

<http://www.allianzinitiative.de/fileadmin/user_upload/redakteur/Musterlizenzvertrag_2013.pdf>; il testo delle linee guida di DFG ("Guidelines for the Purchase of Licences funded by DFG") è disponibile alla URL <http://www.dfg.de/formulare/12_181/12_181_en.pdf> o alla URL

<http://www.allianzinitiative.de/fileadmin/user_upload/redakteur/Grndsaeetze_Allianz_Lizenzen_english.pdf>.

3. The Licensee is further permitted to make such copies or re-format the Licensed Material contained in the archival copies supplied by the Publisher in any way to ensure their future preservation and accessibility in accordance with this Licence.

(...)

4. The Publisher hereby grants to the Licensee and the Institutions, subject to and in accordance with the terms of this Licence Agreement, a non-exclusive perpetual licence for the Metadata associated with the Licensed Material for use in local library catalogues, union catalogues, and such other library and information systems including but not limited to search engines of the Institutions and third parties. The support of Search Engine Advertizing (SEA) and Search Engine Optimization (SEO) is allowed by the Publisher (e.g. xml-sitemap of websites with displays of metadata for inclusion of Search Engine Indexes; support of crawler like Googlebot) as well as the provision of the Metadata as Linked Open Data.

(...)

§ 3 Permitted uses

1. The Licensee and the Institutions may:

(...)

b. (...) The Licensed Material may be used to perform and engage in textmining/data mining activities, including but not limited to full text indexing. Notwithstanding the above-mentioned rights the use of Licensed Material shall be limited to Authorised Users.

(...)

e. Allow Authorised Users to:

(...)

- Deposit in perpetuity the learning and teaching objects as referred to in § 3.1.b on servers operated by the Institution. The use of such material shall be limited to Authorised Users

§ 5 Responsibilities of the Publisher

1. The Publisher agrees to:

a. Make the Licensed Material perpetually available to the Licensee and Authorised Users from the commencement of this Licence Agreement and after termination of this agreement at no additional cost on the Publisher's Platform by Secure Authentication as defined in Schedule 4;

(...)

d. Make the Licensed Material available on request by Licensee for Local Hosting on the Licensee's Hosting Server or on the server of a third party at no additional costs;

e. Deliver the Licensed Material to Licensee as specified below

- in case of current content provided for Local Hosting immediately after publication;
- in case of post-cancellation archival rights six months after cancellation/publication;
- in case of withdrawal of Licensed Material or any part of it before removal from Publisher's Platform;
- in case of termination of this agreement immediately after termination;
- in other cases within three months after the request has been made;

f. Deliver and make accessible the Licensed Material in a format according to the standards as defined in Schedule 5.

g. Supply the correct and complete Metadata as well as updates from time to time in machine-readable form at no additional costs. The Metadata shall be delivered as specified in Schedule 5.

h. Provide electronic product documentation to the Licensee free of charge. The Publisher will allow copies of all documentation to be made and distributed by the Licensee to the Institutions and Authorised Users provided it is either duplicated in full, or a proper ownership acknowledgement and acknowledgement of Source is included;

i. Provide the standards, services and statistics set out in Schedule 4

(...)

§ 8 Term and termination

(...)

6. After termination of this Agreement the Publisher shall provide (at the option of the Licensee) the Licensee, the Institutions and its Authorised Users with access to and use of the Licensed Material. For the avoidance of doubt access and use of archival copies shall be subject to the terms and conditions as set out in § 3 and § 4 of this Agreement. In case of termination of this Agreement by the Publisher due to a material breach access to and use of the Licensed Material shall only be on Licensee's Local Hosting server according to § 2.2.

(...)

§ 9 Acknowledgement and protection of Intellectual Property Rights

(...)

2. For the avoidance of doubt, the Publisher hereby acknowledges that any database rights created by the Licensee or the Institutions as a result of Local Hosting, text mining or data mining of the Licensed Material shall be the property of the Licensee, or the Institution.

(...)

SCHEDULE 4: Standards, Services and Statistics

(...)

8. Transfer of Titles: The Publisher shall comply with the Code of Practice of Project Transfer relating to the transfer of titles between publishers. In addition, the Publisher will use all reasonable efforts to retain a non-exclusive copy of the volumes published and make them available free of charge through the Publisher's Platform. In the event that the Publisher ceases to publish a Part or Parts of the Licensed Material, a digital archive will be maintained of such Licensed Material and will be made available free of charge through the Publisher's Platform or via a third party server and by supplying such material free of charge to the Institution.

(...)

SCHEDULE 5: Standards for Data Delivery

The Licensed Material shall be delivered to the Licensee by using open, standardized formats and accompanied by documentation

- For Metadata this may be: XML (e.g. NLM-DTD for article records) or MARC21/MARCXML/MAB (for eBooks in particular)
- For full texts this may be: PDF and/or HTML/XML

Metadata should be delivered as comprehensive as possible; this includes but is not limited to: ISBN and ISSN; DOI, author, date of publication, publisher, place of publication; if available also tables of contents, prefaces, abstracts, flap texts.

The Licensed Material shall be delivered in its entirety and shall be identical with the licensed product. The organization of the product into logical units (e.g. assignment of data records to products, of articles to journal titles) must be reflected by the data delivered.

The data delivery is accompanied by a documentation of the formats, character sets and MIME types used. Delivery shall be on a mutually agreed medium.

Guidelines for the Purchase of Licences within the Framework of the Alliance Initiative "Digital Information". June 2015 - Estratto

"Guidelines for the Purchase of Licences funded by the DF" -
(...)

A. Purchasing Regulations **(...)**

2. Licensed property

The following applies for completed databases, journal archives and completed e-book packages: They are offered for the purchase of permanent access rights through a one-time payment. The purpose of the licence is to grant to the licensees for an unlimited time the non-exclusive, non-transferable right to permit authorised institutions and authorised users to use the product via secure authentication, particularly for academic and research purposes. Free access to the vendor's server, generally with no time limit, is included in the licence fee.

