

UNIVERSITA' DEGLI STUDI DI PAVIA  
FACOLTA' DI LETTERE E FILOSOFIA

CORSO DI LAUREA SPECIALISTICA IN SCIENZE ARCHIVISTICHE,  
DOCUMENTARIE E BIBLIOTECONOMICHE

LA CONSERVAZIONE DELLA MEMORIA DIGITALE  
QUESTIONI METODOLOGICHE E GIURIDICHE

RELATORE

Chiar.mo Prof. Paul Gabriele Weston

CORRELATORE

Chiar.ma Prof.ssa Gianna Del Bono

TESI DI LAUREA  
di  
ELETTRA DE LORENZO

Anno Accademico 2005-2006

## Indice

<b>1. Conservare la memoria digitale: questioni metodologiche</b>	<b>p. 3</b>
<b>1.1. Conservare e preservare</b>	<b>9</b>
1.1.1. Conservazione/preservazione delle risorse digitali.	9
1.1.2. La conservazione delle risorse digitali è un interesse interdisciplinare	15
<b>1.2. Strategie, metodi, criteri e sistemi di preservazione</b>	<b>34</b>
1.2.1. Strategie di una “politica di conservazione”, reti e modelli di deposito	36
1.2.2. Le risorse digitali e i metodi di preservazione	46
1.2.3. Web archiving: sistemi e criteri di raccolta per la conservazione dei siti Web	53
<b>1.3. Alcuni studi sulla conservazione digitale</b>	<b>60</b>
<b>2. Alcuni esempi di progetti e iniziative per la conservazione dei documenti digitali</b>	<b>65</b>
<b>2.1. Iniziative e progetti internazionali</b>	<b>65</b>
2.1.1. The Internet Archive	65
2.1.2. The Nordic Web Archive	70
2.1.3. In Gran Bretagna: CEDARS, UKWAC, UK Government Web Archivi	74
2.1.4. International Internet Preservation Consortium (IIPC)	77
<b>2.2. Iniziative e progetti nell’Unione Europea</b>	<b>79</b>
2.2.1. Networked European Deposit Library (NEDLIB)	83
2.2.2. Electronic Resource Preservation and Access Network (ERPANET)	85
2.2.3. DigitalPreservationEurope ( DPE)	87
2.2.4. DELOS (Network of Excellence on Digital Libraries)	94
<b>2.3. Australia: PANDORA</b>	<b>96</b>
2.3.1. Politica di raccolta e Linee guida.	98
2.3.2. Archiviazione - Catalogazione – Accesso	101
2.3.3. Copyright e Deposito legale	102
2.3.4. Infrastruttura PANDAS e formazione del personale	104
2.3.5. La preservazione digitale e il ruolo di PADI	109
<b>2.4. Norvegia: il progetto PARADIGMA</b>	<b>110</b>
2.4.1. Il contesto	111
2.4.2. La legge sul Deposito legale	112
2.4.3. La raccolta dei documenti digitali	115
2.4.4. Un archivio digitale progettato per la ricerca: FRBR	117
2.4.5. I documenti Internet: dinamici e specifici	126



# 1. Conservare la memoria digitale:

## questioni metodologiche

Questa tesi è volta ad un'analisi dello stato dell'arte sulla preservazione digitale, per conoscere gli approdi della ricerca e le applicazioni proposte, con l'obiettivo di delineare gli attuali strumenti a disposizione per affrontare la responsabilità della conservazione di questo patrimonio e garantirne il pubblico accesso nel tempo; allo scopo si fornisce un panorama delle strategie e dei criteri possibili, degli studi, di alcuni esempi di sistemi di archiviazione ed delle questioni giuridiche coinvolte.

Questo interesse è stato prepotentemente indotto dalla considerazione di quanto sia stato in concreto trascurato il dovere di conservazione rispetto a questa tipologia di documentazione e di prodotti culturali: molti luoghi deputati istituzionalmente alla conservazione della memoria, mentre si fregiano di aver informatizzato alcuni servizi e di mantenere aggiornate delle proprie pagine web, di aver aperto sale per fornire servizi multimediali, di possedere materiale digitali, di realizzare digitalizzazioni del proprio patrimonio, in realtà hanno spesso disatteso l'esigenza di un più complessivo 'progetto digitale' e il contestuale, dovuto e necessario impegno nella conservazione del digitale.

Nel voler tracciare un sintetico ma efficace quadro di riferimento sugli aspetti teorici, organizzativi ed applicativi, del contesto tecnologico, politico, istituzionale, è stata riscontrata una grande difficoltà, per la complessità del problema e dei percorsi finora battuti, per la frammentazione dei progetti, che spesso sono ancora in piena fase di indagine, sperimentazione e verifiche, per la ricchezza delle fonti a disposizione, di ricerche e bibliografia in continua produzione, indice di una discussione incessante e ben rappresentato dal rilevante riferimento che in questo studio è stato fatto ai convegni, nazionali e internazionali, singoli, ricorrenti o programmati all'interno di iniziative coordinate, di cui (tanto per ribadire l'importanza della documentazione presente su Web) pre-print e relazioni sono spesso reperibili soltanto via Internet. Per esempio, appena terminata questa ricerca, si preparava un seminario in cui Mariella Guercio<sup>1</sup> ha proposto lucide considerazioni sul problema della conservazione digitale che ancora non ha visto raggiunti i risultati sperati.

---

<sup>1</sup> Guercio, Mariella. *I metadati per la conservazione a lungo termine delle risorse elettroniche*. In: *Interoperabilità di contenuti e servizi digitali: metadati, standard e linee guida*. Seminario nazionale di studi : Roma, 3 aprile 2007. A cura del

Quindi la ricerca per affrontare la sfida della sparizione della quasi totalità delle memorie contemporanee è ancora tutta aperta, ma ci sono comunque delle conclusioni che si sono potute trarre e che forse si potrebbero ricondurre ad una concezione globale: responsabilità.

È necessario costruire una nuova consapevolezza culturale e professionale intorno alla conservazione, anche nel sistema delle biblioteche pubbliche, perché possano affrontare nelle realtà locali problemi di sviluppo delle raccolte digitali, anche in assenza di una forte strategia conservativa che imponga soluzioni a livello istituzionale e trovi momenti di raccordo in servizi centralizzati per le politiche di stoccaggio.

Perché la preservazione della memoria digitale è un'azione che passa attraverso un complesso intreccio di problematiche solo in parte risolte, ma che, se si basa su un forte senso della "responsabilità conservativa", rende dirigenti e gestori perseveranti nella ricerca e nello studio di soluzioni, rigorosi nell'adozione delle procedure ormai consolidate, scrupolosi nel doversi fornire dei requisiti necessari per la loro attività.

La preservazione digitale è una branca della conservazione, speciale, ma pur sempre l'adeguamento alla contemporaneità di una tradizionale attività dei conservatori, professionisti in possesso di consapevolezza concettuale, culturale e politica, che, quindi, devono munirsi della padronanza metodologica, organizzativa e regolamentare sul problema della preservazione e della gestione dei depositi di risorse digitali: saper capire i sistemi tecnologici, ricercare la competenza tecnica, provvedere alla diffusione di livelli adeguati di formazione e specializzazione, promuovere azioni di condivisione.

Le istituzioni cui è demandata la responsabilità della conservazione della memoria devono affrontare il compito di raccogliere, organizzare, conservare e mettere a disposizione le risorse culturali e storiche della società, quelle prodotte in passato e quelle prodotte attualmente: tra queste vi sono le risorse digitali. Sicuramente si può vedere una continuità di funzione per le istituzioni designate, che non possono delegare ad altri questa cura e omettere la propria parte di contributo: con la consapevolezza della propria mission istituzionale, occorre far sì che problemi economici non diventino scusanti, saper "essere avanti", guardare ai valori in prospettiva, cercare soluzioni e modi per raggiungere lo scopo.

Il problema del sottovalutato dovere documentale e conservativo dell'attività e della produzione digitale nasce soprattutto presso quelle istituzioni preposte alla ricerca e allo studio, ai servizi culturali, alla diffusione della conoscenza, ma in particolar modo in quelle dedite alla salvaguardia dei beni culturali. Mentre ci si aspetterebbe fossero i primi organismi a sentire questa necessità, invece, quest'impegno è stato, almeno finora, per lo più disatteso.

Tra le risorse da preservare, i siti Web manifestano aspetti più complessi e delicati rispetto alla conservazione delle risorse elettroniche locali: capita facilmente di constatare come istituzioni, che, pure, sin dalla loro origine hanno sempre fatto attività di conservazione nei propri archivi, non hanno dimostrato la sensibilità culturale della preservazione di questa tipologia di oggetti digitali, non hanno dato il doveroso spazio all'esigenza di conservare almeno in parte la propria produzione Web; responsabili di documentazione on line, veri "soggetti produttori" dei siti, non si sono minimamente preoccupate anche della loro preservazione. Inoltre, una variante dello spin over politico ha colpito nelle pubbliche amministrazioni, tanto nei siti Web appartenenti agli organi governativi, ai Ministeri per esempio, quanto nei governi locali; così pure nelle istituzioni di conservazione, centrali e periferiche, come per esempio nelle biblioteche statali, che pure hanno (paradosso dei paradossi) per regolamento l'obbligo di deposito e conservazione della produzione locale, addirittura cestinano la propria passata produzione Web. In questa mancanza si potrebbe ravvisare anche una sorta d'incapacità di riconoscere il valore del documento digitale. Per esempio, non tutti hanno riconosciuto come il sito Web sia uno strumento che obbedisce a determinate regole di progettazione basate sulle esigenze del soggetto che lo attiva, struttura e grafica costituiscano elementi di contesto importanti, capaci in sé di segnalare modifiche strutturali verificatesi al livello del soggetto produttore, evoluzioni o involuzioni, che tradiscono una diversa consapevolezza dell'uso dello strumento e quindi un mutamento sensibile nelle strategie scientifiche e comunicative dell'istituto che lo ha rilasciato on line. Solo soggetti terzi si sono preoccupati, nell'ambito di progetti più o meno strutturati e finalizzati, di raccogliere le risorse della memoria dal Web a futura conoscenza.<sup>2</sup>

Eppure gli oggetti digitali sempre più fanno parte della produzione culturale, del ciclo creativo e lavorativo, a livello individuale e collettivo, pervadendo molte delle nostre attività quotidiane, e le risorse elettroniche sono parte integrante del patrimonio di biblioteche ed archivi, sono stati sviluppati la digitalizzazione delle raccolte e l'accesso elettronico a servizi, uffici, documentazione e iniziative: rappresentano l'elemento sostanziale nella società dell'informazione, quindi l'attenzione alla conservazione a tempo indeterminato dei contenuti digitali deve fare parte integrante dell'idea stessa di gestione del patrimonio digitale, anche perché contano sempre di più nella nostra società l'informazione in generale e i prodotti documentari, sia nel settore pubblico sia in quello privato, per la libera informazione, lo sviluppo delle conoscenze, la comunicazione, la formazione, i processi decisionali, lavorativi, tecnici e di ricerca.

---

<sup>2</sup> Per molti di questi casi l'unico deposito in cui reperire alcune istantanee di vita, con tutte le limitazioni che pure questo strumento comporta, è l'Internet Archive. V. nel Cap. 2.

Ma non si può nascondere che siamo di fronte ad ‘oggetti’ totalmente nuovi: il Subject di cui parliamo non è un oggetto tangibile, ma un concetto extra-ordinario, che non è stato ancora compreso a fondo.<sup>3</sup> Il materiale in questione è molto diverso da quelli fino ad oggi trattati e, per molti versi, appare insidioso e di rottura verso le certezze su cui si basavano le consuetudini conservative. Soprattutto le qualità di queste nuove modalità di trasmissione della conoscenza aprono scenari di cui ancora non si vede bene l’ampiezza: ancora non ci si è del tutto ripresi dalla difficoltà di comprendere a pieno questi nuovi ‘oggetti intangibili’ per la loro gestione, descrizione ed uso, che si sono affacciate l’idea di una responsabilità di preservazione allargata a diversi protagonisti e la necessità di una consapevolezza diffusa del valore e della qualità dell’informazione, forse anche di un’attenzione maggiore, proprio in un momento di surplus e di eccessi, di un vero consumismo dell’informazione e di un atteggiamento ‘usa e getta’ delle risorse. Sin dalla metà degli anni Novanta, con l’avvento e poi il vorticoso incremento della produzione di documentazione e di pubblicazioni elettroniche, si sono presentate le diverse problematiche relative alla loro gestione: la necessità della loro descrizione/catalogazione, per le diverse finalità, e poi gli aspetti legali, connessi all’accessibilità e al deposito legale. A ciò, molto presto, si è aggiunto il problema della loro conservazione/preservazione, finalizzata alla salvaguardia nel tempo della loro integrità, non disgiunta dall’accessibilità e dalla fruibilità necessarie. Per le proporzioni della crescita di documenti, per il ‘rumore’ che connota il mondo di Internet, importante è affrontare, inoltre, il problema dei criteri di scelta del patrimonio da conservare: la selettività è indispensabile per realizzare una vera conservazione della memoria della nostra contemporaneità digitale.

All’adempimento di questo complesso di funzioni si deve aggiungere il quesito su cosa debba essere il deposito conservativo del materiale digitale. Al di là del nominarlo come una biblioteca o un archivio delle risorse digitali, o in altri modi ancora, si può senz’altro affermare che deve trattarsi di un deposito volto all’uso, di un servizio che, più che essere duplice, di conservazione e di documentazione, dovrà essere “completo” per essere efficace. Quindi non si potranno trascurare continui studi sul trattamento e sul processo descrittivo dei documenti digitali, che serviranno alla gestione concreta, per l’uso e per la conservazione, ma, anche, per la possibilità di inserire queste risorse in sistemi informativi, che ne permettano una vera fruizione anche in futuro: non solo per recuperarle in sé, ma, nonostante si tratti di risorse con una forte tendenza alla dispersione e alla ridondanza, per mettere in condizioni di comprendere il contesto dell’informazione che rappresentano, di muoversi

---

<sup>3</sup> Al proposito si rimanda all’esplicativo scritto di Negro Ponte (1995).

nella loro fitta rete di legami con altre conoscenze, di capire l'evoluzione del progetto e dell'idea di cui sono portatrici, come ogni altro documento.

In molti casi, i primi grandi esempi di conservazione di risorse elettroniche non sono stati dettati da una radicata, attenta e complessiva visione conservativa, bensì da necessità contingenti al diffondersi di intere collezioni di giornali elettronici, ove solo la copia digitale e non anche quella cartacea era disponibile, con le problematiche proprie della 'biblioteca digitale' intesa come biblioteca di collezioni digitali. Le collezioni digitali comprendono materiali diversi dal punto di vista bibliografico, originariamente digitali o digitalizzazioni di documenti analogici, di accesso locale o remoto (come periodici elettronici digitali, periodici in versione digitale di equivalenti a stampa, e-book, digitalizzazioni di collezioni cartacee, pubblicazioni informali<sup>4</sup> come gli e-print, sia pre-print, sia post-print). Esse necessitano di un'infrastruttura in grado di permettere un sistema di archiviazione, nonché di una coerente organizzazione per offrire un efficiente servizio; capita spesso, invece, che già nei siti Web, banche dati bibliografiche, risorse Internet selezionate, periodici elettronici, CD-Rom in rete, e-book e quant'altro, si trovino sparse, non siano catalogate, raramente sommariamente descritte, non siano cioè recuperabili dall'utente.<sup>5</sup> Va compreso che costruire una vera biblioteca digitale, organizzare in un sistema informativo l'accesso al sapere digitale (che non possiamo distinguere sempre tra attuale e non più corrente..) è già anche preservare.

E significa preservare anche catalogare: per questo non ha senso il non catalogare risorse digitali, tanto più che la stessa catalogazione si evolve per comprendere nuove risorse, anche quelle che non avendo la qualità della permanenza fino ad un'ufficiale 'nuova edizione', non rispondono alle tradizionali definizioni e al concetto dell'"oggetto della catalogazione", negando il binomio opera-supporto e presentando l'Opera (che, comunque, secondo il modello FRBR resta un concetto astratto) come l'entità che ingloba tutte le sue molteplici manifestazioni che si avvicendano nel corso del tempo.<sup>6</sup> L'evoluzione è nel senso del focalizzare i due obiettivi della catalogazione: descrizione e accesso.<sup>7</sup>

All'attuazione dell'accesso è legata anche la progressiva completa standardizzazione (l'adozione di linguaggi e software standard) e l'uso di linguaggi orientati alla struttura

---

<sup>4</sup> *Manifesto per le biblioteche digitali*. AIB 2005, aggiorn. 2005-12-20.

<<http://www.aib.it/aib/cg/gbdigd05a.htm3>>. Principi, 6: «Le biblioteche digitali diffondono i documenti. Le biblioteche digitali consentono al massimo grado la conservazione e la diffusione, tra gli altri documenti, di pre-print, rapporti interni, dispense, progetti di studio, protocolli sperimentali, pubblicazioni istituzionali, e altri e-print altrimenti relegabili nella cosiddetta "letteratura grigia"».

<sup>5</sup> Tammaro (2005) e (2006), Giordano (2007).

<sup>6</sup> Il problema del trattamento a fini di recupero bibliografico e all'applicazione di FRBR viene trattato ampiamente nel progetto d'archiviazione Web elaborato dalla National Library of Norway, di cui si parlerà nel Cap.2.

<sup>7</sup> Il codice di catalogazione internazionale AACR2 Anglo-American Cataloguing Rules dal 2004 è stato oggetto della revisione che lo ha portato alla sua nuova edizione, che assumerà la denominazione *RDA: Resource Description and Access*. (Fonte: Guerrini (2005)).



logica del documento: questi costituiscono, al momento, lo strumento tecnico fondamentale per la preservazione delle risorse digitali. Essenziale è XML,<sup>8</sup> il linguaggio autodichiarante che reca testo e specifiche: ciò significa, in concreto, riuscire a ‘leggere’ il documento elettronico, prigioniero della macchina e delle tecnologie, riportare alla comprensione naturale umana il contenuto digitale, ma anche permettere la conoscenza di questo ‘materiale virtuale’, perché la descrizione dichiarativa reca informazioni ed esplicazioni sul contesto culturale di riferimento dell’oggetto.

Le soluzioni attuali per intraprendere l’attività di preservazione sono riconosciute da Mariella Guercio, nel suo intervento citato in precedenza, in tre punti sostanziali: il sistema OAIS,<sup>9</sup> dal quale si dovrebbe ricominciare a riflettere per future evoluzioni, l’uso di XML e una rinnovata riflessione dei singoli ambiti disciplinari coinvolti, per trovare risposte di principio e di metodo.

Nel problema della preservazione c’è una forte complessità e ‘interdisciplinarietà’: è trasversale il problema dell’obsolescenza, che riguarda i supporti e le tecnologie digitali tutte, indifferentemente quindi dal contenuto o dalla tipologia o dal contesto (beni culturali, risorse scientifiche, documenti amministrativi) e ormai è sempre più facile dover trattare sistemi complessi dove gestire oggetti e documenti, contenuti e contenitori. Quindi esistono reali bisogni comuni e, anche dal punto di vista metodologico e organizzativo, diviene possibile una crescita dei campi e degli spazi di reciproco o comune interesse; come nel corso della ricerca verrà dimostrato, il dialogo interdisciplinare costituisce un requisito imprescindibile per raggiungere traguardi, la determinazione di strumenti e procedure adeguate ad una funzione così complessa. Ma poi occorre anche il percorso contrario, cioè un ritorno alla concentrazione sui propri ambiti, alla riflessione e all’approfondimento di problemi di principio, di metodo e di definizioni soddisfacenti all’interno delle specifiche discipline, per contribuire ad una maggiore consistenza di criteri, di soluzioni ai nodi ancora irrisolti, di una maggiore consapevolezza e padronanza del proprio settore. Insomma, per organizzare l’accesso alla conoscenza, passata e attuale, si deve riscoprire la disciplina cui si afferisce e trovarvi le risorse per sciogliere i problemi.

Al termine della ricerca, appare questo come un vero e proprio richiamo al proprio compito, a riflettere sulle funzioni proprie, al ritrovare le proprie responsabilità. In una parola, riscoprire il proprio lavoro, con la ricerca specifica, i principi, gli ambiti, le funzioni e le responsabilità. Responsabilità come la conservazione.

---

<sup>8</sup> World Wide Web Consortium. *eXtensible Markup Language*. <<http://www.w3.org/XML/>>.

<sup>9</sup> Si veda al Cap. 1. 2.

Il metodo è quello del “Chi può dire Come”: anche alle prese con nuovi strumenti e oggetti, il documentalista, l’archivista e il bibliotecario, grazie alla loro formazione, mantengono il legame sia con l’oggetto, sia con i fruitori, e possono sciogliere i problemi dell’accesso, anche nel lungo periodo. Essendo gli studiosi di questo tipo di mediazione, conoscono gli strumenti per ‘saper guardare’ l’oggetto con la sua storia, comprenderne limiti e potenzialità e trattarlo per poterlo mettere a disposizione (ben sapendo che continuerà a servire e dovrà poter essere ritrovato anche in un futuro lontano) del pubblico. Pubblico, reale e potenziale, di cui possono studiare le esigenze, prevedere i bisogni, approfondire e stimolare le aspirazioni di conoscenze: cioè occuparsi dell’accesso profondo e completo, puntuale e, se necessario, aldilà del tempo, cioè realizzare lo scopo del loro lavoro.