For current journals and dynamic databases, the following applies: The purpose of the licence is to grant to the licensees for the contractually stipulated term the non-exclusive, non-transferable right to permit authorised institutions and authorised users to use the product via secure authentication, particularly for academic and research purposes. Free access to the vendor's server for the duration of the licensing is included in the licence fee.

In addition, the moving wall principle applies for current journals. Archives that successively build up during the term of the contract become available, after a period to be defined in each case, generally after one year, to all authorised institutions in Germany by way of a national licence, and can then be accessed within each institution's IP range for viewing and for local indexing and archiving. The licence fee includes free access to these archives on the vendor's servers for all authorised institutions, generally without a time limit.

For (continually updated) full-text databases there is also the option to develop appropriate models for making them available, after a date to be determined, to all authorised institutions in Germany by way of a national licence.

For dynamic databases, nationwide licensing is also expected to add value, for example in the form of particularly favourable pricing.

3. The object of the licence for completed databases, journal archives and e-books, as well as of licensed volumes and accrued archive volumes and segments for current journals and full-text databases, includes the right to archive the contents on servers owned by the licensees or by third parties authorised by the licensees, for the purpose of ensuring permanent availability of the contents.

(...)

7. Upon request, the licensor is obligated to physically supply the licensee with the complete product at no additional charge, i.e. including metadata and all full text, including digital objects that are part of the product, on suitable storage media and in suitable data formats as agreed. Exceptions to this rule are only permitted for dynamic databases, in which the product content continually changes without the old versions being archived. For dynamic databases, the following also applies: Should the vendor no longer offer the product, the vendor is obligated to supply the licensees with the last version as an archive version.

a) Licensees may use the data provided to them in any way they deem suitable in order to make the product accessible to authorised users, in compliance with the licence agreements. They may, for this purpose, integrate the data in technical usage/storage systems (hosting and archiving) operated by themselves or by third parties.

b) The licensees may commission third parties (e.g. library network systems and other technical infrastructure facilities of German libraries or commercial operators) with the operation of the technical equipment for secure authentication and for the use of the product by authorised users.

c) The licensees, like the authorised institutions, have the right to use the data provided to them to set up value-added services for the authorised institutions. These include, for example, full-text indexing, automated methods of information extraction and processing in licensed full texts, including digital objects and metadata (data and text mining).³ Products acquired under an Alliance licence can be integrated without restriction in electronic course packs and virtual research environments of all authorised institutions as well as into the specialised portals and virtual specialised libraries operated by authorised institutions.

d) In the case of licences or content aimed at the use of the product for an unlimited time period, the licensees, any third parties commissioned by them, and the authorised institutions also have the right to take any technical measures necessary to safeguard the product for the long term (in particular, but not limited to, the conversion of the data into other formats), or to commission third parties to perform such measures.

e) The data (e.g. metadata, content objects) are delivered in open, standardised formats (if possible XML, PDF and NLM-DTD) and accompanied by documentation.

f) Data shall be delivered in their entirety and in congruence with the licensed product.

g) The breakdown of the product into logical units (e.g. assignment of data records to products, or of articles to journal titles) must be reflected by the data delivered.

8. For metadata, the following also applies:

(...)

b) Data must be provided in standard character sets (preferably UTF-8).

c) Each metadata record must contain a unique, unalterable identification number.

d) Metadata are provided by FTP (File Transfer Protocol), the OAI-PMH method (Open Archives Initiative Protocol for Metadata Harvesting) or the B2B (Business to Business) method via an agreed interface

(...)

13 The following additional criteria apply to the purchase of e-books:

(...)

f) The products are offered in a widely available format and with widely available reader software. Only in exceptional cases may licensing require the purchase of proprietary reader software. In this case, it is mandatory that the vendor provide standard formats for hosting and long-term archiving at no additional charge.

g) Metadata shall be delivered in the MARC 21 or the ONIX for Books format.

4) CRKN/RCDR

Le clausole sono estratte dal modello canadese redatto dal *Canadian Research Knowledge Network (CRKN)*⁴⁹.

(...)

2. GRANT OF LICENSE

(...)

2.2 The Publisher hereby grants to the Consortium and the Members a non-exclusive, royalty-free, perpetual license for Members and their Authorized Users to use, after the termination of this Agreement, the Licensed Materials that were accessible during the term of this Agreement. Such use shall be subject to clauses 3, 4, 5, 6, 12.4, 12.4.1, and 12.4.2 of this Agreement, which provisions shall survive any termination of this Agreement. The means by which the Members shall have access to such Licensed Materials shall be in a manner and form substantially equivalent to the means by which access is provided under this Agreement.

2.3 Notwithstanding any terms or conditions to the contrary in any agreement between authors affiliated with a Member and the Publisher, the Publisher hereby grants to the Members the non-exclusive, irrevocable, royalty-free right to archive or deposit the Content online in institutional, subject-based, national, or other open repositories or archives.

(...)

3. USAGE RIGHTS

3.1 The Consortium and the Members, subject to clause 6 below, may:

⁴⁹ O Réseau Canadien de Documentation pour la Recherche (RCDR). Il testo completo della licenza è reperibile alla URL http://crkn.dev.marsworks.com/sites/crkn.dev.marsworks.com/files/site/crkn-model_license_final_2014-04-07.pdf.

3.1.1 Make such local electronic copies of part of the Licensed Materials by means of caching as may be necessary solely to ensure efficient use of such materials by Authorized Users and Walk-in Users.

(...)

3.1.4 Allow Authorized Users to have access to the Licensed Materials through the Member's Electronic Learning Environment.

3.1.6 Apply automated tools and processes to the Licensed Materials, for the purposes of data mining for purposes of textual analysis and visual mapping of textual relationships, within the context of scholarship or research activities.

(...)

6 PROHIBITED USES

6.1 Neither the Consortium nor Authorized Users nor Walk-in Users may:

6.1.3 mount or distribute any part of the Licensed Materials on any electronic network, including without limitation the Internet and the World Wide Web without the prior written consent of the Publisher or except pursuant to Clause 12 of this Agreement or any separate agreement as the Publisher and a Member may enter into.

8 PUBLISHER'S UNDERTAKINGS

8.1 The Publisher shall:

8.1.1 make the Licensed Materials available to the Consortium and the Members from the Server in the media, format and time schedule specified in Schedule 3. The Publisher will notify the Consortium at least sixty (60) days in advance of any anticipated specification change applicable to the Licensed Materials. If the change renders the Licensed Materials less useful in a material respect to the Consortium and the Members, the Consortium may treat such changes as a material or persistent breach of this Agreement under clause 12.1.2.

12 TERM AND TERMINATION

12.3 On termination all rights and obligations of the parties automatically terminate except for obligations in respect of Licensed Materials for which perpetual access is provided for in clause 12.4.

12.4 Upon termination of this Agreement a Member may provide a notice to the Publisher (the "Access Notice") of its desire to exercise its rights for perpetual access to and use of the Licensed Materials that were made available to the Member pursuant to this Agreement and each previous Agreement between the Publisher and the Consortium concerning all or part of the Licensed Materials (collectively the "Historical Licensed Materials"). Upon receipt of the Access Notice the Publisher shall provide to the Member and its Authorized Users and Walk-in Users access to the Historical Licensed Materials in the same manner in which the Publisher provided access to the Licensed Materials immediately prior to the date of termination of this Agreement all without change. The Member and its Authorized Users and Walk-in Users shall have the same rights with respect to the Historical Licensed Materials as they had immediately prior to the date of termination of this Agreement.

In addition to the above, the Publisher and the Member shall negotiate in good faith for other methods of perpetual access to the Historical Licensed Materials. Options include the following:

- a) continuing online access to archival copies of the Historical Licensed Materials on the Server;
- b) supplying archival copies of the Historical Licensed Materials to the Member in an electronic medium mutually agreed between the parties;
- c) supplying archival copies of the Historical Licensed Materials to an appropriate hosting and archiving facility (e.g. OCUL Scholars Portal) for continuing online access to the Historical Licensed Materials;
- d) granting access to the Historical Licensed Materials through an archival service provider such as, but not limited to, Portico as appropriate;
- e) other options as mutually agreed between the Publisher and the Member.

For the avoidance of doubt, access and use of the Historical Licensed Materials shall be subject to the terms and conditions as set out in clauses 3, 4, 5, 6, 12.4, 12.4.1, and 12.4.2 of this Agreement, notwithstanding the termination of this Agreement. Members are permitted to make copies or re-format the Historical Licensed Materials to ensure their future preservation and accessibility.

12.4.1 In the event that ownership of part of the Historical Licensed Materials is sold by the Publisher or is otherwise transferred to another publisher, the Publisher shall use reasonable efforts to retain a copy of the Historical Licensed Materials in accordance with the UKSG Transfer Code of Practice and make it available to a Member and its Authorized Users and Walk-in Users as required by clause 12.4.

12.4.2 The Historical Licensed Materials made available in accordance with clause 12.4 shall be the complete content except that such Historical Licensed Materials may not contain all the links and other platform features and functionality associated with the Licensed Materials licensed pursuant to this Agreement.

13 GENERAL

13.4 If rights in all or any part of the Licensed Materials are assigned to another publisher, the Publisher shall use its best efforts to ensure that the terms and conditions of this Agreement are maintained.

SCHEDULE 3

LICENSED MATERIALS AND ACCESS METHOD

A schedule dated [DATE] to the Agreement [DATE] between [PUBLISHER NAME] and CANADIAN RESEARCH KNOWLEDGE NETWORK.

THE LICENSED MATERIALS

(...)

ACCESS METHOD

(...)

Local Loading and Local/Alternate Hosting: Publisher agrees to negotiate, at any point in the term of this Agreement, a separate Agreement in good faith with the Member or a duly authorized representative (e.g. a library consortium acting on behalf of the Member) for local hosting on the Member's server or a third party server, as designated by the Member.

PERPETUAL ACCESS:

AS WITNESS the hands of the parties the day and year below first written

FOR THE PUBLISHER: [NAME

(...)

FOR THE CONSORTIUM: CANADIAN RESEARCH KNOWLEDGE NETWORK

(...)

APPENDICE B – Lettera e questionario inviati agli archivi al fine dell’Indagine presso le iniziative di National Hosting

LETTERA

Dear...

I write as coordinator of Group 404 and would like to ask you if you could spend a few minutes regarding your Archive initiative ...considered as one of the most interesting to the purpose of our working group.

Group 404 presentation

[Group 404](#) aims at laying down a national strategy on perpetual access and long-term preservation for scholarly publications in digital, full text edition (journals, books and other scientific publications in full text).

This Group was set up on July 3rd, 2014 on the initiative of the Italian community of the libraries members of [CLOCKSS](#) archive. The [Group](#) counts on specialists from academic and research libraries, the Italian Interuniversity Consortium [CINECA](#) of super computing and the [National Central Library of Florence](#). The Group maintains a strong connection with [CARE-CRUI](#) which is the Italian organization responsible for the national negotiations for electronic resources.

Group 404 accomplishes its objectives through the publication of two reports: the first one already published on September 30th, is available on the web site ([only in Italian](#)); the second will be published on the basis of the [needs analysis](#) by June 30th, 2015 and will mark the end of the Group's activities.

Survey of National Hosting initiatives

In order to publish this final report, the Group believes it essential to carry out an indepth analysis through a survey among those archives offering National Hosting services as well ([ISTEX-Panist](#) for France; [Scholars Portal](#) for Canada; [EJC](#) and [EBC](#) for the consortium OhioLINK USA), among the most recent and innovative pilot projects for National Hosting (the pilot project [SafeNet](#) in the UK with the application of [LOCKSS](#) technology for the realization of the National Hosting) as well as among those initiatives that may be competing with the National Hosting (archive [Portico](#) in USA and archive [e-Depot](#) in the Netherlands) though they are not National Hosting initiatives.