## **1.1. Conservare e preservare**

### 1.1.1. Conservazione/preservazione delle risorse digitali.

Per iniziare la ricerca si è pensato utile fare un breve esame su cosa s’intenda per “conservazione” e “preservazione”, e riportare un *excursus* tra le riflessioni nate negli ambienti biblioteconomici e archivistici nell’affrontare il problema della conservazione della memoria digitale, alla ricerca di approcci precisi ed esaustivi, definizioni, analisi sulle risorse in oggetto, identificazioni di responsabilità. Questo momento di ricerca iniziale sta anche a ricordare quanto il dialogo interdisciplinare costituisca un requisito imprescindibile per l’avanzamento della ricerca e per la definizione di strumenti e procedure adeguate ad una funzione così complessa.

Spesso, proprio in merito alle risorse digitali, è utilizzato anche in Italia il termine “preservare”, al posto del più comune “conservare”: una sottile differenza distingue i due vocaboli, o, meglio, una maggiore completezza di significati assume il termine “preservazione”. Pertanto si ritiene di poterlo utilizzare insieme al consueto “conservazione”, non in contrapposizione, ma come rafforzativo, apportando esso al discorso – e alla volontà d’azione – un bagaglio d’implicazioni e di raccomandazioni che arricchiscono di senso etico questa funzione indispensabile per la trasmissione della conoscenza e del patrimonio contemporaneo alla posterità.

Nel seguente prospetto si riporta una selezione dei risultati di una ricerca terminologica per l'uso di «Preservazione»/«Conservazione».

### digital preservation

E-Lib Project, *Working definitions of commonly used terms for the purposes of the Cedars project*, <<http://www.leeds.ac.uk/cedars>>

#### to preserve

*Guidelines for the selection of online Australian publications intended for preservation by the National Library of Australia*, <<http://pandora.nla.gov.au/guidelines.html>>.

Wikipedia, <<http://en.wikipedia.org/wiki/Digital-Preservation>>

#### Stephen Chapman

(Preservation Librarian di Digital Initiatives al Weissman Preservation Center, Harvard University Library), <<http://www.oclc.org/education/conferences/presentations/2001/preservation/chapman.htm>>.

È il termine usato in ambito anglosassone per indicare le problematiche di immagazzinamento, mantenimento e accesso nel lungo periodo rispetto ad oggetti digitali.

«Digital preservation is storage, maintenance and access to digital objects/materials over the long term».

«Is used to mean taking certain steps including archiving, organising, describing, refreshing and migrating of titles, to ensure providing long term access».

«Digital preservation refers to the management of digital information over time. Unlike the preservation of paper or microfilm, the preservation of digital information demands ongoing attention. This constant input of effort, time, and money to handle rapid technological and organisational advance is considered the main stumbling block for preserving digital information beyond a couple of years. Indeed, while we are still able to read our written heritage from several thousand years ago, the digital information created merely a decade ago is in serious danger of being lost. Digital preservation can therefore be seen as the set of processes and activities that ensure the continued access to information and all kinds of records, scientific and cultural heritage existing in digital formats.

Digital preservation is defined as: long-term, error-free storage of digital information, with means for retrieval and interpretation, for all the time span that the information is required for. "Retrieval" means obtaining needed digital files from the long-term, error-free digital storage, without possibility of corrupting the continued error-free storage of the digital files».

«I trust that this audience will not be surprised by a librarian taking the prerogative of classifying the question before giving an answer! Preservation administrators generally view digital preservation as a sub-topic of library preservation, which has a single overarching mandate: to facilitate use. In the ideal, preservation policies and services would ensure that materials in all formats would be accessible for use in perpetuity».

(cont. **digital preservation** )

Joint Information Systems Committee (JISC)  
<[http://www.jisc.ac.uk/publications/publications/pub\\_digipreservationbp.aspx](http://www.jisc.ac.uk/publications/publications/pub_digipreservationbp.aspx)>.

UKOLN. *Good Practice Guide for Developers of Cultural Heritage Web Service*  
<<http://www.ukoln.ac.uk/interop-focus/gpg/Preservation/>>.

Faiella (2000): rimarca la valenza della coppia terminologica *conservation/preservation*

«Preservation is therefore a more immediate issue for digital than for traditional resources. Digital resources will not survive or remain accessible by accident: pro-active preservation is needed. Digital Preservation requires not only the maintenance and disaster recovery procedures needed for securing the media and its contents (often referred to as 'archiving') but also strategies and procedures to maintain its accessibility and authenticity over time. This is challenging as often it can require collaboration between different stakeholders, good licensing practice, the association of appropriate metadata and creation of sufficient documentation before technical issues can be addressed».

«Basic digital preservation is the **maintenance** of digital material over the long-term with a view to ensuring continued **accessibility**. Digital material refers to any material processed by a computer and includes both that which is “**digitised**” (reformatted to digital) as well as those resources that are “**born digital**”. Long-term in this context should be taken to mean long enough to be concerned with the impact of changing - and should include timescales of decades and even centuries. Digital preservation is the series of actions and interventions required to ensure continued and reliable access to authentic digital objects for as long as they are deemed to be of value. This encompasses not just technical activities, but also all of the strategic and organisational considerations that relate to the survival and management of digital material».

«L’obbiettivo è di assicurare che il contenuto intellettuale di un documento in forma digitale rimanga accessibile per le generazioni future, sia nel caso in cui si tratti di informazioni nate digitali, sia nel caso di conversioni digitali di materiale analogico. Mantenere anche in italiano la distinzione tra i due termini potrebbe essere funzionale per segnalare un binomio semantico che ha i suoi corrispettivi sia nell’ambito delle modalità di trasmissione dell’informazione (analogico/digitale) che di tipologia dell’intervento (curare/prevenire); quindi “preservazione” indicherebbe meglio un’attenzione funzionale non solo a contenere ma a prevenire gli effetti dell’obsolescenza».

## Preservation digitisation

Terminologia presente da anni in molti progetti statunitensi, anche nel settore dei periodici, di digitalizzazione:

CLASS della Cornell University, Open Book della Yale University, o l'esperienza di JStor :

Art And Humanities Data Service. *Digital preservation*. <<http://ahds.ac.uk/digital.htm>> .

S. Michael Malinconico. *Biblioteche digitali: prospettive e sviluppo*. "Bollettino AIB", 1998, 3, p. 275-299.

< <http://www.aib.it/aib/boll/1998/98-3-275.htm> >.

Paul Conway. *Yale University Library's project Open Book*. "D-Lib magazine". February 1996.

<<http://www.dlib.org/Dlib/february96/yale/02conway.html>>.

Kelly Russell. *Digital preservation: ensuring access to digital materials into the future*. 1999.

<<http://www.leeds.ac.uk/cedars/Chapter.htm#3>>.

Conservazione di documenti informativi analogici attraverso la loro digitalizzazione.

«The phrase digital preservation refers to the preservation of digital materials and to the preservation of paper based materials and other artefacts through their digitisation».

## Conservazione digitale

Gloria Cirocchi. *Conservazione di risorse digitali: quali sfide?* "Bollettino AIB", 1999, 3, p. 289-300,

<<http://www.aib.it/aib/boll/1999/99-3-289.htm>>

Indica tendenzialmente il processo di digitalizzazione di documenti analogici: il formato digitale come mezzo di conservazione (che a sua volta diverrà anche oggetto di conservazione).

## Conservazione del digitale

### Conservare

De Mauro. *Il dizionario della lingua italiana*.  
Paravia versione on-line.  
<<http://www.demauroparavia.it/26292>>.

### Conservazione del documento/memoria/informazione/risorsa digitale.

Gloria Ciocchi. *Conservazione di risorse digitali: quali sfide?* "Bollettino AIB", 1999, 3, p. 289-300, già cit.  
Faiella (2000)  
Tammaro (2004):

"*News ICR : la newsletter dell'Istituto Centrale per il Restauro*", [I. C. R., Min. B.A.C.]  
*Conservazione dei documenti digitali - Conservare il digitale: il Web*.  
<[http://www.icr.beniculturali.it/news\\_icr/news\\_icr\\_02.htm](http://www.icr.beniculturali.it/news_icr/news_icr_02.htm)>.  
Uguale spiegazione è data nel Sito Web I. C. C. U., sezione BDI (Biblioteca Digitale Italiana) - Conservazione dei documenti digitali,  
<<http://www.iccu.sbn.it/genera.jsp?s=75#>>.

Lunghi. *DigitalPreservationEurope (DPE)*.  
In Lunghi (2006)

1a: **tenere qcs. in modo che non si sciupi, vada perduto e sim.** 1b: **fig., non far venir meno, mantenere.** 2: **proteggere, difendere; preservare.**

È la più comune traduzione dell'espressione **digital preservation**.

«...La conservazione del digitale consiste nella pianificazione, nello stanziamento delle risorse e nell'applicazione dei metodi e delle tecnologie per la conservazione, necessari ad assicurare che l'informazione digitale di valore duraturo (continuing value) rimanga accessibile e utilizzabile».

«La conservazione, intesa tradizionalmente, consiste nell'assicurare la stabilità nel tempo dei supporti fisici, in modo da assicurare l'accesso all'informazione da essi veicolata. Oggi la crescente quantità di informazione in formato elettronico ha spostato l'attenzione su nuovi fattori di rischio, legati non solo alla labilità dei supporti, ma al rischio di veloce obsolescenza di sistemi operativi e formati. Il problema dell'obsolescenza coinvolge tutti i documenti elettronici, quelli ad accesso locale (Cd-Rom, DVD) e quelli ad accesso remoto, quelli nati digitali e quelli derivati per digitalizzazione da precedenti documenti cartacei.

Il mondo del web è, oggi, il "luogo" della mancata conservazione. Non c'è, forse, ambiente maggiormente volatile: 44 giorni è la durata media di una pagina web».

«La conservazione delle memorie digitali non è meramente una questione tecnica, ma anche una questione di assegnazione di responsabilità e di processi decisionali e discrezionali che implicano anche questioni etiche almeno per le procedure selettive e descrittive dei contenuti stessi».

## Preservazione del digitale

### Preservare

De Mauro. *Il dizionario della lingua italiana*. Paravia versione on-line.  
<<http://www.demauroparavia.it/26292>>.

### Preservazione

Giacomo Devoto - Gian Carlo Oli. *Il dizionario della lingua italiana*. Firenze :  
Le Monnier, 2003.

Anna Maria Tammaro in Tammaro –  
Salarelli (2006)

**1a: proteggere, salvaguardare da danni, pericoli e sim. 1b: premunire qcs. dall'azione nociva di agenti esterni.**

L'azione intesa a garantire dalle insidie dell'ambiente e del tempo.

«La preservazione è quella funzione della biblioteca digitale che comprende le attività di memorizzazione dell'informazione su supporti digitali, il mantenimento delle collezioni digitali, l'accesso permanente e la distribuzione dei documenti digitali. La preservazione gestisce quindi l'intero ciclo di vita del documento digitale, con un'enfasi sull'accesso di lunga durata all'informazione digitale, conservata nella sua integrità...».

«Preferisco usare questo termine, piuttosto che altri diffusi in Italia come conservazione digitale o prevenzione, perché ben rappresenta la dinamicità dell'attività che definisce».

«La preservazione è quindi una funzione critica per la biblioteca digitale, gli ostacoli da rimuovere non sono solo la fragilità del supporto e l'obsolescenza tecnologica. Il problema della preservazione è un problema sia tecnico che politico e verrà trattato da entrambi i punti di vista».

### 1.1.2. La conservazione delle risorse digitali è un interesse interdisciplinare

Conservare la memoria digitale è un'esigenza globale, trasversale e interdisciplinare.

Per quanto riguarda la memoria culturale digitale, per esempio, la cooperazione e lo scambio di conoscenze tra il mondo degli archivi, delle biblioteche e anche dei musei, sono importanti per affrontare le problematiche nel modo più efficace possibile.

Poiché i progetti ad oggi più riusciti sono basati su politiche di cooperazione e di inter/multi-disciplinarietà, sembra opportuno riportare alcune riflessioni nell'affrontare problematiche comuni, maturate negli ambiti archivistici oltre che in quelli biblioteconomici, relativamente alla conservazione e alle sue possibili politiche, alle risorse in oggetto, ai soggetti responsabili. Che l'interdisciplinarietà sia un fattore incisivo nei settori della produzione e tenuta delle risorse digitali è attestato in molti progetti, ai quali si farà cenno in questa ricerca, in quanto è il metodo capace di fornire concentrazione e efficacia nell'utilizzo di principi, strumenti e tecniche presenti nelle diverse discipline (in questo caso biblioteconomiche, archivistiche, documentarie e informatiche).

In particolare, qui si sono volute riportare alcune riflessioni d'ambito dell'archivistica, oltre alle scienze biblioteconomiche, che con le proprie analisi hanno contribuito all'identificazione delle problematiche e alla ricerca di soluzioni.

L'interdisciplinarietà, per la comprensione stessa, l'analisi, lo studio, il reperimento di soluzioni adeguate, con ottiche di lungo respiro e di più ampie prospettive, è percepita come bisogno ormai ineludibile. A ciò ha contribuito anche il fatto che ormai, sempre più spesso, gli oggetti di studio non sono più così nitidamente afferenti ad un ambito disciplinare piuttosto che ad un altro, ma presentano spesso ambiguità e caratteristiche sulle quali solo una convergenza di analisi che partono da differenti punti di vista, può arrecare un reale beneficio:

«Le cosiddette nuove tecnologie hanno colmato o stanno colmando la distanza che fino a ieri separava il documento – inteso come manufatto, oggetto 'fisico' conservato presso un archivio o una biblioteca – dall'edizione del documento – la sua edizione a stampa, esito di una mediazione affidata a pratiche storiografiche correnti, a dimensione specialistica».<sup>10</sup>

I progetti internazionali più avanzati propongono un riconoscimento condiviso dei problemi attraverso una visione multidisciplinare, interdisciplinare e trans-disciplinare,

---

<sup>10</sup> Michele Ansani. *Diplomatica e nuove tecnologie. La tradizione disciplinare fra innovazione e nemesi digitale*. 2001. Disponibile all'URL: <<http://scrineum.unipv.it/rivista/ansani.html>>.



generale e globale.<sup>11</sup> È chiaro che a questo consenso al riconoscimento, o identificazione, delle problematiche, le risposte sono molteplici e legate ai diversi contesti culturali, ma il contributo, che l'unione di 'diversi con gli stessi intenti' fornisce, è insostituibile.

Secondo Federico Valacchi<sup>12</sup> un 'interscambio' è possibile, ad esempio, con l'archivistica. Una disciplina come l'archivistica può svolgere due ruoli nei confronti della preservazione digitale: da una parte indagare nella Rete per identificare e comprendere le nuove forme di tipo archivistico e documentario e, dall'altra, nel campo della conservazione del Web, può essere utile, fornendo strumenti per la comprensione di molti degli aspetti e dei contenuti del Web stesso. Ma ci si domanda, anche, come la comunità archivistica possa portare il proprio contributo alla soluzione della perdita totale della memoria telematica e fornire strumenti per la preservazione di almeno una parte di questa memoria.

Per quanto riguarda l'indagine sui nuovi contenuti, Valacchi, partendo dalla constatazione che esistono nuove pratiche digitali, come i cosiddetti archivi ubiqui (risorse documentarie digitali in Internet), gli archivi nomadi (risorse digitali off line), gli archivi inventati<sup>13</sup> (cioè le collezioni di documenti provenienti da fondi diversi e aggregati su base tematica e volontaria) e riconoscendo che vi è, quindi, la manifestazione di una nuova tipologia di archivio, di aggregazione «... destinata a vivere e, si spera, a sopravvivere solo nel Web, utilizzando le risorse e le modalità di sedimentazione tipiche di questo strumento», afferma che per l'archivistica c'è una marcata esigenza:

«... più approfondite valutazioni delle modalità secondo le quali l'universo documentario viene riorganizzandosi, all'interno del contesto digitale, e delle possibilità che le scienze documentarie hanno di controllare e descrivere queste nuove sedimentazioni».

Nel Web esistono 'archivi' creati diversamente, ma anche destrutturazioni degli archivi, o di parti di essi, per comporre entità che non hanno i tradizionali vincoli archivistici: paternità archivistica, sedimentazione con una continuità temporale dipendente dall'attività del produttore, necessarietà, involontarietà del vincolo, un solido rapporto tra un soggetto produttore chiaramente identificabile e le forme di sedimentazione della sua memoria. Sono invece aggregazioni documentarie che si formano secondo modelli di 'volontà deliberata' di creare sedimentazioni in ragione di una domanda espressa al di fuori degli abituali meccanismi di produzione, gestione e conservazione, ma per contaminazioni documentarie o per attività di un soggetto che ha nel sito Web una porzione del proprio archivio.

---

<sup>11</sup> Luciana Duranti, intervento al convegno internazionale Conservare il digitale (2006).

<sup>12</sup> Valacchi (2002), p. 10-16, da cui sono presi molti degli incisi qui riportati, quando non diversamente specificato.

<sup>13</sup> Citati anche in Vitali (2004 a), p. 107.

«Infatti, nel momento in cui si assume che nel Web prendono forma nuove tipologie di sedimentazioni documentarie e che comunque il Web è uno dei luoghi in cui si manifesta la futura memoria della nostra società, è inevitabile porsi il problema della conservazione di questa memoria. [...] La crescita costante dell'uso delle risorse telematiche e la trasformazione della rete da luogo della comunicazione a teatro della gestione di processi sempre più complessi ha fatto sì che non si possa non preoccuparsi di se, dove e come sia possibile riuscire a fermare nel tempo e nello spazio l'evanescente universo documentario che attraversa il Web, per farne potenziale testimonianza giuridica e "fonte"».

Così, tra le discipline che si devono occupare della conservazione della memoria su Internet, si delinea una nuova specializzazione dell'archivistica:

«[...] l'archivistica telematica, l'ultima nata delle molte specializzazioni che una disciplina necessariamente duttile come l'archivistica deve accogliere nel suo ambito»

e, continuando sempre sulle riflessioni alla ricerca di soluzioni di conoscenza, afferma:

«Prima di preoccuparci di capire se il sito Web sia o meno un archivio nel senso proprio del termine – e la questione è tutto sommato irrilevante di fronte ai rischi di dispersione della memoria – occorre verificare se definizione, modelli concettuali e prassi dell'archivistica servono per gestirlo. [...]

Io credo che siano soprattutto i metodi e l'approccio intellettuale tipici della scienza archivistica, piuttosto che gli strumenti di cui oggi disponiamo, a rivelarsi utili a gestire questi problemi. Non si tratta cioè di tentare di ingabbiare il Web dentro alla griglia codificata del pensiero archivistico ma di capire come certe peculiarità del pensiero archivistico possano contribuire a dipanare la matassa del Web. Un solo esempio: la dimestichezza degli archivisti con le pratiche di selezione della documentazione può dare indicazioni importanti per cercare di aprire sentieri percorribili nella foresta della memoria telematica, dove la selezione è più che mai un elemento qualificante delle fonti».<sup>14</sup>

Stefano Vitali affronta ampiamente il rapporto tra il mondo del Web e l'archivistica.

«Già oggi il Web è un formidabile strumento di comprensione del presente – e tanto più lo sarà nel futuro se di esso resterà testimonianza significativa. Se la conservazione nel tempo del Web può costituire un interesse di generale rilevanza culturale, essa investe, quindi, per lo studioso di storia, un obiettivo di primaria importanza».<sup>15</sup>

Non può sfuggire, per esempio, che vi sono molti fenomeni nati sulla Rete, movimenti che si sono sviluppati in ambito nazionale ed internazionale, come quello dell'antiglobalizzazione: Internet e World Wide Web offrono nuovi canali di aggregazione e di iniziativa civile, associazioni e organizzazioni politiche, etc., che non potranno non essere oggetto di studio per la storia dei movimenti, la scienze sociologiche etc. Inoltre Internet è divenuto abbastanza presto anche strumento comunicazione di enti privati e delle pubbliche amministrazioni, se non addirittura di amministrazione stessa, l'e-gouvernement, come è

---

<sup>14</sup> Valacchi (2002), p. 15.

<sup>15</sup> Vitali (2004 b), p. 106.

propugnata nella programmazione dell'IST<sup>16</sup>: uffici ed enti utilizzano i propri siti Web per comunicazioni con l'utenza e le reti Intranet per la documentazione amministrativa, producendo avvisi, circolari, corrispondenza, relazioni, notificazioni e documenti di vario altro tipo "normalmente prodotto durante la propria vita" (...come dice la stessa definizione di archivio), con materiale in digitale, che finisce per integrare, o sostituire, la costituzione dei propri tradizionali archivi cartacei con questa documentazione di formato diverso, che richiederà propri trattamenti per la duplicazione, la preservazione, lo scarto, etc. Quindi la storiografia deve trattare la documentazione che si trova sul Web e documenti che nascono digitali, ma sono ancora tanti i nodi non risolti su questi documenti, le difficoltà che presentano con le loro caratteristiche. I problemi concettuali e le difficoltà non sono ancora oggi risolti in modo soddisfacente e la riflessione sui processi che permettano di tramandare nel tempo siti e pagine Web sono tutti ancora in discussione.

Ma non solo le fonti sono il problema, anche la loro conservazione, in relazione a metodi e luoghi: la conoscenza del nostro passato è resa possibile dall'esistenza di un sistema di istituzioni e di pratiche che si è andato definendo e consolidando nel corso del tempo, in parallelo all'affinarsi delle metodologie di conoscenza e di studio. Biblioteche, musei, archivi hanno finora provveduto, ciascuno nei rispettivi campi d'intervento, alla conservazione e trasmissione al futuro di prodotti editoriali, di opere d'arte e di manufatti in genere. Oggi le tradizionali politiche di conservazione, generalmente in vigore per una realtà che si riferisce ai materiali cartacei, per i diversi meccanismi di produzione – comunicazione dei documenti – distribuzione – acquisizione, non possono essere applicate alle risorse digitali.

I caratteri di volatilità e di dinamicità, d'immaterialità e di fragilità propri delle risorse digitali pongono problemi di ordine metodologico agli storici, suscitando profonde perplessità sulla loro utilizzabilità all'interno dei tradizionali processi di elaborazione storiografica. Resta però il fatto che, probabilmente proprio a causa della volatilità di questa tipologia di produzione, si sente una forte necessità di trovare sistemi per archiviare il Web e per rendere in futuro disponibili informazioni, documenti e oggetti digitali.

Questi sono alcuni contrasti dell'epoca documentaria attuale: sono stati trasferiti su Web molti dei tradizionali strumenti, inventari, etc., favorendo un potenziale rinnovamento della possibilità di ricerca e ampliando la base documentaria disponibile, offrendo molte applicazioni d'informatica e di tecnologie della Rete alla ricerca storica, ma alla moltiplicazione dei contenuti è corrisposta una labilità dell'informazione e della

---

<sup>16</sup> Information Society Technologies.

documentazione, un'eccessiva velocità di pubblicazione che si contrappone alla lentezza della ricerca approfondita.

I documenti digitali non hanno, per loro stessa natura, i requisiti della 'fonte storica': sono immateriali, sono veicolati, perciò bisognosi di media, sono dinamici, per cui manipolabili, mutevoli, temporanei e fragili; la fluidità del digitale è il contrario delle esigenze di ordine giuridico e pratico che tendono alla stabilizzazione. Insomma esattamente non corrispondono alla tradizionale base per una ricerca scientifica per lo storico.<sup>17</sup>

Connesse al problema della labilità della risorsa Web ci sono le esigenze di verifica e di critica delle fonti utilizzate, basate sul principio della fondatezza delle argomentazioni:

«la più accurata e rigorosa delle forme di citazione non può ovviamente sopperire alla non remota possibilità che la risorsa citata, entro un lasso di tempo più o meno breve, non sia più accessibile. Ciò implica, in linea teorica, che l'affermazione corroborata dal documento citato possa risulterne indebolita».<sup>18</sup>

Per la storiografia, la fonte deve essere sempre verificabile, liberamente accessibile, conservata 'per sempre'. Un espediente praticato è quello di apporre la data di consultazione nelle citazioni e descrizioni bibliografiche, elementare ma essenziale almeno per ancorare nel tempo la presenza e l'accessibilità di una determinata fonte.

Il materiale che si trova su Internet è costituito da documenti, di varia tipologia e formato, recuperabili a partire da siti Web e i documenti sono tra le fonti classiche della ricerca. A proposito delle fonti in Internet per lo storico:

«forse sarebbe più opportuno parlare di disponibilità di risorse – accogliendo una terminologia entrata ormai nell'uso corrente per definire i materiali e i servizi, cui si accede attraverso la Rete. Un termine – quello di risorsa – che in italiano non mantiene l'assonanza che in inglese e in francese conserva con l'equivalente in quelle lingue del nostro fonte (source), ma la cui genericità e ambiguità riflettono bene la difficoltà di stabilire una lineare corrispondenza fra gli strumenti di informazione e di mediazione delle conoscenze tipici dei supporti tradizionali e quelli presenti nel mondo digitale»<sup>19</sup>

Per riconoscere le caratteristiche peculiari dei siti Web è interessante anche partire dalle stesse definizioni che ne vengono date presso alcune istituzioni sicuramente rilevanti per il Web stesso.

---

<sup>17</sup> Vitali (2004 b), p. 2-4, 153-6.

<sup>18</sup> Vitali (2004 b), p. 164.

<sup>19</sup> Vitali (2004 a), p. 72. Sulla specificità semantica del termine anche: Andrea Zorzi. *Comunicazione del sapere ed editoria digitale: problemi e prospettive per gli studi medievali*. In: *Medioevo in rete tra ricerca e didattica*. Atti del seminario di studi organizzato dall'Università degli Studi di Parma e dal Dottorato di ricerca in Storia medievale delle Università di Torino, Bologna e Parma (Parma, 24 gennaio 2001). A cura di R. Greci. Bologna: CLUEB, 2002, p. 186. Disponibile on-line, URL: < <http://www.dssg.unifi.it/scriptorium/az/editoria.htm>>.

Indicative sono le definizioni che si possono trovare sul termine sito Web, raccolte da Stefano Vitali:

	<i>Un sito Web è:</i>
Il progetto di analisi della composizione del Web, <i>Web Characterization Project</i> , promosso dall'OCLC <sup>20</sup> , dà una definizione del sito con una tautologia	«sito Web è un insieme di pagine Web, che, nell'insieme costituiscono un oggetto composito di rilevanza informativa»
Il W3C, il consorzio che da regole e standard di funzionamento del World Wide Web, nel suo glossario <sup>21</sup> afferma che un sito Web è:	«una raccolta di pagine Web interconnesse [...] che risiedono nello stesso nodo di rete»
Il Public Records Office di Londra, cioè l'archivio centrale britannico, da una definizione più articolata	«è una collezione di risorse elettroniche: che è resa accessibile all'interno di uno specifico dominio internet, ai fini della comunicazione di informazioni e /o la condotta di transazioni d'affari; e che condividono un comune nome di dominio, che normalmente appartiene a una singola o a un determinato gruppo di organizzazioni [...]; e che mette a disposizione un corpo di risorse informative interconnesse, navigabile usando un browser».

I siti Web, e le singole pagine di cui questi si compongono, sono intrinsecamente instabili e fondamentalmente volatili, basti pensare alla vita media delle pagine Web. Un'altra caratteristica del Web, che ha una non piccola ricaduta sulle possibilità della conservazione, è sicuramente la sua continua e rapida espansione, senza la benché minima previsione di una qualche limitazione a questa crescita. La quantità è preoccupante per la gestibilità, non per lo spazio: il Web ha contribuito al più generale

«...fenomeno di inflazione quantitativa dell'informazione che caratterizza la nostra epoca, ogni ipotesi di conservazione nel tempo dei materiali che lo compongono non può non confrontarsi con il problema dell'elaborazione di strategie di selezione in grado di individuare ciò che vale realmente la pena di tramandare nel tempo».<sup>22</sup>

<sup>20</sup> *Web Sites: Concepts, Issues, and Definitions*, URL: <<http://wcp.oclc.org/pubs/rn1websites.html/>>. Public record office, *Managing Web Resources: Management of Electronic records on Websites and Intranets*, Vers. 1.0, dic. 2001. URL: <<http://www.pro.gov.uk/recodsmanagement/erecords/website-toolkit.pdf>>.

<sup>21</sup> *Web Characterization Terminology&Definitions Sheet*, W3C Working Draft 24- May- 1999. URL: <<http://www.w3.org/1999/05/WCA- terms>>.

<sup>22</sup> Vitali (2004 b), p. 110.

Nonostante le capacità di stoccaggio dei supporti, in continuo accrescimento, la necessità di mettere in atto processi di selezione documentaria è sempre più grande, a causa dell'“esplosione dell'informazione”, o “overload informativo”, che reca, tra l'altro, ridondanza, ripetizioni, il cosiddetto “rumore” di fondo che nasconde ed appiattisce ogni gerarchia di rilevanza. Ogni strategia di conservazione vede la propria base forte proprio nella distruzione; ciò deriva non tanto da ragioni d'economia di spazio e di costi, quanto dal semplice fatto che non può esistere trasmissione della conoscenza e della memoria, che non implichi, al contempo, dosi variabili di oblio: la memoria, e anche quella documentaria, è per sua natura selettiva.

Sul come affrontare le problematiche poste da questo materiale, i modelli conservativi finora adottati scelgono tra due approcci: o l'uso di strutture nate per la ricerca (come i motori di ricerca) o l'adattamento delle discipline bibliografiche tradizionali. Bisogna, in ogni caso, sottolineare che nel Web esistono realmente tipologie di pubblicazione nuove, difficilmente trattabili secondo parametri tradizionali, quindi è difficile ricondurre prodotti della nuova editoria, nel suo insieme e nelle singole componenti, ad un prodotto o a un insieme di prodotti tradizionalmente bibliografici, facilmente classificabili con i classici strumenti delle discipline biblioteconomiche.

Nel Web si accumulano, come già detto, archivi, o parte di archivi, che sono veri sedimenti documentari dell'attività pratica di un soggetto produttore, materiali di e-government e delle reti civiche, ma c'è anche lo speciale contesto dei movimenti che utilizzano il Web e Internet (cui già si è fatto cenno) e le istituzioni archivistiche pensano a fornire strumenti utili. Come il Public Record Office britannico e il National Archives of Australia, che hanno emanato istruzioni e linee guida sulla gestione e conservazione nel tempo dei siti Web delle istituzioni governative e si propongono di rendere gli stessi soggetti, responsabili dei siti, protagonisti in prima persona di strategie organiche per la loro conservazione nel tempo, e, quindi, non basate sui meccanismi automatici di cattura periodica affidati a terzi. Ci si propone dunque, di estendere in modo consapevole anche ai siti Web le usuali politiche di gestione degli archivi, che vedono di solito i loro soggetti produttori farsi carico della loro sedimentazione fino al versamento nelle istituzioni deputate a conservarli, cioè negli archivi storici.

Pierluigi Feliciati, ne *Gli Archivi del Web: perché, cosa e come conservare*,<sup>23</sup> propone una comparabilità fra il ciclo di vita dei siti Web e quello degli archivi e un approccio archivistico alla conservazione del Web. Si potrebbe, prima di tutto, provare ad applicare le

---

<sup>23</sup> Pierluigi Feliciati. *Il Web sugli archivi, gli archivi del Web*. In: Internet per il restauro (2003), p. 5-6.

categorie archivistiche: la conservazione non indiscriminata ma “valutata”, cioè sottoposta a scarto nel passaggio tra archivio di deposito (o da quello corrente...) ad archivio storico, la necessità di conservare e sottolineare il vincolo archivistico (legato all’attività umana all’origine dei documenti), la necessità di predisporre efficaci strumenti di ricerca, l’attenzione ai problemi di privacy e al diritto d’accesso ai dati.

Alcuni siti Web possono addirittura apparire come una componente essenziale dei sistemi documentari dei loro soggetti responsabili, in quanto svolgono una funzione nell’organizzazione di lavoro e di comunicazione, ma tutti i siti Web possiedono una dimensione genericamente archivistica o, meglio, possono presentarsi ed essere gestiti secondo categorie tipicamente archivistiche, proprio perché si sviluppano secondo dinamiche di accumulazione/sostituzione/aggiornamento dei materiali che si realizzano secondo un flusso temporale teoricamente ininterrotto, così come gli archivi che si sedimentano senza soluzione di continuità. Ogni componente, inoltre, acquista significato in relazione a ciò che lo precede e a ciò che lo segue: come suggerisce Stefano Vitali, i materiali dei siti Web assumerebbero maggiormente un ruolo di fonti affidabili e significative se si rendessero in una prospettiva storica, collocati criticamente all’interno del flusso ininterrotto di evoluzione del sito:

«prospettiva che, teoricamente, può essere pienamente conseguita solo ricorrendo alla conservazione delle condizioni che rendono possibile la completa ricostruzione dei siti “at any point of time”, come recita uno dei documenti dei National Archives of Australia».

È la prospettiva storica che può rendere importante una documentazione, quindi l’archiviazione deve prevedere la documentazione del contesto, decisiva per la comprensione dei materiali aggregati nell’archiviazione, che così possono essere utilizzati come fonti storiche; quindi è necessario permettere una ‘critica delle fonti’, di cui ha bisogno il mestiere di storico, come ricorda Giuseppe Recuperati, contro ogni revisionismo, basato sulla memoria che richiede fedeltà, mentre la storia è soprattutto critica:

«la storia come storiografia non è memoria, non è retorica dello stupore, ma è critica responsabile, anamnesi, ricerca della verità, etica».<sup>24</sup>

Una responsabilità non da poco per i conservatori del digitale, perché viene coinvolta, oltre alla conservazione dell’oggetto, anche l’attenzione alla produzione dell’oggetto stesso, affinché sia possibile la ricostruzione critica di come e dove si produce la documentazione.

---

<sup>24</sup> Dall’intervento di Isabella Ricci Messabò e Giuseppe Ricuperati, *Lo storico e il diritto al passato*, in: *Futuro memoria*, Torino (2004).

Si potrebbe riassumere che, per tutelare gli oggetti digitali, così ‘fragili’ ed intrinsecamente evanescenti, si configurano più ruoli e occorrono più responsabilità, anche nella conservazione.

Nell’ambiente digitale ci sono, dice Maria Guercio,

«...protagonisti numerosi e diversificati rispetto all’ambiente tradizionale. In questo giocano un ruolo crescente gli standard e le procedure utilizzate in fase di formazione delle risorse digitali. Questo accresce la responsabilità dei produttori mentre implica l’assunzione di un’attività conservativa precoce. Essenziale diviene identificare tipologie di produttori e nuove forme per la conservazione. Con una particolare attenzione alle novità connesse con lo sviluppo dell’editoria elettronica, nonché con gli impegnativi programmi di e-governement per le Pubbliche Amministrazioni. [...]. È indispensabile definire ruoli, responsabilità, attività specifiche se si vuole dare vita a una struttura funzionante e a una rete di iniziative tese a sostenere il processo conservativo».<sup>25</sup>

Tra i contributi più importanti per la preservazione delle risorse digitali, ci sono sicuramente gli approfondimenti condotti nell’ambito del progetto InterPARES, lavori e risultati, presentati in corso d’opera, per accogliere e scambiare approfondimenti e discussioni, da diversi collaboratori al progetto, tra cui Maria Guercio,<sup>26</sup> e soprattutto da Luciana Duranti, direttrice del progetto. InterPARES è sicuramente uno degli esempi più illuminanti della multidisciplinarietà necessaria nell’affrontare il tema del trattamento conservativo delle risorse digitali: le discipline coinvolte sin dall’inizio nella ricerca erano archivistica, diplomatica, ingegneria informatica e Computer Science, giurisprudenza, ingegneria elettronica, chimica e statistica; nella seconda fase si sono aggiunte molte altre, tra cui: storia della musica, teoria musicale, composizione musicale, performance, teoria del film, produzione, descrizione; teoria della danza e del teatro, varie scienze sociali e naturali.

Oltre ai risultati specifici che interessano il mondo degli archivi, pubblici e privati, a parere di molti, l’esperienza maturata può costituire un’ottima base di riflessione anche per le attività future, sia formative sia organizzative, nel settore della conservazione e dello studio sul digitale. InterPARES (International research on Permanent Authentic Records in Electronic Systems,) è un progetto internazionale che ha come obiettivo la conservazione dell’autenticità dei documenti prodotti, creati e conservati in Data base e in sistemi di gestione documentale per i settori artistico, scientifico e di e-government,<sup>27</sup> che ha coinvolto in tutto il mondo circa quaranta paesi sulle problematiche del digitale, con il fine di costituire una rete di eccellenza per l’informazione e la formazione nel campo specifico della

---

<sup>25</sup> M. Guercio in: *Futuro memoria*, Torino (2004).

<sup>26</sup> Guercio (2002), p. 118.

<sup>27</sup> Fonte per questa parte sono stati i seguenti contributi: Guercio (2002), Duranti (2006 a) e (2006 b).



conservazione digitale. Avviato nel 1999, si conclude nel 2007 una sua seconda fase; sin dall'inizio InterPARES assume la forma di un consorzio di 7 teams di ricerca (nazionali, multinazionali e di settore).<sup>28</sup>

La prima fase aveva lo scopo di produrre teoria e metodi per la conservazione permanente dell'autenticità di documenti prodotti o mantenuti elettronicamente e, sulla base di queste conoscenze, formulare indirizzi di modello, strategie e standard che potessero assicurare tale conservazione. È stato creato un apposito sito Web, ricco di funzionalità e indirizzato all'utenza esterna e al forum dei ricercatori. Nella seconda fase di InterPARES lo scopo diviene assicurare che la parte della memoria che è prodotta in sistemi digitali dinamici, esperienziali e interattivi nel corso di attività artistiche, scientifiche e di e-government, possa essere generata in forma accurata e affidabile, archiviata e conservata in forma autentica sia a breve, sia a lungo termine, per l'uso di coloro che la producono e di tutta la società, nonostante l'obsolescenza della tecnologia e la fragilità dei supporti. InterPARES 1 ha scelto come oggetto di studio documenti testuali nati digitali, prodotti nel corso di attività amministrative e legali in forma di database e document management systems (DMS); questi documenti erano in stato non corrente e sono stati esaminati dal punto di vista di chi li doveva conservare. InterPARES 2 ha scelto come oggetto del suo studio entità dinamiche, colte nelle diverse fasi della loro vita (durante la produzione, mentre erano 'correnti', una volta 'cessate' e, in alcuni casi, obsolete e inaccessibili), il cui contenuto dipende da dati estratti da una varietà di sistemi che possono avere manifestazioni variabili; entità esperienziali, che incorporano il comportamento del sistema in cui sono, con gli effetti d'interazioni soggettive dell'utente, ed entità interattive, in cui ogni intervento dell'utente e ogni input da parte di un altro sistema causano un cambiamento di contenuto e/o di forma.

I prodotti principali delle due fasi di InterPARES sono di natura teorica e metodologica:

1. includono i concetti di documento digitale, di ciclo vitale del documento digitale, di affidabilità, di accuratezza, di autenticità e di autenticazione in ciascuna delle discipline che partecipano alla ricerca;
2. configurano uno schema di metadati archivistici capace di attestare l'identità e l'integrità dei documenti;

---

<sup>28</sup> Per informazioni dettagliate sulla ricerca si rimanda a: Luciana Duranti. *La conservazione a lungo termine dei documenti elettronici autentici: il progetto InterPARES*. "Archivi per la storia". 1-2, 1999. P. 155-167. Luciana Duranti - Terry Eastwood - Heather MacNeil. *Preservation of the integrity of electronic records*. Dordrecht: Kluwer Academic Publishing, 2002. *The Long-term Preservation of Electronic Records: the InterPARES Project*, edited by Luciana Duranti, San Miniato, Archilab, 2005. Accessibile all'URL: <<http://www.InterPARES.org/book/index.cfm>>. Per l'analisi dei materiali prodotti si rimanda al sito Web: <<http://www.InterPARES.org>>.

3. presentano i requisiti di un sistema documentario affidabile in una varietà di ambienti tecnologici;
4. indicano linee guida per artisti, scienziati e professionisti, per la produzione e il mantenimento di documenti digitali che possano essere conservati nel lungo termine (set di principi complementari per i creatori di archivi singoli famiglie o enti medio-piccoli);<sup>29</sup>
5. propongono metodi di selezione e conservazione;
6. costituiscono un modello che rappresenta la catena di conservazione dei documenti digitali dal momento della loro produzione;<sup>30</sup>
7. fondano principi e criteri per lo sviluppo di strategie di gestione dei documenti digitali sia per enti produttori, sia per enti e programmi preposti alla conservazione permanente (set di principi per i conservatori).<sup>31</sup>

Ci sono alcuni dei concetti e principi sviluppati da InterPARES che, nelle intenzioni dei responsabili, possono essere rilevanti oltre il contesto strettamente archivistico in cui sono stati elaborati:

<p>Il concetto di documento digitale (si ricorda che sono stati presi in considerazione, nella seconda fase della ricerca, solo i documenti nati digitali)</p>	<p>Sulla base degli studi di casi condotti, i ricercatori InterPARES hanno concluso che un'entità digitale, per essere considerata un documento, deve avere una forma fissa e un contenuto stabile e inalterabile. Forma fissa significa che il contenuto binario dell'entità digitale è affisso al supporto (hard drive, disco ottico o magnetico, nastro magnetico, etc.), in modo da rimanere completo e inalterato e con la stessa forma documentaria che aveva quando è stato salvato per la prima volta. Considerando i casi dei documenti digitali la cui forma o il cui contenuto possono variare, con una "variabilità limitata", ove la variabilità non è un prodotto della tecnologia informatica in generale, ma è incorporata in specifiche componenti digitali, come i moduli interattivi, i "software applets" che generano le presentazioni variabili, le regole di flusso documentario, o il software che usa l'input dell'utente per determinare l'output successivo; in questi documenti digitali, il requisito di forma fissa riguarda quegli aspetti formali che l'autore intende e può controllare.</p>
<p>Gli attributi di un documento digitale</p>	<p>Le sue caratteristiche: nome dell'autore, la data o la materia, che gli forniscono un'identità unica, possono essere elementi di forma documentaria sul documento o metadati connessi al documento o possono essere impliciti nei suoi vari contesti. Inoltre, nei documenti digitali, le componenti tecnologiche e quelle intellettuali non coincidono necessariamente, e il concetto di componente digitale (o componente tecnologica) accompagna quello di forma documentaria. Una componente digitale è un oggetto digitale che contiene tutto o parte del contenuto di un documento e/o i dati o i</p>

<sup>29</sup> La relazione su queste linee Guida è stata presentata da Philip Eppard, *Making and Maintaining Digital Materials: Guidelines for Individuals*, a InterPARES (2006).

<sup>30</sup> La relazione su queste linee Guida è stata presentata da Kevin L. Glick, *Guidelines for Records Preservation*, a InterPARES (2006).

<sup>31</sup> Questi criteri sono stati oggetto della relazione presentata da Malcolm Todd, *Principles Guiding Policy for Creators and Preservers*, a InterPARES (2006).