How to fill in the questionnaire

The questionnaire is available at the link:

https://docs.google.com/forms/d/1Ii8qeYn4Jc8fsphJeAhDWXFhrUgTcTjmsa_KbJz5XNg/viewform

If you find you are not able to answer any question, please just write “Answer not possible”.

If you prefer not to answer any question, please write “Confidential matter”.

If any question is not clear, or if you need explanations and/or more details, you could find my contacts below.

Thank you for your collaboration.

Best regards,

Gianluca D’Amato, Group 404 - coordinator

QUESTIONARIO

TECHNICAL QUESTIONS

1. Please insert your full name, organisation and role.
2. Is the software used for managing the Archive proprietary or open source?
3. If it is a proprietary software, what is its name and who is the manufacturer? What is the licensing type (per user, per sites, etc)?
4. If it is an open source software, what is its name? Is the open source community active (eg number of active developers / number of updates released per year etc)? Are source files available?
5. Comments/additions

ARCHIVED DOCUMENTS

6. Which kind of documents are preserved in the Archive (ebooks, ejournals, others)? If others, please specify (i.e. Institutional Repositories, administrative documents)?
7. Does the Archive preserve only Toll Access or Open Access documents as well?

STORAGE

8. What is the actual size of the Archive (TB)? Is it possible to divide the storage by type of document? I.e. percentages for each document type (ejournals, ebooks, etc)?
9. What is the average growth of the Archive storage per year? If possible, indicate the percentage growth of each type of document (ebooks, eJournals, etc.).

STAFF

10. How many FTE people are employed to manage the Archive per year and what skills are required (how many system administrators, developers, librarians, etc)?
11. How many employees work exclusively for the Archive? How many are external contractors? How many secondees are there?
12. What is the organization chart of the staff?

GOVERNANCE QUESTIONS

13. What is the legal status of the Archive (Association, Foundation, Division within a Consortium; division within a university, etc.)?
14. What is (or what are) the governing bodies of the Archive?
15. Who appoints the governing bodies?
16. How long are the members of the governing bodies appointed for?
17. Is the Archive statute available?

QUESTIONS ABOUT FINANCING AND COSTS SHARING

18. How is the Archive financed?
19. Who pays? How are the fees decided?
20. What is the annual cost of the Archive?
21. Are the Archive activities economically sustainable?
22. Are any changes in financing the Archive foreseen in the future? If so in what direction and why?

QUESTIONS ON SERVICES AND USERS

23. What other services does the Archive provide in addition to perpetual access to subscribed content?
24. Are TDM (Text Data Mining) services provided for the archived content?
25. Who are the users allowed to access archived content?

- 26. How is authentication handled. With Shibboleth, user and password, other?
- 27. Does the Archive have to respect any SLA with members institutions? If so, what are they?

QUESTIONS ABOUT DOCUMENTATION

- 28. Is the documentation (i.e. Operation plan; Membership agreement; Submission agreement; etc.) about the Archive public? If so, where is it available?
- 29. If it is not public, may the Group 404 have a copy, providing you with the proper confidentiality guarantee?

MISCELLANEOUS

- 30. How long has the Archive been operating?
- 31. Could the Archive be interested in cooperating with other National Hosting projects within an international network?
- 32. How many institutions use the services of the archive?
- 33. Other

AUTHORIZATIONS

- 34. The 404 Group is authorized to publish the answer to the present survey into the report that the 404 Group are preparing

Yes/No

- 35. The 404 Group is authorized to publish the answer to the present survey only per summary into the report that the 404 Group are preparing

Yes/No

March 2015.

Gruppo 404 (ITALY)

Creative Commons CC BY-NC

Appendice C - RACCOMANDAZIONI

Raccomandazione n° 1. Per realizzare la visione strategica è necessario ottenere in misura sempre più efficace il coinvolgimento di coloro che a diverso titolo partecipano al governo del sistema accademico e di ricerca italiano e in particolare i Ministeri competenti (MIUR e MIBACT), la CRUI e i singoli Atenei (soprattutto i più grandi), i principali Centri di ricerca nazionali.

Raccomandazione n° 2. È essenziale che, per le risorse Toll Access per cui sono stati negoziati diritti di accesso perpetuo, la visione strategica si sostanzi in clausole contrattuali; in particolare esse dovranno essere incluse nel modello di licenza nazionale predisposto dal gruppo CARE presso la CRUI, che oggi in Italia segue le trattative nazionali.

Precisazione. In allegato 1 sono riportate le clausole da prevedere nei contratti nazionali al fine di tradurre, in modo sufficientemente adeguato, la visione strategica negli accordi con gli editori di pubblicazioni accademiche.

Raccomandazione n° 3. È necessario che anche le pubblicazioni Open Access adottino soluzioni adeguate in tema di conservazione a lungo termine (DP), anche a tutela degli autori.

Precisazione. In allegato 2 si riportano più nel dettaglio gli elementi che discendono dalla presente raccomandazione.

Raccomandazione n° 4. Con l'obiettivo di rendere meno costose e più sostenibili le attività che editori e biblioteche devono attuare ai fini della conservazione a lungo termine (DP), dell'accesso perpetuo (PCA) e del deposito legale digitale (DLD), si raccomanda di predisporre accordi e meccanismi automatici che consentano di alimentare il deposito legale digitale, sia per le pubblicazioni Open Access sia per le pubblicazioni Toll Access, direttamente da quegli archivi nei quali gli editori italiani riversano i contenuti ai fini della DP e del PCA (es. CLOCKSS, Portico, ecc.).

Raccomandazione n° 5. Verificare e concordare la possibile evoluzione tecnologica dell'Archivio NERA con la quantificazione dei relativi costi per valutarne la capacità di rispondere adeguatamente alle effettive esigenze qualitative del servizio, l'integrazione con altri servizi gestiti dagli Atenei e la sostenibilità economica.