	<p>metadati necessari a ordinare, strutturare o manifestare il contenuto: richiede un metodo specifico di conservazione.</p> <p>Il documento digitale non esiste come un'entità fisica e, quando viene immagazzinato, si scinde nelle sue componenti digitali. Quindi, la conservazione di un documento digitale consiste della conservazione delle sue componenti digitali e delle loro connessioni, in modo tale che, ogni volta che il documento venga richiamato, le sue componenti si ricompongano in una riproduzione identica al documento inizialmente immagazzinato.</p>
Documento degno di fede	<p>È un documento affidabile e autentico (in ambiente digitale i due concetti non coincidono), accurato:</p> <p>affidabile = capace di rappresentare fedelmente i fatti di cui tratta; in altre parole, è un documento al cui contenuto si può credere</p> <p>autentico = che è esattamente ciò che dice di essere e non è stato falsificato o corrotto; l'autenticità è a rischio durante la sua trasmissione e la sua conservazione</p> <p>accurato = si riferisce all'esattezza e correttezza del contenuto</p>
Documento autentico	<p>L'autenticità è invece una proprietà del documento e lo accompagna per tutto il tempo che il documento esiste, fornita da:</p> <ul style="list-style-type: none"> <li>- identità certa = costituita dagli attributi che lo caratterizzano e lo distinguono univocamente (autore, scrittore, destinatario, originatore e creatore, le date di produzione, trasmissione e ricezione, la materia o l'atto a cui si riferisce, l'espressione della sua relazione con gli altri documenti, e l'indicazione di allegati;</li> <li>- integrità = il documento è intatto, intero e perfetto (per esempio il numero iniziale di bit strings, può essere compromesso, purché l'articolazione del contenuto e i necessari elementi formali rimangano gli stessi).</li> </ul>
Documento autenticato	<p>L'autenticazione è la firma digitale <sup>32</sup> (che corrisponde al sigillo nell'ambiente documentario tradizionale): un metodo per provare che un documento è quello che dichiara di essere in un momento determinato; è una dichiarazione di autenticità che risulta dall'inserimento o dall'aggiunta di un elemento o di un'affermazione al documento, secondo precise norme legislative.</p>

<sup>32</sup> Sulla firma digitale, aspetti tecnologici e normativi si rimanda a G. Michetti in Guercio (2002), p. 184-202, e anche a Salarelli in Salarelli - Tammaro (2006), p. 43-46. M. Guercio, in Guercio (2006), nella sua riflessione sulla certezza documentaria e la funzione archivistica, affronta le problematiche della gestione e conservazione dei documenti digitali e i cambiamenti nelle modalità di consultazione e fa anche un'analisi critica puntuale del Codice dell'amministrazione digitale in Italia. Mette in luce la contraddizione in merito all'art. 42 sulla dematerializzazione dei documenti, che, a suo avviso viene promossa, salvo poi scoraggiarne l'uso, invitando alla considerazione attenta delle potenzialità e criticità del processo di informatizzazione dei documenti. La firma digitale rientra nel trattamento del documento digitale e dei flussi documentari nella Pubblica Amministrazione ed è tra gli strumenti dell'e-gouvernement (in Italia regolato ancora di recente dal *Codice dell'amministrazione digitale*, D. Leg.vo n. 82, 07-03-2005, pubbl. G. U. 16-05-2005 n. 112, S. O., nato nel contesto seguito alla riforma degli anni '90, nella quale spiccano le *leggi Bassanini*, le Leggi 537/93 e 59/97, i DPR 513/97 e 445/00, la Dir.MIT 9-12-2002, le *Linee guida in materia di digitalizzazione* dal 2001, le delibere AIPA/CNIPA 15/ 1994, 24/1998, 42/2001, 11/2004). A proposito della preservazione dei documenti informatici di cui per legge è prescritta la conservazione, va ricordato che, per gli usi correnti, possono essere archiviati anche con modalità cartacee, ma per la conservazione permanente, questa deve avvenire con modalità digitali (art. 43, c. 3, del Cod. dell'Amm. Dig.), elemento che comporterà, quindi, l'implementazione di sistemi di sicurezza e la programmazione di interventi periodici.

Una conclusione fondamentale raggiunta da InterPARES è quella della necessità di parlare di una “catena ininterrotta di conservazione”, poiché i processi d’immagazzinamento e reperimento dei documenti digitali comportano trasformazioni fisiche e di rappresentazione tali che il concetto tradizionale di conservazione deve includere i procedimenti necessari ad assicurare la trasmissione, inalterata nel tempo, del documento.

La catena continua di conservazione deve:

- ≈ garantire che i documenti siano prodotti in un sistema digitale affidabile
- ≈ dimostrare tutti i cambiamenti subiti dai documenti nel corso della loro vita, oltre a tutte le fasi dei processi di acquisizione, riproduzione e conservazione portati avanti dal custode affidabile.

Infine, InterPARES ha dimostrato che la fonte più importante per stabilire l'autenticità dei documenti conservati dal custode affidabile è la descrizione archivistica, che acquisisce la funzione di attestazione collettiva dell'autenticità dei documenti di un fondo e di tutte le loro relazioni. Tale descrizione offre anche una prospettiva storica sui documenti e sulle loro trasformazioni, e così li autentica attraverso il tempo, piuttosto che solo in un momento determinato, come avverrebbe con l'autenticazione di ciascun singolo documento per mezzo dell'affissione del sigillo.

Molte sono le responsabilità sui documenti (che vengono anche inserite nei metadati relativi al documento stesso) che vengono identificate nella catena ininterrotta della conservazione che InterPARES ha delineato per la preservazione dei documenti autentici:

il produttore, il conservatore, l'autore, il responsabile dell'inserimento nel sistema di tenuta dei documenti del produttore.

Le loro responsabilità ‘seguono’ il corso delle fasi di vita del documento.

Dall'*excursus* finora seguito in ambito archivistico si possono dedurre almeno questi temi:

- ≈ l'utilizzo di tecniche proprie per la descrizione specialistica del materiale,
- ≈ il problema dell'integrità del documento,
- ≈ il progetto di uno scenario di azioni ma anche di responsabilità.

Questi temi, circa i compiti di descrizione specialistica, d'integrità e di responsabilità, si potranno ritrovare anche negli studi biblioteconomici.

Un esponente della corrente che presenta forti dubbi e un profondo scetticismo sulla possibilità e sulla convenienza dell'archiviazione del Web e di una sua ‘catalogazione’ con

strumenti alternativi ai codici classici, come i set di metadati, è Michael Gorman. Uno scetticismo che muove dalla constatazione che l'evanescenza del documento digitale e il susseguirsi incessante, estemporaneo e spesso imprevedibile di pubblicazioni effimere, rendano illogico sottoporre ad un controllo bibliografico o ad una difficoltosa opera di conservazione queste risorse, per le quali, neppure gli stessi autori dimostrano interesse o cura, una volta considerate 'scadute' contenute o vestite.

Efficace è la concezione che esce dalla similitudine che Gorman fa tra i siti Web e le strutture per i cartelloni pubblicitari:

«I siti Web con contenuti continuamente mutevoli, d'altra parte, non hanno alcun parallelo nel mondo della stampa o dell'editoria di altri oggetti tangibili. Assomigliano di più a quelle curiose strutture che si vedono per le vie di Parigi – erette per ricoprirle di manifesti. I manifesti cambiano, si strappano e vengono ricoperti da nuovi manifesti, vengono staccati e deturpati con graffiti, tanto che il contenuto e l'effetto visivo sono diversi di settimana in settimana. Solo il sito resiste. Finché il contenuto è così instabile e la forma volubile e finché i siti stessi, diversamente dalle strutture parigine possono svanire da un giorno all'altro, è molto difficile, se non impossibile conservarli e sottoporli al controllo bibliografico».<sup>33</sup>

Quindi il sito Web è solo un supporto, di per sé molto precario, di forme e comunicazioni transitorie. Il Web come inno dell'effimero.

L'istinto (o la 'nevrosi') del bibliotecario catalogatore, da quando è iniziata l'avventura del Web, è lo sforzo frustrante di realizzare l'idea di catalogare tutto, anche nella Rete, ma i risultati sono scarsi, faticosi, insoddisfacenti, probabilmente perché il Web è il Web e perché la catalogazione è un lavoro faticoso, già solo quando tratta di oggetti tradizionali: con questo atteggiamento anche provocatorio (Gorman parla di "ossimoro" per la volontà di catalogazione del Web anche in uno dei suoi più recenti interventi su questi temi)<sup>34</sup> l'autore cerca di fare il punto sulla situazione e sulle possibilità future del controllo bibliografico su documenti e risorse elettroniche.

Mentre, dopo l'esperienza di un decennio, si dibatte ancora a causa dei dubbi e dei tanti aspetti irrisolti offerti dalle alternative alla catalogazione convenzionale, riuscire a selezionare il materiale valido nel marasma del Web è più che mai un'impresa (ma "the gargantuan task" è l'espressione di Gorman) che obbliga ad affrontare problemi enormi: finanziari, di identificazione di chi si debba assumere la responsabilità di un'operazione costosa di lavoro (molto oneroso per l'applicazione intensiva di principi e di standard), di selezionare (e quindi di scartare, con tutte le implicazioni culturali ed etiche del caso), di scegliere i test di verifica per accertare il grado di conservazione e recupero.

---

<sup>33</sup> Gorman (2003), p. 3.

<sup>34</sup> Gorman (2006).

Resta sempre critico nei confronti di soluzioni alternative alla catalogazione, che finora è l'unica che riesce a fornire standard qualitativamente elevati, comprensione tra i contenuti e le strutture che li contengono, rispetto ad altri standard, come i metadati, surrogati che pretendono di ovviare la difficoltà del lavoro di catalogazione, non danno la qualità della catalogazione. Anzi, se uno standard vuole durare a lungo, deve rifarsi proprio alla catalogazione.

La catalogazione - oggi risultato di una ricerca bicentenaria passata tra varie difficoltà, errori, frammentazioni, difformità, problematico controllo dei vocabolari - ha raggiunto, secondo Gorman, risultati che le permettono di essere perfettamente adottata per le risorse elettroniche ed attualmente rappresenta l'unica soluzione per indicizzare alla perfezione poche risorse digitali, degne di essere catalogate e conservate dalle biblioteche, secondo regole motivatamente molto complesse, ma all'altezza del raggiungimento di un risultato qualitativamente elevato.<sup>35</sup>

Al contrario, i metadati si sono moltiplicati per quantità e complessità d'applicazione, si sono rivelati una vera delusione rispetto alle alte attese di semplificazione e velocizzazione del lavoro, per non parlare della qualità dei risultati. Se si prendono in considerazione i motori di ricerca, che usano le metodiche dei metadati (come Google per esempio), questa ricerca non è il processo autentico che risponde con attendibilità, sicurezza e funzionalità alle esigenze della ricerca stessa: pur con il pregio della velocità, tale meccanismo non riporta risultati affidabili, con garanzia di coerenza, ma una massa immensa di dati, spesso neanche tutti vagliabili e in un ordine non utile al ricercatore. Non costituiscono quindi la possibile magia "third way" che apporta taglio dei costi, velocità, precisione e competenza nel recupero.

Anche Riccardo Ridi, per affrontare il problema della conservazione di quello che chiama "retaggio digitale", prospetta una riflessione sull'oggetto digitale e una riscoperta delle più tipiche funzioni degli specialisti: s'intende crescere la consapevolezza che attraverso strumenti tradizionali, come la propria missione conservativa, il deposito legale e l'inserimento nelle bibliografie nazionali, si possono consentire la trasmissione alle generazioni future di una dimensione storica del Web e tentare un controllo bibliografico di parte dei documenti digitali prodotti, attestando nel tempo le varianti dei documenti dinamici instabili e in continuo aggiornamento.

Ribadendo il ruolo delle istituzioni preposte alla conservazione e trasmissione ai posteri del patrimonio culturale, come i musei conservavano oggetti originali o in copia unica, gli

---

<sup>35</sup> Gorman (2001), (2003) e (2006).

archivi documenti originali, le biblioteche esemplari come concretizzazioni di pubblicazioni, di cui diventano “garanti”, ora invece, le biblioteche devono fronteggiare oggetti digitali, con caratteristiche e problematiche ben diverse dal consueto:

«Le biblioteche, finchè si occupano di supporti elettronici portatili restano, tutto sommato, nel solco della tradizione... Il vero salto concettuale appare solo col network publishing, con l'editoria elettronica in rete, con Internet e la conseguente radicale riorganizzazione della classica catena documentaria sempre “locale” in una vera e propria rete documentaria ipertestuale in gran parte “remota”».<sup>36</sup>

La differenza con le “pubblicazioni” precedenti la rivoluzione informatica è significativa per comprendere le difficoltà di trattamento: un esemplare manoscritto viene reso pubblico attraverso la diffusione di copie tutte diverse, una pubblicazione a stampa attraverso copie uguale ad un certo numero di altre copie, ma nel Web essere pubblicato significa giacere su un server telematicamente collegato a tutti i computer del mondo, essere messo a disposizione e migrare rapidamente di server in server, essere distribuito in “file”, prigioniero di hardware e software. Ridi riporta la citazione da J. Rothenberg, che anche qui volentieri si ripete:

«Di per sè un file non è un documento, è soltanto la *descrizione* di un documento, che acquista realtà solo quando il file viene interpretato dal programma che l'ha prodotto. Senza questo programma (o un software equivalente), il documento rimane prigioniero della propria codifica».<sup>37</sup>

C'è quindi una dipendenza tecnologica per la conoscenza del documento, ma le tecniche per salvare questo nuovo patrimonio vanno trovate, raffinate, indirizzate, perché il compito di conservare almeno parte della produzione su Web deve essere compiuto:

«Se nessuno si occuperà di “archiviare Internet”, ovvero di preservare e catalogare almeno le principali varianti dei documenti disponibili in rete che si sono susseguite nel corso del tempo, il World Wide Web sarà per sempre condannato a vivere in un eterno presente di documenti aggiornati, privi di qualsiasi dimensione storica. Per non parlare, anche a prescindere da "edizioni" e "varianti" di documenti comunque ancora disponibili, di quelle pagine web che scompaiono repentinamente nel nulla dopo un periodo più o meno lungo di permanenza in rete, senza lasciare alcuna traccia».

Fra l'illusione di una ‘museificazione’ automatica e globale dei progetti come Internet Archive e “il fatalismo” di confidare in una ‘selezione naturale’ per cui i documenti

---

<sup>36</sup> Ridi (2003), p. 2. Ridi qui cita anche il concetto di “docuverso” di Nelson: Theodor Holm Nelson. *Literary machines 90.1: il progetto Xanadu*. Traduzione di Valeria Scaravelli e Walter Vannini, revisione di Giancarlo Mauri. Padova: Muzzio, 1992. Ed. originale: *Literary machines 90.1*. Swarthmore, T. H. Nelson, 1990. Riccardo Ridi. *Xanadu: l'ipertesto globale fra utopia e realtà*. In: *Università: quale biblioteca?* Atti del seminario-dibattito a cura di Rodolfo Taiani. Trento: Università degli studi di Trento, 1995, p. 153-161. Anche in: AIB-WEB Contributi, febbraio 2001, <<http://www.aib.it/aib/contr/ridi1.htm>>.

<sup>37</sup> Rothenberg (1995), p. 18.

realmente importanti sopravviveranno comunque, grazie ad aggiornamenti, riedizioni, citazioni, copie locali e trasferimenti, la anche proposta di Ridi è quella di conservare una selezione di risorse e recuperarle sotto il controllo bibliografico:

«[...] almeno i documenti elettronici disponibili in Internet (non solo pagine web ma anche archivi di mailing list e newsgroup, gopher ecc.) che si considerano più stabili, compiuti, identificabili e descrivibili, proprio come in quasi tutti i paesi del mondo si cerca di ottenere in ambiti più tradizionali il controllo bibliografico universale e la disponibilità universale delle pubblicazione mediante le due armi delle bibliografie nazionali e del deposito legale, che coprono una vasta percentuale, ma mai la totalità, dei documenti prodotti. [...] In ambiente reticolare [...] torna d'attualità il controllo bibliografico universale almeno delle principali versioni dei documenti che incessantemente vengono aggiornati in rete».

In un quadro di cooperazione e condivisione delle risorse e limitazione degli sforzi duplicati e inutili, ci sarebbe poi la risoluzione del problema bibliografico e conservativo, oltre che gestionale, delle collezioni digitali, aspetto sul quale molti approfondimenti hanno offerto Tommaso Giordano e Anna Maria Tammaro (v. anche nel Cap. 1, Introduzione).

Per giungere ad una conclusione su quale sia l'aspetto che va privilegiato nella preservazione, anche per elaborare possibili strategie 'politiche' di conservazione, si ritorna ora ancora sulla riflessione circa il documento digitale. Non è un esercizio teorico la ricerca d'approfondimenti, ma è la ricerca di valori e contenuti che si vogliono salvare e tramandare rispetto al documento prodotto con nuove tecnologie, ed è alla base della coscienza necessaria per motivare scelte politiche sulla conservazione, decisioni per affrontare i problemi conservativi, soluzioni gestionali rispetto alla quotidianità della gestione delle risorse digitali e alla determinazione su cosa conservare tra le risorse esistenti; ma anche per scegliere come descrivere un documento, quali aspetti del documento devono essere documentati per stabilire la sua autenticità, quali metadati vanno elaborati per accompagnare gli oggetti digitali nel tempo.

Se pare ovvio che, dato un fine, si definiscano i limiti per raggiungerlo e si trovino gli strumenti, nell'approccio concreto al problema della conservazione digitale l'applicazione di questo metodo di "ovvietà" non è stata così attuabile, a causa delle complessità proprie sia del materiale elettronico e multimediale, sia della definizione del contenuto con valore documentale delle risorse; nonostante che in questi ultimi dieci anni la ricerca abbia comunque fornito alcune soluzioni tecnologiche adeguate, è stato difficile avviare vere politiche di conservazione e diffondere metodologie di gestione tempestiva del problema conservativo, con consapevolezza e capacità di selezionare sistemi e criteri di raccolta di risorse più congrui allo scopo che ci si prefigge. Ciò anche perché non esiste un accordo universale su cosa debba essere preservato e per quale scopo. La metodologia scelta per la



preservazione è influenzata dalla differente opinione circa il “quando” e il “quale parte” dell’oggetto digitale debbano essere interessati dalle azioni di preservazione, e ciò è conseguenza di una diversità di concezione dell’oggetto digitale, o meglio, di concezione dell’integrità e dell’autenticità dell’oggetto digitale.

«Le argomentazioni sulla metodologia, emulazione vs. migrazione, ad esempio, spesso mascherano una profonda differenza di opinione su cosa costituisce l’autenticità degli oggetti digitali».<sup>38</sup>

E, al proposito, sempre Tammaro riporta la contrapposizione tra questi due esempi alla base del processo d’identificazione dell’oggetto della preservazione: la valutazione che un oggetto digitale sia un programma eseguibile che deve conservare tutte le sue funzionalità originarie nel tempo per poter sembrare autentico, il che implica che bisogna conservare il supporto di memorizzazione ed il suo contesto, oppure che, per fornire il contesto necessario per rendere l’oggetto accessibile in futuro, sia necessario catturare il documento digitale al momento della sua creazione, e, quindi, bisogna conservare la funzionalità dell’accesso alla risorsa come è in origine.

Come già in ambiente archivistico, viene anche qui identificato il fulcro della preservazione nella salvaguardia dell’integrità del documento: nell’interrogarsi su quale sia il vero oggetto della preservazione, si arriva a chiedersi ciò che identifica nella sua interezza e nella sua unicità un documento elettronico, e, quindi, si giunge alla definizione delle caratteristiche che determinano la sua ‘integrità’.

Nel primo documento di ricerca, il *Report of the Task Force on Archiving of Digital Information*, della Commission on Preservation and Access – RLG (Research Libraries Group), del 1996, si identificano cinque caratteristiche (content, fixity, provenance, reference, context) contro i pericoli di alterazione del documento; si auspica la realizzazione di una rete di sicurezza in grado di garantire l’integrità nel corso del suo ciclo di vita e di assicurarne l’accesso nel tempo attraverso la realizzazione di un sistema di archivi digitali secondo una struttura distribuita, piuttosto che centralizzata, nelle varie forme della corporazione, della federazione e del consorzio. I creatori/editori/proprietari dell’informazione vengono poi identificati come responsabili attivi della conservazione, o ingaggiando altre parti (archivi digitali certificati), o interagendo e collaborando con le biblioteche durante e dopo il ciclo di vita attivo dell’informazione.

---

<sup>38</sup> Tammaro – Salarelli (2006), p. 201

Non distanti da queste caratteristiche, sono state identificate ancora da A. M. Tammaro le prerogative che garantiscono l'integrità dell'informazione in un contesto digitale, che qui si ripropongono:<sup>39</sup>

<i>Contenuto:</i>	Il contenuto comune a tutti i documenti digitali è un insieme di byte, che sono dipendenti dall'ambiente tecnologico di lettura. Metodi come la migrazione preservano il contenuto intellettuale del documento ma non necessariamente la forma di presentazione. La preservazione del contenuto infatti è indipendente dal formato di memorizzazione. Sono stati studiati alcuni meccanismi speciali che garantiscono l'autenticità del contenuto, come firme crittografate o marcature (watermark).
<i>Stabilità:</i>	Gli oggetti digitali devono avere un contenuto stabile per essere preservati. La stabilità deve essere inoltre indipendente dall'HW e dal SW. Per alcuni oggetti digitali soggetti a continuo cambiamento, la stabilità deve essere legata alla versione di un dato momento. L'integrità degli oggetti digitali è misurata dalla somma cifrata (encrypted checksum), che è calcolata sommando tutti i bit di un documento digitale (i bit sono interpretati come numeri per questo processo).
<i>Consistenza della citazione:</i>	Ogni oggetto digitale deve essere accompagnato da metadata che descrivono il suo contenuto e la sua natura. I metadata possono essere integrati nel documento stesso o archiviati in archivi separati. Non c'è ancora un accordo universale su standard per la preservazione. La rete di collegamenti creata dalle citazioni dovrà poter recuperare l'edizione citata del documento e non, come avviene spesso la sola edizione più recente.
<i>Contesto:</i>	I metadata devono anche indicare il contesto tecnologico da cui l'oggetto digitale dipende per la sua funzionalità ed i collegamenti ad altri oggetti digitali, oltre al contesto sociale in cui la risorsa è stata creata ed usata, cioè ogni individuo coinvolto nella pubblicazione, distribuzione o creazione della risorsa nel suo ciclo di vita.

---

<sup>39</sup> Tammaro – Salarelli (2006). Nel prospetto qui presentato il testo è stato attinto largamente dal capitolo di A. M. Tammaro, p. 202-203

## 1.2. Strategie, metodi, criteri e sistemi di preservazione.

### 1.2.1. Strategie di una “politica di conservazione”, reti e modelli di deposito.

Il problema della conservazione della produzione intellettuale digitale è una materia dai molteplici aspetti, teorici, politici e tecnici. È quindi una questione molto più complessa del, pur già complicato, livello tecnico.

«Le biblioteche digitali si fanno carico, tramite la cooperazione, della conservazione permanente dell'eredità culturale digitale. Alcuni repository, in base alle vigenti normative sul deposito legale e tramite la cooperazione e l'integrazione delle proprie funzioni, sono deputati alla conservazione permanente (ovvero allo scarto necessario) dell'eredità culturale digitale, e a questo fine adottano gli standard più opportuni (es. Open Archival Information System, OAIS ISO 14721). Allo stesso modo gli istituti depositari adottano le migliori tecnologie per la conservazione dei dati e per la loro consultabilità a lungo termine».

Questo recita la tesi n. 26, nella sezione delle Funzioni, del *Manifesto per le biblioteche digitali*.<sup>40</sup> Si può ormai considerare emersa una certa consapevolezza culturale del problema della conservazione digitale, ma a questo avviso, nella maggioranza dei casi, non fa riscontro un comportamento conseguente sul piano politico e pratico e, pur riconoscendo il valore strategico della conservazione, molte biblioteche ritengono che questo sia un compito da demandare ad altri (biblioteche nazionali, trusted third part repositories, consorzi, editori). Una comprensione del problema, quindi, ancora non capace di adeguarsi alle esigenze di elasticità e prontezza:

«Tutto il gioco si svolge all'interno di una visione di corto termine, dove le statistiche di uso e uno sconto anche modesto contano più di qualsiasi considerazione di ordine culturale e politico».<sup>41</sup>

Per la preservazione dell'informazione digitale, ma soprattutto per assicurare che l'autenticità di una risorsa sopravviva nel tempo, nonostante i cambiamenti tecnologici, con la perdita minima possibile rispetto al suo contenuto informativo, alla sua funzionalità ed accessibilità, è indispensabile affrontare un insieme di decisioni strategiche e procedurali, insieme ad altri aspetti più tecnici che riguardano la gestione dei dati, tecniche e standard. Sono, infatti, possibili scelte metodologiche molto diverse tra di loro, ma anche decisioni che comprendono una combinazione di metodi e strategie.

---

<sup>40</sup> *Manifesto per le biblioteche digitali*. ©AIB 2005, agg. 2005-12-20.  
<<http://www.aib.it/aib/cg/gbdigd05a.htm3>>.

<sup>41</sup> Giordano (2006)

Anna Maria Tammaro<sup>42</sup> dice che, al di là dei problemi tecnici, c'è tutta una serie di problematiche politiche, a causa del fatto che non c'è una profonda consapevolezza ed un'attenzione generalizzata al problema della preservazione dell'informazione digitale. Queste “problematiche politiche” cui si riferisce

«...hanno a che fare con una chiara strategia per la preservazione e con una necessaria collaborazione tra tutti gli interessati, incluso una diversa cornice legislativa per i diritti di proprietà intellettuale».

Quindi, per un'ottimale pianificazione della preservazione, diventa necessaria anche una strategia di comunicazione e di coinvolgimento di tutte le parti, anche dei produttori delle risorse.

Documenti “politici” possono essere considerati anche la Risoluzione del Consiglio d'Europa<sup>43</sup> *Conservazione della memoria del domani - conservazione dei contenuti digitali per le generazioni future*, 25 giugno 2002, che ribadirà la necessità di monitorare le iniziative di conservazione realizzate in Europa e le cooperazioni risultate efficaci per scambi di esperienze su politiche, programmi e problemi, di sostenere le istituzioni deputate alla conservazione, con infrastrutture organizzative (Reti) e norme tecniche necessarie a creare e mantenere nel tempo le raccolte digitali, ma anche la *Carta per la conservazione del patrimonio digitale* dell'UNESCO (17 ottobre 2003), nella quale si affermava esplicitamente il rischio gravissimo di perdita del patrimonio, riconoscendo attraverso una serie di principi tutti quegli elementi che hanno uno specifico valore di documento e di testimonianza, ai quali va indirizzata salvaguardia e conservazione, mirando a creare una pluralità di sistemi accessibili nel lungo termine e muovendo a tal fine gli Stati membri, organizzazioni intergovernative, associazioni internazionali, società civile e settore privato, quindi con misure tecniche, ma anche politiche.<sup>44</sup>

Nello specifico della programmazione per affrontare la preservazione digitale, Michael Day<sup>45</sup> elencava, nel 2001, le aree in cui occorre un'attenta considerazione; tra queste sono

---

<sup>42</sup> Salarelli – Tammaro (2006), p. 198.

<sup>43</sup> Dal sito Web di Europe, Releases Press (URL:< <http://europa.eu/rapid/pressReleases/>>): «Il Consiglio ha adottato la risoluzione sulla conservazione della memoria del domani - conservazione dei contenuti digitali per le generazioni future (*doc. 9206/02*). Tale risoluzione sottolinea l'esigenza di continuare a definire metodi e linee Guida per la conservazione a lungo termine di quelle registrazioni, documenti, collezioni e archivi che sono indispensabili per la salvaguardia del patrimonio europeo. Il Consiglio è quindi convinto che è indispensabile darsi come obiettivo l'attuazione di interventi concreti, condivisi da tutti gli Stati membri, che superino la grande frammentazione metodologica fra i diversi campi della cultura tenendo conto non soltanto delle sfide tecnologiche, ma anche delle più vaste implicazioni socioeconomiche». Sul dibattito v. anche Santangelo (2002).

<sup>44</sup> Altre note e riferimenti su questo importante documento si trovano nell'introduzione al Cap. 3.

<sup>45</sup> Michael Day. *Preservation 2000*. “Ariadne”, 26 (Jan. 2001). E anche: Id. *Collecting and preserving the World Wide Web. A feasibility study undertaken for the JISC and Wellcome Trust. - Version 1.0 (25 February*

segnalate: la necessità di strategie per la preservazione da mettere in atto al momento stesso della creazione della risorsa digitale, l'opportunità di accordarsi su specifiche licenze o assicurazioni per la preservazione al momento della negoziazione con produttori e fornitori di risorse digitali, l'organizzazione di uno staff di operatori sempre aggiornati, consapevoli e con capacità specifiche.

Non si può non notare (e non solo negli interventi fin qui riportati) un ribadire di alcuni aspetti che riguardano le scelte politiche di chi ha le responsabilità della funzione conservativa, che rimandano ad una visione di cooperazione, politica e scientifica: per l'archiviazione del Web occorrono studi, progetti, confronti e sviluppare una cooperazione nazionale ed internazionale; occorre collaborazione con produttori ed editori di informazione, rendendo esplicita la vitale importanza della comunicazione tra chi crea la risorsa e chi la preserva (e per questo il modello OAIS si pone come modello di riferimento per la corretta archiviazione della risorsa digitale).

Ma vi sono anche aspetti più inerenti alla normale, quotidiana, mansione di responsabilità propria dell'amministratore-gestore-conservatore, come quello della determinazione, fondamentale, che è costituito dalla "politica delle collezioni": ossia, rispetto alle analisi dei patrimoni (il posseduto - il posseduto nei contesti di riferimento, territoriale o per altre ragioni - l'acquisibile), quella funzione decisionale che fornisca uno studio di settore, dei piani di fattibilità, delle strategie anche di medio e lungo termine, con il progetto di una "carta" delle collezioni anche digitali, che trovi l'equilibrio ottimale tra il realizzabile e il desiderabile, definendo chiaramente criteri - tempi - risorse per una cernita su cosa va conservato e per quanto tempo (qui è importante l'eventuale termine o anche la sospensione temporanea di interventi di preservazione digitale, in quanto questi coinciderebbero, ovviamente, con una reale dismissione della preservazione del materiale stesso).

La questione della competenza e della responsabilità decisionale è molto delicata, anche perché investe l'aspetto, meno praticato in alcune realtà come quella italiana, dell'apertura alla collaborazione concreta, alla cooperazione sui compiti effettivi e non solo teorica, il capovolgimento di alcune consuetudini gestionali: però la conservazione a lungo termine dei materiali digitali ha proprio bisogno di nuovi approcci - difficili da conciliare con il modello tradizionale di biblioteca - e la questione della conservazione è sicuramente una delle maggiori sfide che le biblioteche devono affrontare oggi, con esito determinante, non solo per la loro missione, ma bensì per il futuro dell'informazione, dell'educazione e della ricerca.

---

2003). <[http://www.jisc.ac.uk/uploaded\\_documents/archiving\\_feasibility.pdf](http://www.jisc.ac.uk/uploaded_documents/archiving_feasibility.pdf)>. (Citaz. desunte da Tammaro - Salarelli (2006)).

A causa della sua peculiare caratteristica di ricerca continua e di applicazioni, anche tecnologiche, ripetute e periodiche, la conservazione del materiale digitale deve essere considerata un processo continuo, come già accade per le altre attività che implicano la gestione di risorse elettroniche. Dunque, la preservazione digitale va considerata un processo di lavoro e, come tutti i processi di lavoro, va esaminata nel contesto, progettata e pianificata, anche nei minimi particolari.

Il primo fondamentale aspetto per affrontare l'archiviazione digitale è la consapevolezza che esiste una differenza metodologica considerevole rispetto alla conservazione del materiale cartaceo. Per esempio:

«in sostanza, con i documenti cartacei ci si può permettere una certa discontinuità nella gestione [...] che invece è del tutto improponibile quando si tratta di documenti digitali, dove il procrastinare può essere esiziale alla loro sopravvivenza [...]».<sup>46</sup>

Inoltre, può essere utile valutare chiaramente alcuni 'limiti' della questione:

- ≈ se l'istituzione che vuole applicarsi alla preservazione vede nettamente previsti nei propri compiti istituzionali l'attività conservativa e, se sì, rispetto a cosa;
- ≈ l'obiettivo della conservazione è ottenere la longevità a lungo termine per due scopi: permettere la conoscenza di come era il prodotto digitale quando è nato, autentico ed integro, e avere la disponibilità della sua fruizione, visto che la preservazione richiede l'accesso;<sup>47</sup>
- ≈ la preservazione ha costi molto elevati, pertanto ci si trova a dover identificare con precisione l'oggetto<sup>48</sup> della conservazione, decidere cosa acquisire e cosa preservare; si potrebbero scartare alcune risorse, ad esempio quelle che vengono già preservate a livello nazionale, per il deposito legale;
- ≈ i "tempi" sono pressanti e molto impegnativi per chi gestisce questo materiale, in quanto il problema della preservazione non è alla fine del ciclo del documento, come succede per il materiale tradizionale, e, inoltre, sarà necessario un atteggiamento di cura dell'usabilità e dello stato di conservazione continuativo nel tempo e, comunque, sempre... 'molto prima di quanto si potrebbe pensare';
- ≈ tra gli oggetti digitali che si dovranno trattare sono comprese le risorse elettroniche locali e quelle remote, nate digitali e convertite da formato analogico:

---

<sup>46</sup> Giordano (2006), p. 97.

<sup>47</sup> Seamus Ross nel suo intervento in *Cultural heritage*, Firenze (2006), afferma: «Preservation requires access».

<sup>48</sup> Faiella (2000) presentava in modo molto stimolante lo scenario della preservazione, identificando delle Strategie di metodo (how to), l'Identificazione dell'oggetto (what), le Strategie di responsabilità (who), e qui, per la sua pregnanza, si riprende in parte questo approccio.

ogni tipologia presenterà esigenze tecniche differenti, ma per le risorse born digital opportune strategie per la preservazione vanno previste al momento stesso della creazione della risorsa digitale, visto che è dal suo nascere che la risorsa può avere già in sé caratteri che le garantiscano una certa longevità d'uso;

≈ occorre, una volta deciso di intraprendere l'attività di preservazione, disegnare policy)<sup>49</sup> e strategy<sup>50</sup> da seguire, identificare delle responsabilità.

Queste possono essere identificate come le istanze primarie di una politica di conservazione.<sup>51</sup>

1. La preservazione deve essere identificata come una funzione specifica a sé stante.
2. La politica di preservazione deve essere dichiarata esplicitamente, collegata anche alla politica delle licenze.
3. Poiché il documento elettronico è instabile è inadatto alla consuetudine del metodo dell'«autoconservazione», o ad una conservazione «per inerzia», metodi molto applicati da sempre nelle biblioteche e nei luoghi di conservazione.
4. L'instabilità e la caducità della risorsa richiedono di strutturare in anticipo un piano d'azione, come un «progetto di preservazione», con fasi di conservazione bene definite e non procrastinabili né «diluite» accidentalmente nel tempo, non avendo la garanzia, fino ad oggi scontata, della coincidenza fra attuale accesso e la conservazione; il piano deve essere sostenibile nel lungo periodo, per quanto riguarda le risorse, finanziarie, umane, logistiche etc.
5. La risorsa da conservare va identificata già nella fase d'acquisizione, mentre è ancora basata sulla licenza stabilita di uso temporaneo, anche in considerazione delle possibilità di accordo per l'ottenimento della copia di archivio.
6. Nella stipula del contratto tra le parti, le licenze elettroniche devono prevedere almeno questi elementi, rilevanti ai fini della conservazione digitale: per esempio, per i giornali elettronici, l'annata corrente- le annate precedenti – la copia sottoscritta fornita

---

<sup>49</sup> «Una linea d'azione (policy) è la determinazione formale, motivata da interessi o programmi specifici, della direzione o dell'orientamento con cui un'organizzazione dovrà esercitare il suo mandato, le sue funzioni e attività». In: Maria Guercio, *Prefazione*. In I.C.C.U. – Università degli Studi di Urbino – ERPANET.

*Normative e linee d'azione per la conservazione delle memorie digitali. Un'indagine conoscitiva*. Firenze: 2001

<sup>50</sup> «Una strategia (strategy) è il complesso dei mezzi pratici definiti da un soggetto per raggiungere un determinato scopo, ovvero un piano o una mappa per lo sviluppo di linee d'azione». In: Maria Guercio, *Prefazione*. In I.C.C.U. – Università degli Studi di Urbino – ERPANET. *Normative e linee d'azione per la conservazione delle memorie digitali. Un'indagine conoscitiva*, Firenze: 2001.

<sup>51</sup> Tommaso Giordano fa un'analisi molto precisa della differenza degli approcci alla conservazione, tradizionale ed elettronico, da cui sono stati desunti i termini riportati negli incisi e hanno tratto spunto alcune considerazioni qui liberamente riportate nell'elenco; v. Giordano (2007).

dall'editore in un determinato formato, che possa essere installata sul sito locale come copia di archivio, cioè il «diritto di archiviare».

7. La biblioteca, o l'istituzione in causa, deve prevedere un'infrastruttura specifica per le risorse conservate in grado di permettere un sistema di archiviazione e poi di ricerca, non essendo coincidente la preservazione con l'accesso.
8. Per l'instabilità e il carattere dinamico della pubblicazione digitale, nonché per l'obsolescenza tecnologica cui è particolarmente esposta, il documento digitale richiede una cura attiva e costante, che non ammette fratture e discontinuità, e non la semplice custodia, ma, nella gestione del ciclo di vita del documento digitale occorre una *digital curation*, che prevede un'attenzione per la risorsa non solo sulla sua forma (o formato), ma anche sui contenuti informativi di cui è portatrice.

In questo impegno sono contenuti due termini, maturati in ambito anglosassone, molto significativi delle nuove attività:

- ≈ life cycle, nell'ambito della gestione delle collezioni, indica l'insieme delle attività mediante le quali le biblioteche esercitano la loro responsabilità di conservazione dei documenti nel lungo termine;
- ≈ digital curation indica diversi impegni: manutenzione, conservazione del documento digitale lungo il suo ciclo di vita, capacità di dare valore aggiunto alle collezioni, stabilendo collegamenti disciplinari, evidenziando i contenuti e contestualizzandoli.<sup>52</sup>

Per realizzare politiche di conservazione e studiare sistemi in grado di soddisfare la preservazione nel tempo di risorse digitali, per le quali, per ordini di grandezza e di impegno, non è considerata possibile una politica di autonomia della singola istituzione, è, come si visto, auspicata e ricercata una politica di cooperazione. In questo senso si sono mosse le progettazioni di reti, considerandole lo strumento per affrontare al meglio problematiche complesse come quelle della preservazione, e si sono organizzati diversi 'modelli', depositi e sistemi di archiviazione digitale per la conservazione.

L'uso della rete<sup>53</sup> come ambiente di collaborazione e cooperazione indirizzate alla ricerca ma anche all'applicazione pratica di servizi e infrastrutture, permette di raggiungere più velocemente sviluppi della ricerca, distribuendo lo sforzo ed evitandone l'inutile duplicazione, creando soprattutto delle comunità di professionisti, ricercatori e operatori,

---

<sup>52</sup> Si veda anche DIGITAL CURATION CENTRE (DCC), <<http://www.ac.uk/about/what/>>, di cui al cap. 2.1.3.

<sup>53</sup> Seamus Ross in: Conservare il digitale (2006). Ma anche Guercio (2004), p. 13.



accrescendo responsabilità, conoscenza, orientamento e formazione. Con la presenza di partecipanti e di contatti diversi si crea una situazione di controllo reciproco e di verifica continua dei risultati, che può portare a proficui aggiustamenti e reindirizzamenti delle policies avviate.

Un esempio di cambio di strategia nella politica della preservazione è, secondo Seamus Ross, la “Roadmap di Ricerca per la Preservazione”: in seguito alla constatazione che anni di discussione non hanno portato ad una strategia ben definita e concisa a livello europeo, si definisce una Roadmap di Ricerca, guidata dalla DPE e basata su ‘NSF/DELOS Workgroup Report – Invest to Save’, con la previsione di identificare dei centri di competenza (gli European Competence Centres).

Ross descrive le reti come “ambienti collaborativi” in grado di:

- ≈ facilitare gli sviluppi di ricerca non altrimenti realizzabili,
- ≈ creare comunità dei soggetti operanti,
- ≈ promuovere la crescita della consapevolezza ed una informazione distribuita,
- ≈ distribuzione di carichi e impegni tra i partners,
- ≈ eliminazione della ridondanza degli sforzi.

Una distinzione ampia si può fare tra le reti che offrono servizi tecnici e le reti di collaborazione/cooperazione, mentre, rispetto all’organizzazione, si possono distinguere reti che rispondono a progetti di modelli centralizzati e reti che si rifanno a modelli distribuiti.

Tra alcuni esempi di reti esistenti nel panorama internazionale, si potrebbe così schematizzare:

Reti con servizi tecnici		LOCKSS (Lots of Copies Keep Stuff Safe), <sup>54</sup> Grid-based storage e servizi di preservazione, Federazioni di depositi
Reti di collaborazione/cooperazione	Connotazioni degli obiettivi: consapevolezza e orientamento Formazione Ricerca	InterPARES PADI ERPANET Digital Curation Centre (DCC) DigitalPreservationEurope (DPE) DELOS Network of Excellence and DPC Digital Preservation Cluster. <sup>55</sup>

<sup>54</sup> Giordano (2006), Ridi (2004). Nel 2007 il gruppo per il sistema di controllo formato da OCLC e da grandi biblioteche, riunisce anche grandi nomi storici d’istituzioni e di editori (ad es.: New York Public Library, le Università di Stanford, Indiana, Virginia, Edinburgh, l’American Chemical Society, American Physiological Soc., American Medical Association, Blackwell, Elsevier, Springer, SAGE, J. Wiley & Sons, Taylor and Francis, Oxford University Press, etc.). URL: <<http://lockss.stanford.edu>>.

<sup>55</sup> Tranne InterPARES, di cui si è già detto, di questi progetti viene dato un approfondimento nel cap. 2.