Precisazione. Dal punto di vista tecnologico le disponibilità di Scholars Portal, Portico, ISTEEX, Global LOCKSS Network, acquisite durante i lavori del Gruppo 404, possono essere la base da cui ripartire per una reingegnerizzazione dell'archivio. La collaborazione con Scholars Portal potrebbe essere particolarmente significativa in quanto potrebbe aiutare non solo per le scelte tecnologiche e per l'organizzazione della produzione, ma anche per il recupero dei contenuti, in quanto Scholars Portal ha avuto una genesi simile a quella di NERA, ossia è una migrazione di ScienceServer.

Tempi. Produrre una proposta entro 15 mesi dall'inizio della verifica.

Raccomandazione n° 6. Intraprendere l'esperienza di una PLN LOCKSS con sei università italiane che diano sufficienti disponibilità e garanzie. Questo è molto auspicabile che avvenga in coordinamento con CRUI in quanto riferimento per le trattative nazionali.

Precisazione. Ciascuna delle sei università, da individuare preferibilmente tra gli atenei italiani che già aderiscono all'archivio CLOCKSS che usa la medesima tecnologia e che sono già in contatto con lo staff del programma LOCKSS dell'Università di Stanford, metterà a disposizione:

- a) l'hardware da inserire nella propria rete;
- b) il personale (un bibliotecario, un sistemista, uno sviluppatore, tutti part time).

Se interessata, CRUI metterà a disposizione un budget da cui attingere per le necessità legate all'iniziativa.

Tempi. Realizzare entro dodici mesi dall'inizio delle attività l'accesso perpetuo alternativo rispetto all'accesso su server dell'editore per almeno un editore estero importante che non riversa i dati in NERA e li riversa in GLN (ad esempio uno tra Taylor and Francis, Sage, OUP e simili) ovvero per un editore come

Springer (che riversa in NERA e aderisce anche a GLN) e per almeno un editore italiano (ad esempio uno tra Il Mulino, Franco Angeli, Vita e Pensiero, ecc.). Nel caso dell'editore italiano quest'azione, sempre entro i dodici mesi, deve sperimentare il popolamento in automatico del deposito legale digitale di Magazzini Digitali, purché la BNCF garantisca il necessario supporto. Nel caso degli editori stranieri si negozi con GLN e CLOCKSS un accordo per lo scambio dei plugin in modo da non dover replicare il lavoro già fatto da altre reti LOCKSS, creando economie di scopo che deriverebbero da un ecosistema a livello internazionale che usa la medesima tecnologia LOCKSS.

Raccomandazione n° 7. Mantenere le relazioni con gli altri archivi nazionali e internazionali in modo da realizzare uno scambio di esperienze e di best practice.

Precisazione. Particolarmente interessante potrebbe risultare la collaborazione con i britannici di SafeNet (ad esempio per l'Entitlement Registry) e con i francesi di ISTEEX. Tendere a una collaborazione con i partner europei è auspicabile anche per configurare infrastrutture interoperabili in grado di dare servizi ai cittadini europei con conseguente accesso a fondi dell'Unione Europea.

Raccomandazione n° 8. Ai fini del finanziamento, cercare fondi anche oltre le risorse messe a disposizione eventualmente da CRUI-CARE, cogliendo eventuali opportunità offerte da bandi (pubblici o privati) nazionali e /o europei, e/o internazionali.

Precisazione. Con la migrazione dai formati analogici ai formati elettronici distribuiti via web, ai fini di una migrazione sostenibile e professionalmente adeguata, è importante sottolineare come al finanziamento degli archivi digitali devono essere indirizzate, almeno in parte, le spese previste nel mondo analogico per acquisto-affitto di magazzini; acquisto della scaffalatura; rilegatura; spolveratura, ecc. La spesa per la conservazione dei documenti nel passaggio dall'analogico al digitale deve essere riqualificata ma non deve essere soppressa se non al prezzo di scelte errate sotto il profilo professionale e gestionale.

Raccomandazione n° 9. Continuare (CRUI-CARE) ad offrire la possibilità di un contratto nazionale alle migliori condizioni per le università e gli enti di ricerca italiani per l'Archivio Portico, in quanto garanzia ulteriore e non sostitutiva del National Hosting ai fini della tutela dei diritti di accesso perpetuo da offrire alle Università italiane interessate. Negoziare (CRUI-CARE) un accordo per una partnership strategica anche con l'Archivio CLOCKSS, sia per rafforzare la presenza della comunità italiana (biblioteche ed editori) in questo importante archivio, sia per esperire possibili integrazioni con altri archivi digitali italiani, ad iniziare dal deposito legale digitale della Biblioteca Nazionale Centrale di Firenze (BNCF).

GLOSSARIO

Molti termini del presente glossario riprendono, riadattandole, definizioni riportate nei documenti di cui alla sezione “Bibliografia e Standard”, dando precedenza, laddove possibile, alle definizioni riportate negli standard internazionali. I termini necessari per discutere di DP e PCA sono molti: in questa sezione sono presenti solo quelli più tecnici, riportati nel rapporto, al fine di chiarirne l’accezione con cui sono stati usati. I termini seguono una successione logica, per cui non sono ordinati alfabeticamente.

Conservazione a lungo termine (Digital Preservation o DP o Long term preservation). Gli atti (i processi e le procedure) necessari per mantenere le informazioni in un formato comprensibile nel lungo termine, ossia un periodo di tempo tanto lungo quanto è necessario perché ci si trovi nelle condizioni di dover superare problemi derivanti dai limiti di storage, dai fallimenti dei media o dai cambiamenti tecnologici. La DP non è ad esclusivo vantaggio delle biblioteche che sottoscrivono contenuti a pagamento quanto un’attività a vantaggio della società nel suo insieme, poiché tende ad assicurare che il materiale digitale continui ad essere accessibile alle future generazioni.