Reti che danno vita a modelli di tipo centralizzato	Tra i modelli centralizzati il più famoso è Jstor, <sup>56</sup> l'organizzazione che si preoccupa della conservazione a lungo termine del retrospettivo di periodici cartacei in diversi campi dello scibile, per lo più in lingua inglese, dei file digitali e anche degli originali cartacei raccolti per l'operazione di riproduzione
Reti che danno vita a modelli di tipo distribuito	<p>Tra quelli distribuiti viene citato come ottimo esempio LOCKSS, come già sopra citato, che offre servizi tecnici. Dichiaro come scopo quello di "rivitalizzare il ruolo centrale delle biblioteche offrendo un approccio cooperativo che permette a ciascuna biblioteca di salvaguardare la propria autonomia e di mantenere le proprie collezioni digitali a costi accettabili con un sistema di archiviazione facile da usare". Ciò attraverso: software distribuito gratuitamente e architettura distribuita con annessa procedura di controllo e di salvaguardia, che si basa sulla reciproca verifica dell'integrità degli archivi da parte dei partner; oltre ad essere assicurata l'integrità dei documenti posseduti in comune dai singoli partecipanti, consente anche di eliminare le procedure dispendiose di back-up locale.</p> <p>Si può citare anche il programma DOM (Digital Object Management Programme), sviluppato dalla British Library per integrare le collezioni in una strategia di conservazione e accesso, con la riduzione al minimo di perdita degli oggetti archiviati; prevede un'architettura il cui primo requisito è la scalabilità a lungo termine (dove scalabilità è intesa come capacità di integrare nel tempo componenti eterogenei evitando tecnologie proprietarie); il secondo requisito è la soluzione basata su siti geograficamente distanti tra loro, autonomi amministrativamente e dotati di tecnologie differenti che replicano i contenuti degli archivi. DOM si colloca nel quadro di azione della British Library, che da alcuni anni ricerca strategie e programmi per la gestione e la conservazione a lungo termine del materiale digitale, dirigendo i suoi sforzi su due piani convergenti: l'infrastruttura tecnologica e il cambiamento normativo adeguato (che si è concretizzato nel Legal Deposit Act del 2003, il provvedimento legislativo del deposito legale anche per le pubblicazioni digitali).</p> <p>L'adeguamento tecnologico si fonda su due punti considerati irrinunciabili:</p> <ol style="list-style-type: none"> <li>1) assicurare l'autenticità e l'integrità dei documenti digitali;</li> <li>2) definire un modello in grado di supportare una vasta gamma di materiali e di corrispondere ai requisiti di conservazione definiti dalla British Library.</li> </ol>

Anche in ambito archivistico italiano si è pensato a centri di deposito per la preservazione, e qui si riporta una proposta, maturata in ambito accademico, di deposito come centro cui convogliare la preservazione da parte di diverse istituzioni: in questo caso, si potrebbe parlare sicuramente di repository dedicato o di repository istituzionale, per un modello di servizio di tipo centralizzato. Per quanto riguarda la conservazione documentaria amministrativa della pubblica amministrazione, invece, è stato presentato un apposito disegno di legge (privo di ogni riferimento tecnico) per la costituzione di una società "Banca della memoria digitale", riportato in una nota nell'introduzione al Cap. 3 di questo studio.

<sup>56</sup> <<http://www.jstor.org/>>.

La proposta d'ambito archivistico è la creazione di strutture riservate alla conservazione digitale, i “Centri di Conservazione Digitale (Ce.Co.Di.)”<sup>57</sup> o “Depositi Digitali”, strutture dedicate alla conservazione della memoria digitale di più soggetti produttori, perché la realizzazione di ‘sistemi’ sembra riuscire a fronteggiare la complessità delle procedure da svolgere, i requisiti giuridici da soddisfare e le competenze professionali necessarie per la corretta conservazione dei documenti digitali e degli archivi digitali, tutti fattori che richiedono alti costi e impegno di risorse spesso non disponibili da strutture limitate. Il modello di riferimento preferito, anche in questo caso è OAIS (Open Archival Information System).<sup>58</sup>

Anche secondo questa proposta, la struttura del “Centro di Conservazione Digitale” deve intervenire a partire dai processi che attengono alla produzione documentaria e alla formazione degli archivi dei soggetti produttori, per evitare che siano presentate entità elettroniche non compatibili con il processo di conservazione digitale per lungo termine o mancanti delle necessarie informazioni di contesto, quindi con un impegno alla preservazione che inizia già dalla formazione del documento, richiedendo che siano specificati i formati elettronici abilitati, individuati i metadati da gestire per ogni entità, stabiliti i tempi e le modalità trasferimento delle unità, evidenziate le rispettive responsabilità, definite le modalità di accesso e di fruizione del patrimonio informativo e documentario conservato.

Essendo questa proposta volta alla documentazione di tipo archivistico, tutto il sistema prevede le funzionalità adeguate alle prerogative di questo tipo specifico di documenti, come protocollazione, classificazione e fascicolazione, la gestione dei flussi documentali e dei procedimenti amministrativi, la generazione di strumenti di corredo finalizzati all’accesso di natura storico-culturale, la registrazione dei metadati relativi ai documenti cartacei, ed eventualmente di una loro copia digitale, per mantenere la visione complessiva e organica dell’archivio di un soggetto produttore. Si può, ad ogni modo, sottolineare la richiesta esplicita di alcuni passaggi:

- ≈ l’integrità degli oggetti digitali acquisiti
- ≈ la compatibilità delle unità documentarie con il processo di conservazione digitale
- ≈ una forte valorizzazione dell’insieme dei metadati di corredo.

Funzioni della struttura di questo Deposito Digitale conservativo:

---

<sup>57</sup> Stefano Pigliapoco in *Conservare il digitale* (2006).

<sup>58</sup> NASA, Consultative Committee for Space Data System. *Recommendation for space data system standards. Reference Model for an Open Archival Information System (OAIS)*. CCSDS 650.0-B-1 BLUE BOOK. January 2002. < <http://public.ccsds.org/publications/archive/650x0b1.pdf>>.

A seguito dell'acquisizione dei pacchetti informativi trasferiti dai soggetti produttori, completarli in modo da attuare il modello conservativo predefinito e immetterli nel sistema di conservazione;
Generare gli strumenti di accesso finalizzati al recupero dei documenti;
Monitorare l'evoluzione delle tecnologie e dei formati elettronici;
Provvedere all'esecuzione di eventuali processi di migrazione nel rispetto della normativa vigente;
Monitorare le registrazioni di sistema per accertare e dimostrare il regolare funzionamento dell'impianto e la puntuale esecuzione delle procedure provvedere alle operazioni di selezione e scarto se necessario;
Deve garantire le attività nel rispetto delle norme vigenti in materia di copyright, deposito legale, etc.
Deve assicurare la continuità della funzione di conservazione a lungo termine, anche sul piano giuridico.
L'impianto informatico del "deposito digitale" deve presentare il più alto livello di standardizzazione ed indipendenza dalle piattaforme tecnologiche per dare al responsabile della conservazione la possibilità di affrontare, con ragionevole sicurezza, i processi di migrazione o altro tipo di intervento che nel tempo si renderanno necessari. Al sottosistema di storage è richiesta la massima affidabilità, sicurezza e longevità e deve essere garantita la sicurezza fisica e logica dell'impianto e adottate le misure per il disaster recovery.

L'OAIS (Open Archival Information System), il modello concettuale di sistema informativo per l'archiviazione e la conservazione delle risorse digitali, è esattamente uno strumento per pianificare la preservazione digitale.<sup>59</sup>

OAIS nelle organizzazioni più accreditate (come molte tra quelle citate in questo paragrafo) e negli studi degli specialisti (v. Day, cit.) è diventato il fondamentale modello di riferimento<sup>60</sup> per la preservazione delle raccolte digitali, perché considerato come modello di riferimento ottimale o comunque preferito per la corretta archiviazione della risorsa digitale. Nel 2002 è stato definitivamente approvato come standard ISO 14721.

Il modello Open Archival Information System per la gestione di un archivio è un modello disegnato per sostenere le funzioni di accesso e di conservazione a lungo termine e non di un semplice bit storage, è stato sviluppato assumendo come unità minima di riferimento il bit, ed è il lavoro svolto dalla Consultative Committee for Space Data Systems, utilizzando il linguaggio di modellazione UML.<sup>61</sup>

<sup>59</sup> Brian Lavoie. *The Open Archival Information System reference model: introductory Guide*. "Digital preservation coalition technology watch series report". January 2004. Da questo documento sono tratte le fig. riportate nelle pagine seguenti. (URL: <[www.dpconline.org/docs/lavoie\\_OAIS.pdf](http://www.dpconline.org/docs/lavoie_OAIS.pdf)>).

<sup>60</sup> Ridi (2004), p. 26.

<sup>61</sup> Giovanni Michetti. *Metodologie di analisi per l'automazione dei sistemi documentari*. In Guercio (2002), p. 252-255. A proposito di modelli UML, p. 239: «Il linguaggio UML (Unified Modeling Language) fornisce il potere espressivo necessario per modellare un sistema in un'ottica Object Oriented: esso consente di rappresentare le classi e gli oggetti, lo stato degli oggetti e il loro comportamento. Prima della definizione di UML, la modellazione OO richiedeva una scelta fra linguaggi diversi ma talora molto simili, con minime differenze di potere espressivo: spesso i concetti di base erano sostanzialmente gli stessi, ma diversamente interpretati e rappresentati, in misura tale da generare confusione e scoraggiare l'impresa della modellazione ad oggetti. UML è il risultato dello sforzo di unificazione condotto a partire dal 1994 per l'individuazione di un linguaggio standard per la modellazione OO: proposto nel 1997 per la standardizzazione da parte dell'Object Management Group (OMG), ha ottenuto l'approvazione nella sua versione 1.1 ed è ora di dominio pubblico». Sulla modellazione Object Oriented (OO), p. 238-239: «... [diversamente dall'analisi tradizionale di un sistema] L'approccio Object Oriented è invece basato sulla definizione estensionale di classe, interpretata come

L'Open Archival Information System (OAIS reference model) sviluppato dal Consultative Committee for Space Data System Panel 2, propone il modello sviluppato a partire dal 1995 ma proposto nella sua versione finale nel 2002, descrive un framework concettuale per un sistema di archiviazione con un modello informativo complementare che ha rappresentato il punto di partenza di molte ricerche e molto lavoro di implementazione il modello è stato usato in progetti di ricerca quali Inter-PARES<sup>62</sup> e nella progettazione e realizzazione di archivi digitali<sup>63</sup> e di biblioteche digitali universitarie, nelle quali è stato spesso usato unitamente alle raccomandazioni RLG/OCLC (anch'esse sviluppate sul modello OAIS). Questo modello definisce la terminologia e i concetti rilevanti per l'attività di archiviazione digitale, identifica le componenti chiave e i processi ad essa endemici e propone un modello informativo per gli oggetti digitali e i loro metadati. Non specifica alcun'implementazione ed è dunque neutrale riguardo al tipo di oggetti e agli aspetti tecnologici: è possibile applicarlo ad ampio livello a file di immagini, oggetti *born/digital* e persino ad oggetti fisici. Questo tipo di modello opera in un ambiente composto da quattro entità: produttori (forniti dell'informazione), consumatori (coloro che utilizzano l'informazione archiviata), management (entità responsabile delle politiche a lungo termine dell'archivio) e lo stesso archivio delle risorse. Un'interessante implementazione è rappresentata anche dal progetto NEDLIB, con il modello di funzionamento Delivery and Kapture,<sup>64</sup> che descrive molto bene le varie fasi del data flow nell'assegnazione dei metadati: ricezione della pubblicazione, controllo della qualità, creazione dei metadati, aggiornamento del sistema esterno, creazione e sottomissione di SIP (Submission Information Package).

Nella fase di registrazione nel sistema di una nuova risorsa proveniente dall'esterno, vengono considerati i metadati (descrittivi, strutturali, di accesso, tecnici, di definizione su componenti e unità logiche) già presenti in quanto forniti da chi effettua l'invio della risorsa. I metadati esterni aggiunti nella fase d'integrazione per completare la descrizione della risorsa ed inserirla nel sistema, comprendono i metadati bibliografici e i record bibliografici

---

astrazione di un insieme di oggetti che condividono gli stessi attributi. Gli oggetti sono quindi elementi della classe, e non delle istanze. Ogni oggetto ha un'identità, uno stato e un comportamento (una caratteristica dinamica di un oggetto: descrive come esso interagisce con altri oggetti e come gli attributi dei suoi valori cambino nel tempo), e, a partire da tali caratteristiche, è possibile analizzare il sistema a seconda di diversi profili (strutturale, comportamentale) e diversi livelli di astrazione. Quindi, la filosofia Object Oriented permette di considerare sia l'aspetto statico sia l'aspetto dinamico dei sistemi: dati e funzioni sono integrati».

<sup>62</sup> Il progetto, la cui prima fase è durata dal 1999 al 2001, si trova attualmente nella seconda fase di ricerca (2002-2006) e si occupa del ciclo di vita dei record elettronici (dalla creazione alla conservazione permanente, prodotti in ambienti digitali sperimentali, dinamici e interattivi).

<sup>63</sup> Tra le iniziative di archivi digitali si veda United States National Archives and Records Administration's Electronic Records Archives (ERA), <[http://www.Archives.gov/electronic\\_records\\_archives/index.html](http://www.Archives.gov/electronic_records_archives/index.html)> e l'OCLC Digital Archive, <<http://www.oclc.org/digitalpreservation/about/archive>>.

<sup>64</sup> <<http://www.kb.nl/coop/nedlib>>. Della descrizione di questo esempio è stata fonte Canali (2006).

forniti, per esempio dalla bibliografia nazionale, oltre ai metadati forniti dal sistema di acquisizione.

Le fig. riportate sono tratte dall'URL: <[www.dpconline.org/docs/lavoie\\_OAIS.pdf](http://www.dpconline.org/docs/lavoie_OAIS.pdf)>.

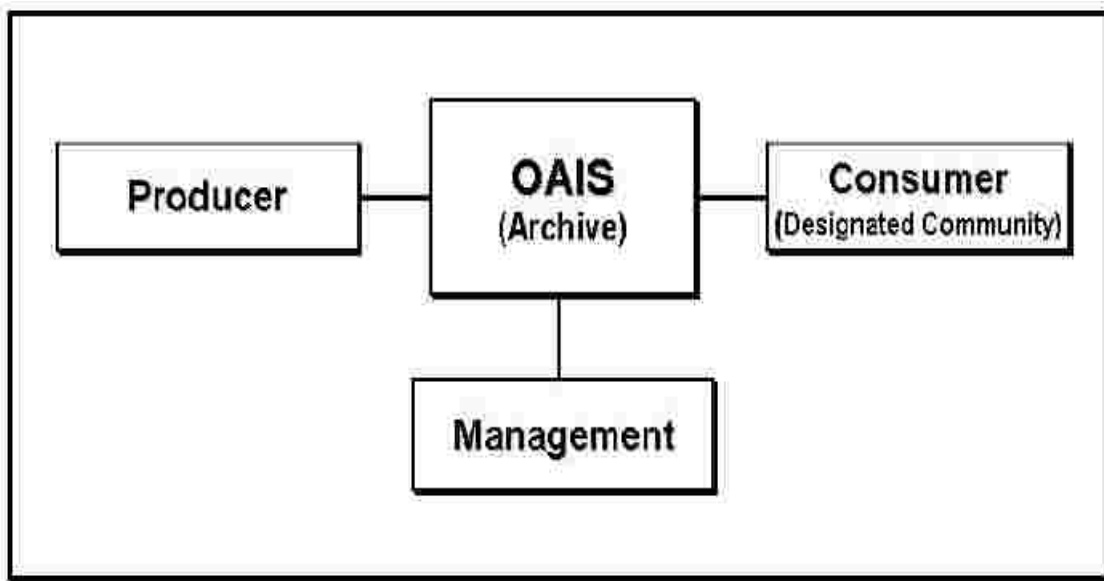


Figure 1: OAIS Environment

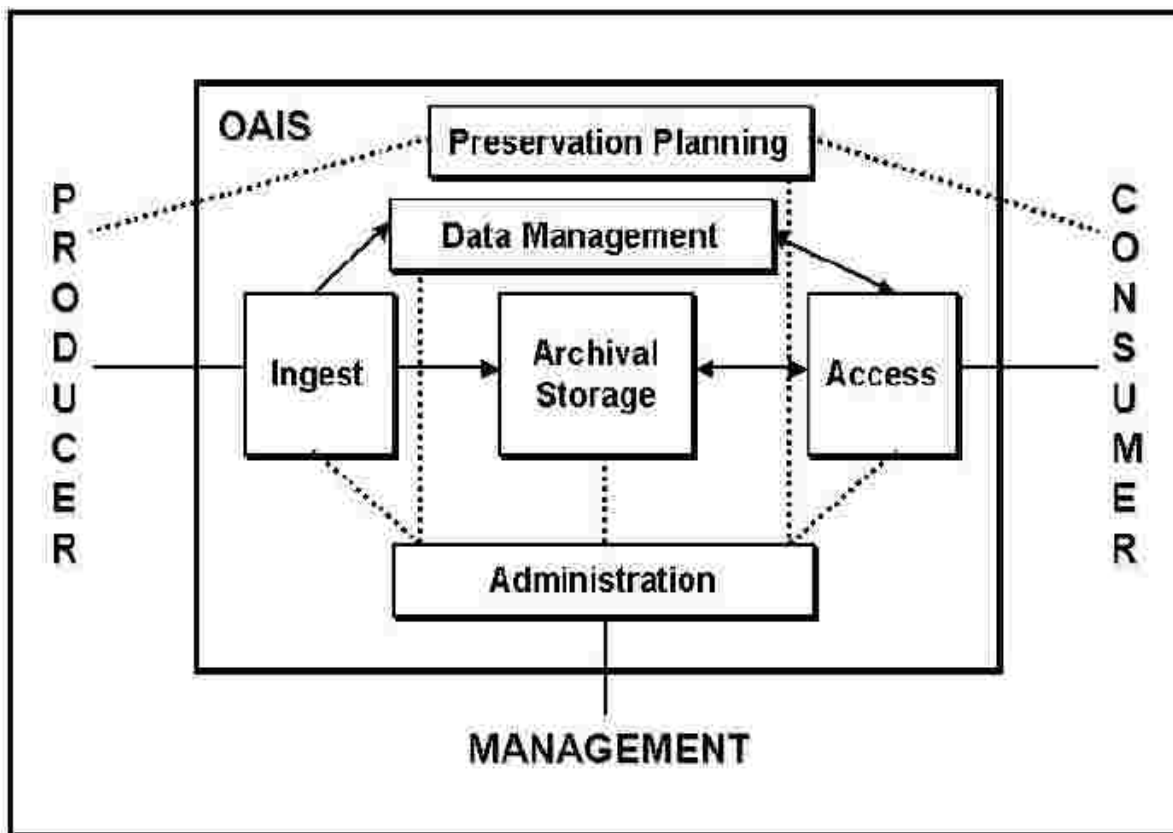


Figure 2: OAIS Functional Model

## 1.2.2. Le risorse digitali e i metodi di preservazione

Riccardo Ridi<sup>65</sup> individua quattro livelli d'invecchiamento delle risorse elettroniche:

- a) obsolescenza fisica dei supporti per la memorizzazione dei dati (cd-rom, floppy, nastri);
- b) obsolescenza dell'hardware per la decodifica dei supporti (lettori, drive);
- c) obsolescenza del software per l'interpretazione dei dati (word processor, programmi di grafica, browser);
- d) obsolescenza dell'hardware per l'esecuzione dei programmi d'interpretazione (microprocessori, computer) (un tipo di obsolescenza che è più "commerciale" che "tecnologica").

Accanto ai documenti elettronici<sup>66</sup> su supporti digitali ma tangibili, le risorse elettroniche locali, l'era del digitale ha permesso la creazione di nuove tipologie informative: risorse a volte dinamiche, con un aggiornamento continuo (come le banche dati e le pagine Web), risorse interattive che sono per definizione contingenti (come le liste di discussione), documenti ipermediali complessi composti da reti di collegamenti ed immagini.

Per una chiarificazione sull'impatto che tipi diversi di pubblicazione elettronici possono avere sulle procedure di conservazione, si riporta questo prospetto elaborato da Giuseppe Vitiello:<sup>67</sup>

<i>Tipologie</i>	<i>Criterio di classificazione</i>	<i>Impatto sulle procedure di conservazione</i>
Pubblicazioni on-line / Pubblicazioni off-line	Contenuto reperibile su rete vs. su supporto elettronico autonomo	Irrilevante (il contenuto va conservato in ogni caso come <i>file</i> elettronico)
Pubblicazioni professionali / Pubblicazioni occasionali	Professionalità del produttore	Gli editori professionali si adeguano ai formati standard di conservazione e alle procedure di informazione bibliografica
Pubblicazioni digitali fin dalla loro creazione (born digital) / digitalizzate	Create elettronicamente vs. scannerizzate	Impatto sui formati di conservazione e sulle procedure legali connesse al diritto d'autore
Pubblicazioni statiche / Pubblicazioni dinamiche	A contenuto "chiuso" vs. a contenuto "aperto"	Per un contenuto "aperto" conservazione simultanea di contenuto e delle strategie di interrogazione del contenuto
Pubblicazioni su Internet	Contenuto "aperto", varietà di contenuto, molteplicità dei formati	Procedure di selezione del materiale e del formato di conservazione

<sup>65</sup> Ridi (2003).

<sup>66</sup> Sulle caratteristiche del documento digitale si rimanda agli scritti di Salarelli in Tammaro – Salarelli (2006), in particolare alle p. 27-39.

<sup>67</sup> Vitiello (2002).

Nella letteratura più recente, sebbene questa materia non sia ancora del tutto ‘assestata’, sono focalizzate tre principali tecniche di preservazione, il refreshing, la migrazione,<sup>68</sup> l’emulazione, alle quali si aggiungono altre ipotesi, sulla conservazione della tecnologia, il riversamento su supporti analogici, l’uso di standard informatici, etc.:

1. <i>Refreshing</i>	Copia dell’informazione digitale in supporti più nuovi, prima che i vecchi si deteriorino fino a compromettere la lettura e il riversamento dei dati: riversamento periodico.
2. <i>Migrazione</i>	Memorizzazione del supporto digitale dell’informazione nell’ambiente corrente e secondo procedure concordate di immagazzinamento e trattamento: trasferimento periodico dei contenuti nei linguaggi e software correnti.
3. <i>Emulazione</i>	Emulazione dell’ambiente hardware e software necessario e conservazione del supporto d’informazione. È possibile la creazione di raccolte di emulatori software dell’hardware e di piattaforme tecnologiche originali.
4. <i>Museo delle tecnologie- Computer Museums</i>	Preservazione della tecnologia: praticabile per brevi periodi di tempo.
5. <i>Copia su supporti tradizionali: microfilm e su carta</i>	Il progetto del paper disk prevede la conversione dei dati su carta (in un linguaggio ad altissima densità che ricorda le schede perforate e i codici a barre), ricatturabili e riversabili su pc con uno scanner. Stampa e conservazione di copie su carta non acida (la carta permanente, a bassa acidità ed altissima resistenza, è oggi lo standard ISO 9706) o su microfilm: cosiddetto “backup atomico”.
6. <i>Affidamento agli standard</i>	Adozione di linguaggi e software standard (standardizzazione). Adozione di linguaggi orientati alla struttura logica del documento.

<sup>68</sup> Maria Guercio, in Guercio (2002), usa il termine ‘migrazione’ per quattro diverse strategie di migrazione: dei supporti, su piattaforme o con prodotti che garantiscono la compatibilità retroattiva (backward compatibility, come è nel caso di una nuova versione di videoscrittura, per esempio Word 97 per migrare documenti in Word 6.0), su piattaforme o con strumenti che sostengano l’interoperabilità, mediante l’utilizzo di applicazioni orientate, di formati standard.



7. <i>Reti e depositi federativi</i>	Preferenza per la distribuzione in rete rispetto a quella su supporti portatili; istituzione di un deposito legale federativo per i documenti distribuiti in rete. Moltiplicazione di copie conservate in depositi multipli di sicurezza assai distanti fra loro; applicazione di tali metodi non a singoli file, ma a loro insiemi organici (contestualizzazione). <sup>69</sup>
--------------------------------------	---

1. *Refresching. Copia dell'informazione digitale in supporti più nuovi*, prima che i vecchi si deteriorino fino a compromettere la lettura e il riversamento dei dati.

Si occupa della sicurezza dei supporti al fine di mantenere “i bit in buona salute”. Un esempio di applicazione di strategia di refreshing è quello fornitoci dalla Bibliothèque nationale de France che, nell’ambito del progetto Gallica, si è avvalsa di un nuovo speciale supporto, il Century Disk, prodotto dalla Digipress e pensato per durare più di un secolo; Questo CD ha una struttura in vetro temperato, oro, alluminio, rame, nichel, resistente a shock termici, scalfitture e raggi ultravioletti; permette un’archiviazione di documentazione multimediale a bassi costi e alto livello di standardizzazione.<sup>70</sup> Dalla Norsam di Palo Alto viene invece la proposta dell’“HD Rosetta”, *The Digital Rosetta Stone*,<sup>71</sup> un disco in silicone e nichel resistentissimo e leggibile per mezzo di uno speciale microscopio ottico a campo luminoso (il microscopio Iris 2000 System) che permetterebbe all’High density disk di essere letto sia attraverso una lente che per mezzo di un monitor in formato digitale.

Faiella<sup>72</sup> al proposito dice: «L’evidente debolezza di questa strategia sta nel fatto che la buona salute dei bit e l’integrità dei supporti abbiano in realtà poca significanza senza la preservazione dell’ambiente hardware e software necessari per accedere al contenuto informativo... Pur essendo una tecnica di vasto impiego, presenta una serie di problemi, il primo dei quali è senza dubbio quello della possibile perdita di informazioni – nel senso della funzionalità e del contesto – in corso di conversione. A ciò si aggiungano la difficoltà di pianificare questo tipo di interventi nel medio e lungo termine – vista l’impossibilità di prevedere i tempi e i modi dell’obsolescenza – e l’alto costo di applicazione».

<sup>69</sup> Ridi (2003).

<sup>70</sup> Sull’applicazione del Century Disk (v. <<http://www.mosarca.com/CDINFO/CENTURY.htm>>) al progetto Gallica: <[http://www.culture.gouv.fr/culture/mrt/numerisation/fr/seminaire\\_du\\_161198/projet\\_n11.htm](http://www.culture.gouv.fr/culture/mrt/numerisation/fr/seminaire_du_161198/projet_n11.htm)>.

<sup>71</sup> Su questo supporto si rimanda a Alan R. Heminger - Steven B. Robertson. *The digital rosetta stone: a model for maintaining - long-term access to static digital documents*. Communications of the Association for Information Systems. 3, 2 Jan 2000. URL: <[www.norsam.com](http://www.norsam.com)>.

<sup>72</sup> Faiella (2000).

2. *Migrazione. Memorizzazione del supporto digitale dell'informazione nell'ambiente corretto e secondo procedure concordate di immagazzinamento e trattamento.*

I dati sono memorizzati in un formato indipendente dal software e migrati attraverso diversi ambienti tecnologici. La strategia della migrazione si concentra sull'oggetto che viene periodicamente trasferito da una configurazione hardware/software ad un'altra. È la soluzione più costosa e può essere giustificata solo per documenti particolarmente preziosi.

3. *Emulazione dell'ambiente hardware e software necessario e conservazione del supporto d'informazione: l'aspetto e la funzionalità di una risorsa digitale sono ricreati con l'emulazione delle caratteristiche e dell'ambiente tecnologico superato.*

L'emulazione si basa sul sistema di metadati a corredo del documento digitale che descrivono l'ambiente tecnologico. In questa situazione non si trasforma l'oggetto originale (lasciato come stringa di bit) ma si ricrea l'ambiente. I costi precisi dell'emulazione sono per lo più ignoti, ma ogni costo viene rimandato al momento dell'uso dell'oggetto. L'emulazione è l'opzione "just in time" (mentre la preservazione della tecnologia che richiede di conservare il necessario software/hardware è una metodologia "just in case) e si pone l'obiettivo di mantenere inseparabili contenuto e contesto. L'emulazione si fonda sulla sperimentata possibilità degli emulatori di riprodurre il comportamento di hardware obsoleti in altri sistemi; esiste una comunità di sviluppatori software molto attiva che si occupa di creare emulatori in grado di far rivivere nei moderni computer videogiochi creati per macchine di una ventina di anni fa. Centinaia di siti Web offrono gratuitamente emulatori quali ad esempio MAME (Multiple Arcade Machine Emulator) in grado di emulare un buon numero di piattaforme, <<http://mame.retrogames.com>>. L'idea portante di questa strategia sarà dunque di "incapsulare", insieme con i documenti digitali, dei file che rappresentino il software e il sistema operativo originali, le specifiche dell'emulatore della piattaforma hardware, e metadati relativi al documento e alle modalità di fruizione delle informazioni contenute nella "capsula". I punti di forza di questa proposta stanno nell'idea di trattare il documento elettronico senza separarlo dal contesto in cui è stato creato – preservandone quindi sia il contenuto sia l'integrità – e di aspirare ad essere una strategia a lungo termine, neutrale rispetto alla tipologia del documento, indipendente da successivi interventi umani e flessibile ad una economia di scala. Ma si riconosce che, sebbene tali interventi richiedono risorse elevate, non eliminano le rischiose e impegnative attività di migrazione né riducono

le difficoltà dell'utenza costretta a misurarsi con strumenti assai diversificati e spesso obsoleti anche dal punto di vista della presentazione e della modalità di ricerca.<sup>73</sup>

4. *Preservazione della tecnologia*: i dati sono preservati insieme all'hardware e al software da cui dipendono.

Per alcuni oggetti digitali, almeno nel breve periodo, questo può essere la migliore soluzione poiché assicura che il materiale è accessibile attraverso gli strumenti che riescono a leggerlo. In questo caso la biblioteca diventa un misto tra un centro avanzato d'informazione, che usa le più moderne tecnologie, ed un museo delle tecnologie. Ma nel lungo periodo per il mantenimento di hardware e software obsoleti, questa soluzione potrebbe diventare non più praticabile.

5. *Stampa su supporto analogico riconosciuto*.

Le esperienze di preservazione che si concentrano nella conservazione dell'oggetto digitale, finora più numerose, tentano di trovare un supporto di memorizzazione che sia duraturo. Il supporto digitale ha un'aspettativa di vita relativamente breve, non superiore a cento anni. Convertire il contenuto digitale ad un formato analogico con qualità di preservazione a lungo termine noto, è sembrato ad alcuni una soluzione e c'è chi usa, paradossalmente, la carta per procurare un elementare backup. Viceversa la soluzione microfilm/accesso digitale è stata esplorata per la conservazione dei libri a stampa. Ma con la crescente complessità dell'informazione digitale, queste strategie sono limitate e il contenuto elettronico e la funzionalità della risorsa possono andare persi. Nuovi supporti di memorizzazione con una durata di vita molto maggiore sono state sperimentate dalle biblioteche.<sup>74</sup>

---

<sup>73</sup> Il metodo emulativo è sostenuto dai progetti condotti presso l'Università del Michigan da Margaret Hedstrom ed è stato oggetto di un accurato studio da Jeff Rothenberg (1998), <<http://clir.org/pubs/reports/rothenberg/contents.html>>. Fresko – Tombs. *Digital preservation Guidelines* e il documento proposto come standard ISO dal Consultative Committee for Space Data System, Reference, Model for an Open Archival Information System (OAIS).

<sup>74</sup> La Biblioteca Nazionale Francese ha realizzato un progetto di digitalizzazione di oltre il numero di 86.000 volumi sul Century Disc. Il Century Disc (CD in vetro temperato rivestito di oro e alluminio e da una lega di nichel e rame) ha la caratteristica di resistere a quasi tutti i possibili disastri ambientali e a temperature da -150° a +350°, oltre che, ovviamente, garantire la longevità che - come dice il nome - va oltre il secolo. Il progetto è consultabile all'URL <[http://vds.cnes.fr/pin/presentations/2002/cd\\_choix\\_bnf/index.htm](http://vds.cnes.fr/pin/presentations/2002/cd_choix_bnf/index.htm)>. Altre informazioni su attività di riversamento per l'archiviazione sono disponibili a partire dalla pagina del Gruppo di lavoro per la "Pérennisation de l'information numérique" (PIN) presso il Centre Spatial de Toulouse au Centre National d'Etudes Spatiales (CNES), dell'Associazione Aristote, che dal 1988 raggruppa i grandi organismi ed enti francesi per l'utilizzo evolutivo delle telecomunicazioni e della trasmissione dei dati, come INDRIA (Institut National de Recherches en Informatique et en Automatique), CEA (Commissariat à l'Energie

Nella memorizzazione, inoltre, si usa un formato dei dati indipendente dall'ambiente tecnologico.<sup>75</sup>

#### 6. Affidamento agli standard.

Il passaggio in formati standard, particolarmente adatto per grandi e complessi archivi digitali, è tra le proposte che hanno finora ottenuto i consensi maggiori e promettono sviluppi interessanti e utilizzabili in contesti operativi diversificati anche di piccole dimensioni. La conversione in formati standard, indipendenti dalle tecnologie, basati sull'uso di linguaggi di marcatura (SGML/XML),<sup>76</sup> della rappresentazione originaria dei documenti e dei metadati di contesto e di relazione sembra destinata, nel medio e lungo periodo, ad una implementazione diffusa. Si tratta di una soluzione che risponde ai requisiti di base dell'Open System Interconnection Reference Model (OSI),<sup>77</sup> poiché utilizza standard non proprietari, applicabili a sistemi informatici diversi, largamente diffusi, comprensibili anche da parte di non specialisti, indipendenti dall'hardware, ben documentati. Questo metodo presenta, in generale, il vantaggio (rilevante per le istituzioni preposte alla custodia) di ridurre enormemente il numero di formati da gestire e di contenere il numero di interventi di migrazione.

Tra gli strumenti standard XML si è ormai imposto all'attenzione generale, poiché offre un metodo diffuso, a basso costo e scalabile per affrontare la diversificazione e la frammentazione della produzione documentaria e delle sue articolazioni. Lo standard apre ulteriori e rilevanti possibilità per lo sviluppo di sistemi documentari informatici soprattutto perché consente, oltre alla gestione dei riferimenti esterni al documento e alle sue partizioni, anche il trattamento della struttura logica e semantica dei contenuti. Per svolgere una funzione significativa nei processi di automazione del settore documentario in relazione sia

---

Atomique), EDF (Electricité de France), URL: < <http://www.aristote.asso.fr>>. Gruppo PIN, URL: <[http://vds.cnes.fr/pin/pin\\_stockage.html](http://vds.cnes.fr/pin/pin_stockage.html)>. Vedi anche il *Report*, MINERVA (2005), p. 42.

<sup>75</sup> I formati utilizzati sono ad esempio ASCII, SGML, XML. Tom Shepard ha presentato UPF, Universal Preservation Format, formato di memorizzazione definito perenne. Shepard, con il patrocinio del WGBH Educational Foundation e il finanziamento del US National Historical Publications and Records Commission of the National Archives ha realizzato l'Universal Preservation Format per la conservazione di lungo periodo di tutti i supporti multimediali

<sup>76</sup> La decisione sul formato da utilizzare per la migrazione (XML, HTML, PDF, etc.) è una scelta di natura sia tecnica che organizzativa determinata anche dalle esigenze di conservazione: in alcuni casi è rilevante concentrare l'attenzione sulle funzionalità del sistema, in altri è essenziale mantenere integra la rappresentazione estrinseca del documento, utilizzando il formato PDF. Alcune indicazioni sono state fornite anche dalla normativa tecnica del CNIPA (ex AIPA) in materia di riproduzione ottica e di conservazione dei documenti informatici.

<sup>77</sup> Si tratta di un modello di riferimento di interconnessione di sistemi aperti realizzata in accordo a standard ISO per lo scambio di dati e la comunicazione tra sistemi diversi.

al contenimento dei costi, sia all'efficienza dei risultati, deve essere accompagnato da un uso diffuso di uno strumento complementare, le Document Type Definition normalmente identificate con l'acronimo DTD. Si tratta di modelli di documenti (strutture di elementi costitutivi e attributi) che descrivono tipologie e classi di entità documentarie, migliorando in questo modo «la permanenza, la longevità e l'ampio riutilizzo dei propri dati, insieme alla prevedibilità e all'affidabilità della loro elaborazione, che necessitano di progettazione logica e concettuale» e, naturalmente, richiede un approccio seriamente interdisciplinare per la definizione di regole comuni.<sup>78</sup>

---

<sup>78</sup> Sulla metodologia XML per la conservazione e l'accesso ai documenti informatici si basa il progetto statunitense SDSC-NPACI-NARA (Supercomputer Center di S. Diego-National Partnership for Advanced Computational Infrastructure-National Archives & Records Administration), Persistent Archives and Electronic Records Management, "Methodologies for Preservation and Access of Software – Dependent Electronic Records". Documentazione disponibile all'URL: <[http://www.sdsc.edu/NARA/Publications/nhprc\\_latest.pdf](http://www.sdsc.edu/NARA/Publications/nhprc_latest.pdf)>. Fonte per gli approfondimenti di questa parte è stata Guercio (2002).

### 1.2.3. Web archiving: sistemi e criteri di raccolta per la preservazione dei siti Web.

Ad oggi si possono distinguere,<sup>79</sup> dal punto di vista tecnico, tre sistemi di raccolta e, dal punto di vista metodologico, tre approcci per le scelte della raccolta, normalmente utilizzati dai migliori esempi internazionali di archiviazione del Web.

#### *Sistemi di raccolta*

1. selezione manuale	selezione manuale è totalmente all'intervento umano
2. harvesting <sup>80</sup> automatico senza selezione	harvesting automatico sullo spazio Web senza selezione: si appoggia a software chiamati crawler che setacciano la rete e raccolgono pagine Web sotto forma di istantanee (snapshots)
3. harvesting automatico con selezione o parametrizzazione manuale	vengono impostati i crawler per puntare su siti precedentemente selezionati indicati come rilevanti

Si è qui riportato il sistema dell'Harvesting automatico in quanto, a livello internazionale, è considerato tecnologicamente il più efficiente e sostenibile, come dimostra anche l'esperienza di motori di ricerca come Google; alternative all'harvesting sono prese in conto solo dove esso non è in grado di dare i risultati ricercati (per esempio, nel caso del deep Web,<sup>81</sup> come specificato nelle pagine seguenti).

<sup>79</sup> Questa distinzione è anche in Di Giammarco (2004); AIB (2004).

<sup>80</sup> Harvester è un software che realizza la funzione di harvesting, ossia la raccolta automatica, in un unico deposito digitale, dei metadati presenti nelle basi dati remote. I metadati descrivono le risorse informative disponibili ai relativi indirizzi. L'harvesting può riguardare anche i dati: per full-text harvest si intende una raccolta di testo volta a costruire un indice a pieno testo con collegamenti alle risorse, per full-content harvest si intende una raccolta che registra parti estratte dalle pagine per presentare, all'interno di un indice a pieno testo, i termini di ricerca nel loro contesto.

<sup>81</sup> Si parla di "Web invisibile" o "deep Web" costituito da file inaccessibili che non vengono rintracciati dai motori di ricerca (e quindi non raggiungibili nemmeno da un crawler). Infatti, i miliardi di pagine e documenti indicizzati dai motori di ricerca rappresentano solo una piccola porzione delle risorse potenzialmente fruibili su Internet. È nominato Web profondo quello di risorse pubblicamente accessibili, ma non incluse negli archivi dei motori di ricerca. Sono compresi materiali difficili da indicizzare, ad es. basi dati, alcuni formati di file pagine successive a quella in cui si richiede una autenticazione, pagine dinamiche. Tra questi si indicano di solito: siti non accessibili liberamente (per esempio a pagamento e protetti da password); siti che presentano le pagine HTML come risultato di una interrogazione da parte di un utente (per esempio la ricerca in un catalogo dove l'utente inserisce il titolo, l'autore ecc di un libro). Nel primo caso l'harvesting potrà funzionare solo se il sito "aprirà le porte" al crawler (per esempio fornendo la password all'istituzione depositaria). Nel secondo caso occorrerà una forte collaborazione tra il produttore dell'informazione e la biblioteca depositaria. Per un'analisi delle dimensioni del Web profondo cfr. Guerrini – Gambari (2002), p. 288-291, ma anche le interessanti

Soluzioni diverse dall'harvesting sono il download del sito e il plug-in lato Server Web:

≈ il download del sito in un formato standard è considerato più costoso e poco efficiente, perché i produttori del sito Web dovrebbero dotarsi di un software per lo scarico in formato standard (tra l'altro, in caso di deposito l'istituzione depositaria sarebbe anche impegnata in un'accurata verifica), mentre, nel caso dell'harvesting se lo ritengono necessario, preparano solo un file (robots.txt) da mettere nella radice (root) del sito con le istruzioni per il crawler;

≈ il plug-in lato Server Web è considerato "più invasivo",<sup>82</sup> come, per esempio, è PageVault<sup>83</sup> che, installato sul Server, registra tutte le risposte alle richieste pervenute.

*Metodi risultanti dai diversi criteri di selezione*

A	Dominio completo	Il raccoglitore scorre tutto il proprio spazio Web, collezionando tutto il possibile.	Sistema di raccolta dell'harvesting automatico (sistema 2 cit. preced. tabella)
B	Deposito	Deposito legale o deposito volontario di contenuti Web proprietari.	Sistema di raccolta dell'harvesting automatico per il dominio pubblico nazionale (sistema 1 cit.) e accordi per un deposito (obbligato o volontario) di records esportati da banche dati che alimentano il deep Web e di documenti non più attuali, possibilmente in formato standard e con metadati per garantire l'accessibilità nel tempo.

considerazioni di Vitali (2004 b). Il deep linking è l'operazione di inserire all'interno di una pagina di un sito Web un collegamento (*hyperlink*) a una pagina *interna* di un sito differente. Il termine significa letteralmente "collegamento in profondità", indicando appunto che si scende in profondità nel sito remoto; può pertanto essere tradotto anche con *collegamento diretto* (in quanto tralascia la prima pagina del sito e tutte quelle intermedie). Per contrapposizione, talvolta il collegamento alla prima pagina (home page) di un altro sito viene detto surface link (collegamento di superficie). Punto di riferimento essenziale per il deep web è considerato il testo di Michael K. Bergaman, *The deep web: surfacing hidden value*. "The journal of electronic publishing" 7 (2001), 1. URL: <<http://www.press.umich.edu/jep/07-01/bergman.html>>.

<sup>82</sup> AIB (2005), § 3.

<sup>83</sup> <<http://www.projectcomputing.com/products/pageVault/>>.

C	Selettivo per qualità o tematico	Vengono individuati settori di interesse, un argomento o uno specifico insieme, definite porzioni dello spazio Web o particolari risorse.	Sistema di raccolta di selezione manuale (sistema 1 cit.)
D	Selettivo per tipologia di Web contents	Vengono individuati tipi di contenuti Web di interesse: per es., il Virtual Remote Control (VRC) per i siti considerati a rischio di perdita d'informazione.	Sistema di raccolta dell'harvesting automatico con selezione o parametrizzazione (sistema 3 cit.)

Esempi dei criteri considerati:

#### A. Dominio completo

- ≈ Internet Archive negli USA:<sup>84</sup> è rivolto ad una copertura di livello internazionale, con il proposito di poter raccogliere e preservare l'intera sfera Web;
- ≈ il progetto Nordic Web Archive,<sup>85</sup> implementato da Danimarca, Finlandia, Islanda, Norvegia e Svezia, volto ad offrire un accesso comune alla raccolta di livello nazionale svolta dai singoli Paesi.

Va precisato che ogni harvesting ha tutti i limiti della periodicità del campionamento e della navigazione "anonima", a fronte di una personalizzazione che l'utente effettua durante la navigazione, quando accede ad un sito Web: ad esempio, vi sono siti come Amazon che presentano pagine personalizzate se un utente effettua il login.