Accesso perpetuo (o accesso permanente o Post cancellation access o PCA o Post termination access o PTA o perpetual access). Il diritto di un sottoscrittore e dei suoi utenti ad avere l’accesso permanente ai materiali elettronici per i quali si è pagato. Il PCA entra in gioco quando una biblioteca ha necessità di tutelare se stessa e i propri utenti per mantenere l’accesso ai contenuti pagati e per i quali ha negoziato un diritto di accesso perpetuo per similitudine con il modello di acquisti seguito nel mondo analogico, dove non era necessario ripetere l’acquisto di un libro o di un’annata di periodico (se non in limitati casi residuali).

Deposito Legale Digitale (o Digital Legal Deposit o DLD). Il Deposito legale è una importante invenzione legislativa che accompagna l’età moderna dal XVI secolo e consiste nell’obbligo (in Italia in capo agli stampatori) di consegnare una copia (in Italia più copie) del materiale pubblicato nella nazione di riferimento (nel nostro caso l’Italia) alla Biblioteca Nazionale. Il Deposito legale ha sostanzialmente due finalità: conservare l’eredità culturale di un paese; consentire a tutti i cittadini di avere accesso ai documenti quale base fondamentale per esercitare i diritti di cittadinanza. Con la rivoluzione dell’editoria elettronica distribuita via web questo importante istituto richiede interventi legislativi e innovazioni organizzative e tecniche. Solitamente i materiali pubblicati in edizione elettronica devono essere consegnati alla Biblioteca Nazionale e l’accesso deve essere consentito ai cittadini usando apposite macchine ubicate in edifici dedicati della Biblioteca Nazionale. In Italia il deposito legale digitale è disciplinato dal D.P.R. 252/2006 e ha nella Biblioteca Nazionale Centrale di Firenze l’attore impegnato in prima linea. Ogni altra informazione è disponibile alla URL: <http://www.depositolegale.it/>.

Long term. Un periodo di tempo lungo abbastanza per cui si incorre nella necessità di far fronte all’impatto di cambiamenti nelle tecnologie (incluso il supporto per nuovi media e formato di dati), oppure al cambio della comunità degli utenti per fare in modo che le informazioni in archivio digitale continuino ad essere accessibili.

Trigger event. Questo termine è usato di norma per determinare i casi al verificarsi dei quali si rende necessario attivare l’accesso sull’archivio digitale. I trigger event sono riconducibili a casi dipendenti a) dall’editore o b) dal sottoscrittore.

a) Nel primo caso, sulla base dell’esperienza maturata fino ad oggi, si possono verificare quattro tipologie di casi: 1) un editore ha deciso di non continuare più la sua attività economica; 2) il singolo particolare contenuto non è più reso disponibile; 3) i back files di un contenuto corrente (anche ad esempio i contenuti

pubblicati sotto un altro titolo o dentro una diversa collection) non sono più resi disponibili per ragioni commerciali o tecniche (si pensi alle risorse che mettono a disposizione le ultime 10 annate di un periodico: ogni anno se ne perderebbe una); 4) una continuata e permanente impossibilità e incapacità dell'editore di ripristinare l'accesso alla piattaforma editoriale.

b) I casi dipendenti dal sottoscrittore sono due: 1) quando per ragioni economiche o di altro tipo il sottoscrittore decide di non rinnovare i suoi abbonamenti; 2) oppure quando le sottoscrizioni sono state rinnovate, ma in generale il sottoscrittore ritiene importante – nell'ipotesi che possa sempre accadere un evento nuovo causa di un disservizio agli utenti – mantenere un accesso alternativo e concorrente ai contenuti per i quali ha acquisito un diritto di accesso perpetuo, offrendo ai propri utenti l'accesso alternativo sui server dell'archivio digitale. I trigger event devono essere identificati e disciplinati in modo chiaro e rigoroso nei testi contrattuali con i titolari dei diritti, per chiarire quando l'accesso all'Archivio Digitale è consentito.

Archivio Digitale (Digital Archive o Digital Repository). Un'organizzazione che intende conservare le informazioni digitali per l'accesso e l'uso da parte di una Comunità Designata. Ai fini del tipo di archivi trattati nel presente report si possono distinguere almeno tre differenti tipi di Archivio Digitale in base ai trigger event che determinano l'attivazione dell'accesso sui server dell'Archivio digitale.

Dark Archive. Un Archivio digitale che non è accessibile al pubblico se non quando i contenuti conservati nell'archivio sono soggetti ad uno dei Trigger Event dipendenti dall'editore. I contenuti soggetti a Trigger Event vengono allora resi accessibili alla Comunità Designata secondo determinate regole. Un Dark Archive serve per la conservazione a lungo termine. CLOCKSS, Rede Cariniana e Portico sono un esempio di Dark Archive.

Dim Archive. Un Archivio digitale che dà l'accesso ai contenuti sottoscritti al verificarsi dei Trigger Event dipendenti dall'editore o dal sottoscrittore, limitatamente al caso b1). Un Dim Archive serve sia per la DP ma soprattutto per il PCA, e ha la caratteristica di fornire un accesso ai contenuti sottoscritti nel momento in cui non si rinnova l'abbonamento e non si vuole più pagare l'editore per avere l'accesso sui suoi server. Ne sono un esempio Global LOCKSS Network e Portico (limitatamente a quegli editori che hanno firmato con Portico un accordo che preveda anche il PCA).

Light Archive. Il Dim Archive diventa Light Archive quando si attiva, oltre che nei casi del Dim Archive, anche nel caso b2) ossia quando l'archivio offre un accesso alternativo concorrente sempre funzionante, ossia quando il Dim Archive offre sempre agli utenti autorizzati un accesso alternativo e concorrente rispetto all'accesso sul sito degli editori, anche in presenza di un abbonamento corrente e con una piattaforma editoriale perfettamente funzionante. Questo è il caso di Scholars Portal; era il caso di SDOS ed Emeroteca Virtuale.

Comunità Designata (o Comunità di riferimento o Designated Community). Un identificato gruppo di utenti potenziali che dovrebbe essere in grado di comprendere un determinato set di informazioni e che è servito dall'Archivio digitale. La Comunità Designata può essere composta da una pluralità di comunità utenti che si uniscono a formare la Comunità Designata servita dall'Archivio in generale.

Archivio OAIS (OAIS Archive). Acronimo che sta per Open Archival Information System. Si tratta di uno Standard ISO nella sua seconda edizione (ISO 14721:2012; prima edizione del 2003) sviluppato dal Consultative Committee for Space data Systems per l'acquisizione, la conservazione, la diffusione dei contenuti digitali. Un Archivio OAIS è un'organizzazione di persone e sistemi che ha accettato la responsabilità di conservare le informazioni e di renderle disponibili per una Comunità Designata seguendo il modello logico-funzionale delineato dallo standard. In particolare un Archivio OAIS deve far fronte alle responsabilità definite nella sezione 3.1 dello standard ISO 14721:2012.

Archivio digitale affidabile (Trustworthy Digital Repository). Un archivio digitale affidabile è un archivio che ha conseguito la certificazione di affidabilità rilasciata da un ente abilitato (standard ISO sugli enti certificatori, ISO 16919:2014) utilizzando lo standard ISO 16363:2012. Questo standard si basa sul documento TRAC pubblicato da RLG e NARA nel 2003, oggi mantenuto da una task force internazionale.

L'ultima edizione di TRAC ha il copyright di CRL e OCLC. CRL si sta accreditando (soprattutto in USA) come ente certificatore di fatto usando TRAC.

Accordo per la consegna dei dati e dei metadati (Submission Agreement). L'accordo tra l'Archivio digitale e l'editore che disciplina le modalità di consegna dei metadati e dei contenuti. L'accordo può riflettere diverse modalità.

Archivio con approccio harvesting. Si tratta di un Archivio che conserva i contenuti così come sono pubblicati sulla piattaforma dell'editore ("look and feel"). La tecnologia LOCKSS adotta questo approccio utilizzando due modalità: facendo l'harvesting dei metadati e dei contenuti direttamente dal sito dell'editore; oppure facendosi consegnare dall'editore i source files del suo sito web mediante FTP.

Archivio con un approccio di normalizzazione o di standardizzazione. L'archivio normalizza i datafeeds che riceve dall'editore e li ripubblica nell'archivio medesimo. È l'approccio seguito ad esempio da Portico.

Migrazione. Il processo di conversione di un file da un formato diventato (o a rischio di diventare) obsoleto ad un altro formato in uso corrente (non obsoleto). La migrazione può essere realizzata in diversi modi: Portico e LOCKSS ad esempio effettuano le migrazioni di formato in modi differenti.

Emulazione. Una soluzione che ricostruisce l'ambiente software (e talvolta anche hardware) che rende possibile la rappresentazione delle informazioni che si trovano in un formato obsoleto.

Conservazione digitale distribuita (Distributed Digital Preservation DDP). Il processo di creare copie di file digitali e di conservazione degli stessi in ubicazioni distribuite geograficamente per i propositi della conservazione a lungo termine. Le best practice in DDP sono:

- o I luoghi di conservazione (siti) non devono trovarsi in un raggio di azione compreso tra 120-200 km l'uno dall'altro;
- o I siti devono essere distribuiti in aree geografiche al di fuori dei luoghi interessati storicamente da disastri naturali (alluvioni, terremoti, uragani, tornado, ecc.);
- o I siti devono essere distribuiti in luoghi serviti da differenti reti elettriche;
- o I siti devono essere sotto il controllo di differenti system administrator;
- o I contenuti conservati nei siti distribuiti devono essere in linea e devono essere controllati regolarmente per bit-rot per verificarne lo stato di decadenza;
- o I siti di conservazione devono essere come minimo tre, scelti secondo le specifiche di cui ai punti precedenti.

Backup. Il processo di backup serve per realizzare copie di dati che possono essere usati per il ripristino dei file originali persi a causa del verificarsi di un evento. Il backup dei contenuti dei server è una procedura standard usata da molto tempo nelle pratiche di amministrazione dei sistemi. Il backup semplice tuttavia non è sufficiente per accreditare un Archivio Digitale come OAIS compliant e certificato secondo lo standard ISO 16363:2012. Ad esempio è sempre necessario il coinvolgimento di un altro ente a cui affidare copie dell'archivio (siano esse online che offline).

Local Hosting (o LH o Local Loading o LL). Il Local Loading è la copia in locale dei metadati e dei contenuti per i quali è stato acquisito un diritto di accesso perpetuo. Il local loading può essere fatto da una singola istituzione (ad esempio quanto fa la Technical University of Denmark, cfr. Charles Beagrie, *Consultancy report...*, p. 12) o da più istituzioni associate (ad esempio i servizi SDOS ed Emeroteca Virtuale).

National Hosting (o NH). Il National Hosting è una forma di Local Loading ma con una rilevanza nazionale, ossia un archivio digitale ai cui servizi, soprattutto di PCA, possono rivolgersi le organizzazioni di una determinata nazione. Il progetto francese ISTEK è una forma di National Hosting; il progetto britannico

SafeNet lavora in una prospettiva di National Hosting; Scholars Portal è una forma di Local Hosting che evolverà in National Hosting una volta raggiunto un accordo tra OCUL e CRKN.

Text Mining e Data Mining (o TDM). Il TDM è un processo di ricerca che, attraverso l'uso di software dedicati e di tecniche di analisi e di estrazione di contenuti e di dati, è in grado di ricombinare le informazioni permettendo di identificare nuovi modelli, ipotesi, tendenze, creando nuova conoscenza. Se per esempio ci si riferisce ai periodici e ai libri elettronici, è importante che la ricerca venga effettuata su una gran quantità di documenti a testo completo, pubblicati da diversi editori, incrociando diversi campi. Queste attività stanno assumendo un'importanza crescente per i ricercatori. In base a questa definizione si comprende come gli Archivi Digitali per la DP, per il PCA, per il DLD, curando la conservazione a testo completo di diverse migliaia di documenti editi da diversi editori, si prestano a fungere come base documentale su cui applicare tool di TDM. La posizione dei diversi archivi oggetto dell'indagine è differente e per i dettagli si rimanda al contenuto del rapporto.

BIBLIOGRAFIA E STANDARD

Standard internazionali

ISO 14721:2012 *Space data and information transfer systems — Open archival information system (OAIS) — Reference model*, 2. Ed., Geneva : The International Organization for Standardization), 2012, URL: http://www.iso.org/iso/catalogue_detail.htm?csnumber=57284.

Una copia dello standard nella versione non ufficiale è disponibile sul sito della CCSDS alla URL: <http://public.ccsds.org/publications/archive/650x0m2.pdf>

ISO 16363:2012 *Space data and information transfer systems — Audit and certification of trustworthy digital repositories*, 1. Ed. Geneva : The International Organization for Standardization), 2012, URL: http://www.iso.org/iso/catalogue_detail.htm?csnumber=56510

Una copia dello standard nella versione non ufficiale è disponibile sul sito della CCSDS alla URL: <http://public.ccsds.org/publications/archive/652x0m1.pdf>

ISO 28500:2009 *Information and documentation — WARC file format*, 1. Ed. Geneva : The International Organization for Standardization), 2009, URL: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=44717

Una copia dello standard nella versione non ufficiale è disponibile sul sito della BNF alla URL: http://bibnum.bnf.fr/warc/WARC_ISO_28500_version1_latestdraft.pdf

ISO 16919:2014 *Space data and information transfer systems – Requirements for bodies providing audit and certification of candidate trustworthy digital repositories*, URL: http://www.iso.org/iso/catalogue_detail.htm?csnumber=57950.

Buone pratiche, pratiche raccomandate, alcuni standard nazionali

A Guide to Distributed Digital Preservation, edited by Katherine Skinner and Matt Schultz, Atlanta, Educopia Institute, 2010, URL: <http://www.metaarchive.org/GDDP>

Trustworthy Repositories Audit & Certification (TRAC): Criteria and Checklist, Chicago : CRL, The Center for Research Libraries and Dublin : OCLC Online Computer Library Center, 2007 URL: http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf

Digital Repository Audit Method Based On Risk Assessment (DRAMBORA), by Digital Curation Centre (DCC) & DigitalPreservationEurope, URL: <http://www.repositoryaudit.eu/>

PREMIS, *Preservation Metadata: Implementation Strategies*, vers. 2.2, Washington : Library of Congress, 2012, URL: <http://www.loc.gov/standards/premis/>

Altri standard sui metadati a cura dalla Library of Congress (METS, MODS, ecc.) URL: <http://www.loc.gov/standards/>

ANSI/NISO Z39.96-2012 *JATS, Journal Article Tag Suite*, vers. 1.0, Baltimora : National Information Standard Organization, 2012, URL: <http://www.niso.org/standards/z39-96-2012/>

NISO-RP-9-2014, *Knowledge Base And Related Tools (KBART)*, Baltimora : National Information Standard Organization, 2014, URL: http://www.niso.org/apps/group_public/download.php/12720/rp-9-2014_KBART.pdf

OAI-PMH, *The Open Archives Initiative Protocol for Metadata Harvesting*, Open Archives Initiative, URL: <http://www.openarchives.org/OAI/openarchivesprotocol.html>

ANSI/NISO Z39.88-2004 (R2010), *The OpenURL Framework for Context-Sensitive Services*, URL: http://www.niso.org/apps/group_public/project/details.php?project_id=82

Bibliografia

La bibliografia sull'argomento diventa ogni anno più lunga e articolata. In questa sezione sono riportati solo i testi a cui viene fatto riferimento nel rapporto o usati come fonte per le definizioni riportate nel glossario.

André Dazy, *ISTEX: a powerful project for scientific and technical electronic resources archives*, in *Insights*, November 2014, pp. 269-273

Blue Ribbon Task Force on Sustainable Digital Preservation and Access, *Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information*, Final Report of the, Blue Ribbon Task Force on Sustainable Digital Preservation and Access, February 2010
http://blueribbontaskforce.sdsc.edu/biblio/BRTF_Final_Report.pdf

Center for Research Libraries, *Certification Report on CLOCKSS*, luglio 2014
http://www.crl.edu/sites/default/files/attachments/misc/CLOCKSS_Report_2014.pdf

Center for Research Libraries. *Report on Scholars Portal Audit*, febbraio 2013
http://www.crl.edu/sites/default/files/attachments/pages/ScholarsPortal_Report_2013_%C6%92.pdf

Center for Research Libraries, *Report on Portico Audit Findings*, 10 gennaio 2010
<http://www.crl.edu/sites/default/files/attachments/pages/CRL%20Report%20on%20Portico%20Audit%20010.pdf>

Charles Beagrie, *Ensuring Perpetual Access: Establishing a Federated Strategy of Perpetual Access and Hosting of Electronic Resources for Germany*, prepared by Charles Beagrie Limited in association with Globale Informationstechnik GmbH, a study funded by The Alliance of German Science Organizations, Bonn : Deutsche Forschungsgemeinschaft, 2010 https://www.mpg.de/230683/access_hosting_studie_e.pdf

International Publishers Association, *Digital Legal Deposit : An IPA Special Report*, 24 luglio 2014
<http://www.internationalpublishers.org/images/news/2014/digital-legal-deposit-2014.pdf>

Maryanne Wolf, *Proust and the Squid. The Story and Science of the Reading Brain*, New York, HarperCollins, 2007, trad. it. *Proust e il calamaro: storia e scienza del cervello che legge*, Milano, Vita e Pensiero, 2009

Neil Beagrie, *Permanent Access to e-Journals in Denmark, Final report (Version 3.0)*, 8 ottobre 2014

Stewardship of our Scholarly resources. The report of the Perpetual Access Task Group (PATG) to CRKN, 15 febbraio 2012 http://crkn.ca/sites/crkn.ca/files/file/PATG%20report_OCUL%20comments.pdf