#### B. Deposito legale o deposito volontario di contenuti Web proprietari

La questione del deposito determina un'altra categoria di modelli di archiviazione Web: molti paesi l'affrontano: per esempio, Canada, Francia, Finlandia, Germania, Danimarca, Italia, Norvegia, Regno Unito, Svezia;<sup>86</sup> è interessante anche la sperimentazione di soluzioni

<sup>84</sup> <<http://www.internetarchive.org>> - <<http://webarchive.org>>

<sup>85</sup> <<http://www.nwa.nb.no/>>.



di deposito volontario, regolata tramite accordo tra pubblicazioni elettroniche ed editori (Olanda), o di entrambe le modalità: per recuperare il deep Web e per il progresso (i numeri non correnti) dei seriali, come in Francia e nel Regno Unito.

Va sottolineato che di per sé il deposito dovrebbe garantire nel lungo periodo l'accesso alla collezione così raccolta, l'autenticità dei documenti archiviati e la persistenza nel tempo dei riferimenti: ciò non solo per via dell'applicazione di standard tecnologici di descrizione e trattamento, ma anche tramite tutto il 'procedimento' dell'acquisizione delle risorse.

### C. Modello selettivo:

- ≈ Pandora, l'archivio, realizzato dalla National Library of Australia, delle pubblicazioni e materiale online, selezionato in base ad una Carta delle collezioni;
- ≈ l'archivio del Web britannico<sup>87</sup> segue un'archiviazione basata sui settori d'interesse afferenti alle sei istituzioni culturali che aderiscono all'iniziativa:
  1. *lead partner* è la British Library,<sup>88</sup> che attua una raccolta più generale basata su siti che rivestono una particolare importanza culturale, storica e politica;
  2. la Wellcome Trust Library,<sup>89</sup> si occupa dei siti di medicina;
  3. l'Higher Education Funding Council of England, dell'istruzione e dell'educazione;
  4. il Joint Information Systems Committee (JISC);
  5. la National Library of Wales;<sup>90</sup>
  6. la National Library of Scotland, che colleziona siti che rispecchiano la vita contemporanea del paese.
- ≈ USA Minerva (Mapping the Internet Electronic Resource Virtual Archive)<sup>91</sup>: progetto nato per la conservazione di particolari contenuti online, come le tornate elettorali americane 2000 e 2002, i giochi olimpionici invernali del febbraio 2002, etc. Si tratta, quindi di un Web archiving tematico, cioè di una forma di archivio solitamente basata su materiali *born digital*, centrata su un particolare argomento od evento.

---

<sup>86</sup> Legislazione sul deposito legale delle risorse digitali nei vari paesi e un'ampia bibliografia si possono trovare alla pagina di PADI. *Legal deposit*. <<http://www.nla.gov.au/padi/topics/67.html>>. Un'indagine conoscitiva è stata quella di Andrew Charlesworth. *Legal issues relating to the archiving of Internet resources in the UK, EU, US and Australia A feasibility study undertake for the JISC and Wellcome Trust*. (2003). <[http://www.jisc.ac.uk/uploaded\\_documents/archiving\\_legal.pdf](http://www.jisc.ac.uk/uploaded_documents/archiving_legal.pdf)>.

<sup>87</sup> <<http://www.webarchive.org.uk/>>; fonte Di Giammarco (2004).

<sup>88</sup> <<http://www.bl.uk>>.

<sup>89</sup> <<http://www.wellcome.ac.uk>> .

<sup>90</sup> <<http://www.llgc.org.uk>>.

<sup>91</sup> <<http://www.loc.gov/minerva>>.

## D. Sistemi combinati per l'archiviazione dei contenuti Web

E' il sistema dove viene affiancato ai sistemi di harvesting automatico generale anche un harvesting con forti possibilità di parametrizzazione "manuale"<sup>92</sup>: visite più frequenti del crawler a siti segnalati come "più importanti", dove l'importanza è determinata dai "link in entrata" o altre referenzialità, in modo che si superino anche le problematiche che seguono alla raccolta di disomogenee tipologie di Web contents. Per esempio, nel mondo universitario, c'è il caso dell'Università di Heidelberg, con il progetto Digital Archive for Chinese Studies,<sup>93</sup> oppure della Cornell University con il sistema VRC<sup>94</sup> che ha lo scopo di monitorare, nel tempo, i cambiamenti dei siti Web, catturando quelli a rischio di perdita d'informazioni. Sono tecniche usate anche in Francia e in Olanda; presuppongono una ben definita 'politica di conservazione' alla base della pianificazione degli interventi.

-

A margine di questa rassegna sui sistemi di raccolta del Web si riporta, nel tentativo di fornire un contributo per un'immagine d'insieme sull'archiviazione del Web, un resoconto dallo studio condotto nel 2005 da Niels Brügger,<sup>95</sup> che nella sua pubblicazione *Archiving Websites : General Considerations and Strategies Archiving Websites*, affronta il problema dell'archiviazione del Web in una prospettiva che aiuta ad un completamento del quadro generale di ricerche sulle metodiche dell'archiviazione, e, perciò, si ritiene possa essere interessante fare qui presente.

La pubblicazione rientra nel progetto di ricerca danese MODINET (Media and Democracy in the Network Society) e ha avuto il contributo di studio della *Faculty of Humanities* e della Research Foundation dell'Università di Aarhus, in Danimarca, progetto che ha come obiettivo l'approfondimento delle tematiche sull'archiviazione che trova un primo approdo in questo report, l'implementazione di un software adeguato per l'uso e un'analisi preliminare all'avvio di una ricerca continuativa su questa materia.

---

<sup>92</sup> È tra le ipotesi di lavoro più sostenute dall'International Internet Preservation Consortium (IIPC).

<sup>93</sup> <<http://www.sino.uni-hedelberg.de/dachs/>>.

<sup>94</sup> Virtual Remote Control, <<http://irisresearch.library.cornell.edu/VRC/>>.

<sup>95</sup> Niels Brügger è associate professor PhD of Media Studies all'Institute of Information and Media Studies, University of Aarhus (Danimarca) e co-fondatore del Centre for Internet Research e ha collaborato a progetti per il NETARKIVET danese. La pubblicazione è del Centre for Internet Research Institute of Information and Media Studies Aarhus, Denmark. Si rimanda anche: Niels Brügger - Niels Ole Finnemann. - Birte Christensen-Dalsgaard (et al.). *Experiences and Conclusions from a Pilot Study: Web Archiving of the District and County Elections 2001. Final Report for The Pilot Project "netarkivet.dk"*. Copenhagen: netarkivet.dk, (2003). URL: <<http://www.netarkivet.dk/rap/webark-final-rapport-2003.pdf>>.

La ricerca ha come scopo il reperimento di una strategia d'archiviazione e di metodi adeguati per la creazione di Web archives ad uso ed implementati direttamente da parte di ricercatori, studenti o altri senza specifiche conoscenze tecniche, che usando computer standard di media capacità, desiderano salvare siti Web per successivi studi, ricerche, documentazione, con una prospettiva temporale di breve-medio termine. Tralasciando in questa sede le riflessioni teoriche e di metodo che sono emerse nella ricerca si già detto, si riportano qui invece le conclusioni sulle metodiche possibili, che partono da una distinzione tra metodi che cercano di salvaguardare la dinamicità del documento oppure permettono una conservazione 'statica' delle pagine. Va sottolineato il fatto che queste conclusioni sono strettamente legate agli studi condotti dal team del progetto sulle tecniche permesse dai software esistenti per l'archiviazione e che la strategia d'archiviazione che è stata così composta è basata proprio sul lavoro di testing tenuto sui software d'archiviazione.

I testi delle seguenti tabelle sono tratte dalla pubblicazione e tradotte dalla lingua inglese.

SW che archiviano l'intero sito				
Possibile	Non possibile	Considerazioni importanti		
		spazio	tempo	Montaggio
La struttura del sito	link a elementi per i quali è necessaria una connessione online continua	Dimensione in Mb e complessità	Serve per la programmazione dell'aggiornamento. Se fatto manualmente non è necessario, ma a volte può essere utile.	È un processo più difficile in quanto gli elementi scaricati non sono precisamente definiti in termini di tempo e di spazio
I suoi elementi				
Link				
link a elementi per i quali non è necessaria una connessione online continua				

SW per archiviazione statica di singole pagine o per screenshots				
Possibile	Non possibile	Considerazioni importanti		
		spazio	tempo	Montaggio
Singole pagine	Qualsiasi dinamica tra gli elementi	Dimensioni ridotte generalmente in Kb	È richiesta la presenza dell'operatore durante l'intero processo	È un processo più facile in quanto gli elementi scaricati non sono precisamente definiti in termini di tempo e di spazio

SW per screenrecording (film)				
Con movimento <sup>96</sup> sul sito Web				
Possibile	Non possibile	Considerazioni importanti		
		spazio	Tempo	Montaggio
Elementi e struttura del sito.  Singoli elementi per i quali l'attivazione dei movimenti richiede sempre la connessione online (streaming)	La possibilità di muoversi liberamente all'interno del sito	Complessità e dimensioni: generalmente in Mb e cresce rapidamente con la qualità delle immagini	È richiesta la presenza dell'operatore durante l'intero processo	Gli elementi scaricati sono precisamente definiti in termini di tempo e di spazio
Senza movimento sul sito Web				
Possibile	Non possibile	Considerazioni importanti		
		spazio	Tempo	Montaggio
Elementi e struttura del sito.  Singoli elementi per i quali l'attivazione dei movimenti richiede sempre la connessione online (streaming)	La possibilità di muoversi liberamente all'interno del sito	Complessità e dimensioni: generalmente in Mb e cresce rapidamente con la qualità delle immagini	Generalmente è controllato dall'operatore, anche se teoricamente non sarebbe necessario	Gli elementi scaricati sono precisamente definiti in termini di tempo e di spazio

Forme combinate: le forme in cui l'archiviazione è fatta attraverso i tre tipi di software. Nell'insieme permettono di salvare una maggior quantità di informazione.

	Uso	Consiste di	Tempo e spazio
documentazione	Assicura che tutti gli elementi vengono effettivamente archiviati e che vengono archiviati come e dove appaiono.	SW di archiviazione per l'intero sito più immagini statiche (fotografie) e screen recording.	Per i siti piccoli: possono essere scaricati interamente. Per i siti più grandi: devono essere selezionate aree specifiche, come quelle da sottoporre ad analisi successive o particolarmente importanti.
Esemplificazione	Per essere in grado di mostrare esempi di specifici elementi animati o tipi di elementi che richiedono una connessione online continua, senza comunque farne un oggetto di analisi.	SW di archiviazione per l'intero sito e screen recording.	Lo spazio del sito è meno importante. È importante sapere dove si trovano, nella struttura, gli elementi che hanno movimenti da esemplificare.
Contestualizzazione	Archiviare il sito Web così com'è, integralmente, senza discontinuità; per esempio strutture, elementi e animazioni, incluse quelle per cui l'attivazione richiede la connessione continua online.	SW di archiviazione per l'intero sito e screen recording.	Lo spazio del sito è meno importante. È importante sapere dove, nella struttura, gli elementi che hanno movimenti da esemplificare si trovano.

<sup>96</sup> Per movimento s'intende qualsiasi elemento della pagina Web contenente animazioni, anche non realizzate con tecniche vettoriali.

### 1.3. Alcuni studi sulla conservazione digitale

Non sono pochi gli studi sul tema della preservazione digitale, prodotti dall'inizio dell'interesse per questo problema, ma si è pensato utile presentarne qui una piccola cernita per rendere l'idea di cosa è stato scritto sull'argomento o si sia tentato di mettere in evidenza sui reali termini della questione o sullo stato dell'arte; sono stati aggiunte anche segnalazioni di contributi allo studio in seminari e conferenze internazionali, nonché alcuni repertori biblio-sitografici su Web.

È ad un gruppo di lavoro, costituito dai rappresentanti delle maggiori università statunitensi, insieme con Library of Congress, OCLC (Ohio College Library Center), RLG, IBM, gli editori Bell Communication ed Elsevier Science, che si deve il primo tentativo di creare una reale consapevolezza riguardo al tema della preservazione della memoria elettronica. A conclusione del loro sforzo congiunto fu reso pubblico, nel maggio del 1996, il Report Commission on Preservation and Access (CPA) and RLG. *Preserving digital information: report of the Task Force on Archiving of Digital Information*. Commissioned by the Commission on preservation and access and the Research Libraries Group. Washington: D.C, 1996. URL: <<http://www.rlg.org/ArchTF>>.

Feliciano Faiella<sup>97</sup> riferisce che questo rapporto sul tema della fragilità della memoria culturale nell'era digitale (ricordando l'esempio del primo messaggio di posta elettronica – spedito nel 1964 non si sa se dal MIT, dalla Cambridge University, o dal Carnegie Institute of Technology – che non è sopravvissuto semplicemente perché nessuno si era preoccupato di conservarlo, comportando il fatto che, perciò, non esiste alcuna prova documentaria che possa determinare quale dei gruppi sia stato il primo a sperimentare una tecnologia destinata a trasformare in maniera epocale il nostro modo di comunicare) lo studio osserva che sapere come conservare – anche operativamente – gli oggetti digitali, dipenda in parte dal sapere che cosa di essi debba essere conservato.

È del fondatore di Internet Archive, il seguente contributo, pubblicato a meno di un anno dopo l'apertura del sito Web: Brewster Khale, *Conservare Internet*, pubblicato col titolo originale *Preserving the Internet*, su “Scientific American”, 276 (March 1997), p. 82-83.

Margaret Hedstrom – Sheon Montgomery. *Digital preservation needs and requirements in RLG member institution*. December 1998. Tutto lo studio è disponibile all'URL: <<http://www.rlg.org/preserv/digpres.html>>.

Nel 1998, sono stati condotti due studi commissionati dalla Comunità europea:

---

<sup>97</sup> Faiella (2000)

Mark Fresko – Kenneth Tombs. *Digital preservation guidelines: the state of the art in libraries, museums and archives*. European Commission DG XIII/E-4, 1998, seguito da *Digitisation of library materials. Report of the concertation meeting and workshop*. European Commission DG XIII/E-4. Luxembourg 14 dicembre 1998, nel quale si denunciava la disattenzione dei creatori a prodotto terminato, la carenza di strategie a lungo termine, mancanza di standard e di linee guida per l'orientamento delle istituzioni responsabili, una sorta di "Medioevo digitale".

In Italia, Giancarlo Buzzanca. *Digit fugit ovvero osservazioni sulla conservazione del Web* e, Id. *Frammenti di storia dell'evoluzione del Web: un aggiornamento*. In *Manuale per la qualità dei siti Web pubblici culturali*. Ed. 2005, p. 209-234, che tratta, con ampi riferimenti bibliografici, della storia del Web in Internet, e delle prime esperienze di conservazione: le grandi organizzazioni internazionali Iccrom (International Centre for the Study of the Preservation and Restoration of Cultural Property), Icom, Icomos (International Council on Monuments and Sites) e il Getty Trust di Los Angeles, tutti già tra il 1993 e il 1996. Per Buzzanca la coscienza della criticità della long term preservation del digitale matura quasi contemporaneamente all'adozione stessa di strumenti informatici e quindi data almeno ai primi anni Sessanta, ma a questa generica coscienza non ha corrisposto una razionale politica di migrazione di dati tra piattaforme, sistemi operativi e diverse versioni dei software utilizzati.

Uno studio con una prospettiva particolare sull'archiviazione del Web è quello di Niels Brügger. *Archiving Websites : General Considerations and Strategies Archiving Websites*, Aarhus University Press, Århus, c2005.<sup>98</sup> Come presupposto del tema dell'archiviazione è posta la necessità che la ricerca su Internet sia in grado di fornire una capacità di stabilizzare e salvare l'oggetto dell'analisi e degli studi. Brügger inquadra la possibilità di scelta dell'archiviazione a due livelli, uno circoscritto, parziale, l'altro 'globale', identificato come "macro-archiviazione", e ha concentrato la ricerca sull'archiviazione di medie-piccole dimensioni, definita "micro-archiviazione" dei siti Web e sui metodi adeguati per la creazione di Web archives ad uso ed implementati direttamente ad opera di ricercatori e studenti. Prima di riportare gli schemi di programmazione di questo tipo di Web archiving, l'autore dà però conto delle riflessioni preliminari che hanno guidato alla stesura del tipo di metodologia applicabile nei casi specifici. Sono riportate riflessioni sulle caratteristiche del sito-Web come documento (con la distinzione tra 'documento', 'monumento' e 'imprint') e, partendo dalla constatazione che Internet è corredato di certe dinamiche "media-specifiche"

---

<sup>98</sup> Di questo studio si è trattato nel cap. 1.2.3.

fondamentali, che fanno sì che la stabilizzazione sia difficile, analizza la caratteristica peculiare primaria di Internet, identificata nella sua dinamicità, che si riscontra sia dal punto di vista dei contenuti, sia da quello prettamente formale, e si pone l'interrogativo se questa caratteristica vada salvaguardata nei processi di archiviazione o se si può optare per una conservazione 'statica' delle pagine. Interessante anche le riflessioni sulla necessità, ma anche la problematicità, dell'archiviazione, che può essere necessaria anche ad un livello 'inferiore' o limitato rispetto ai grandi progetti di ordine cooperativo e nazionale.

Ricognizione importante quella di Ingeborg Verheul. *Networking for digital preservation: current practice in 15 national libraries*. Munchen: K. G. Saur, 2006. (IFLA publications, 119). Il testo può essere consultato on-line all'URL: <<http://www.ifla.org/VI/7/pub/IFLAPublication-No119.pdf>>.

È il rapporto finale della ricerca che la Koninklijke Bibliotheekmm, la biblioteca nazionale olandese, ha svolto fra il 2004 e il 2005, su incarico dell'IFLA-CDNL Alliance for Bibliographic Standards (ICABS),<sup>99</sup> al fine di conoscere lo stato dell'arte e le prospettive di sviluppo dell'archiviazione digitale nel mondo delle biblioteche. La ricerca si è basata sull'indagine condotta solo sulle seguenti quindici biblioteche nazionali: Australia, Austria, Canada, Cina, Danimarca, Francia, Germania, Giappone, Nuova Zelanda, Olanda, Portogallo, Regno Unito, Stati Uniti, Svezia, Svizzera. La ricognizione ha compreso rilevamenti sui dati più generali (per esempio sulle legislazioni sul deposito dei documenti digitali, on-line e su supporto fisico), sugli archivi digitali (infrastrutture, servizi, software, metadati), sulle strategie di conservazione attuali e programmate e sulle attività in corso, con i progetti nazionali e internazionali in cui le singole biblioteche, concludendo con una mappa organizzativa delle biblioteche esaminate. Dall'analisi della realtà conservativa e dalla consistenza degli approcci nei Paesi presi in esame, sono state tratte considerazioni sulla condivisione di alcuni standard (OAIS e PREMIS), per lo sviluppo e la costruzione di archivi digitali e per trovare garanzie di un possibile accesso permanente, pur essendo questo ambito di ricerca in rapida continua evoluzione.

In Internet si trovano anche raccolte di risorse bibliografiche e sitografiche, nate proprio con lo scopo di fornire informazioni e documentazione di approfondimento nel settore.

Si riportano qui due tra gli esempi più conosciuti, implementati uno in Australia, e l'altro in Europa, per la precisione in Austria:

---

<sup>99</sup> Trattato più ampiamente all'inizio del cap. 2.3.

PADI, il portale australiano di risorse internazionali sulla preservazione digitale, aggiorna la sezione *DPC/PADI What's new in digital preservation*. Fornisce molti servizi, tra cui una selezione di risorse in continuo aggiornamento, relativa alle attività più recenti nel settore, una redazione sulla digital preservation, Padiforum e una mailing list. URL:<<http://www.nla.gov.au/padi/index.html>>.

*Web Archiving Bibliography*. A cura di The AOLA (Austrian On-Line Archive) Project - The AOLA Team. URL: <<http://www.ifs.tuwien.ac.at/~aola/links/WebArchiving.html>>.

AOLA è un progetto del Department of Software Technology, presso la Vienna University of Technology, in Austria. La pagina qui segnalata è un ricchissimo repertorio che comprende: *Presentations and Publications, Initiatives and their Publications, Further Publications, Crawling, Analysis+Accessibly, Tools, Guidelines and Best Practices for Creation of web material* («...in order to leverage preservation of material from the web over the long term»), *Newspaper Articles, Workshops, Meetings, Subject Gateways and Bibliographies*. Purtroppo non è stato aggiornato negli ultimi due anni (la data riportata sulla pagina web risale all'aprile del 2004), ma un'alta percentuale dei link è tutt'ora valida e sicuramente questa rassegna assume un valore 'storico'.

In Italia, è stata rilasciata nel 1998 e curata fino al 2003, ma ancora on-line, una Metarisorsa, a cura di Antonella De Robbio, sulla conservazione di documenti elettronici:

*Conservazione di documenti elettronici*. Pagina a cura di Antonella De Robbio. Data creazione 11/10/98, aggiorn. 26/10/03.

URL: <<http://www.math.unipd.it/~derobbio/preserv.htm>>.

Come indicato nella pagina Web, si tratta di una raccolta di risorse miste sulla conservazione di documenti elettronici: records, files, periodici elettronici, multimedia, «... una miscellanea eterogenea di informazioni selezionate da siti specializzati sull'argomento». Nelle intenzioni della curatrice lo strumento è rivolto a bibliotecari ed a quanti si occupano di gestione e conservazione di materiale su supporto digitale. Vi sono link siti di interesse generale, linee guida, raccomandazioni, riviste e newsletters specializzate, reports, atti di workshop, qualche bibliografia ed alcune informazioni su norme e formati.

Si vogliono qui riportare anche le pubblicazioni seguite ad alcuni convegni e seminari nazionali ed internazionali, ma svoltisi in Italia, che hanno dato vita ad incontri e scambi importanti e di cui poi sono stati, per l'appunto, pubblicati Atti che ne attestano la ricchezza e la varietà degli argomenti affrontati.



*Futuro delle memorie digitali e patrimonio culturale : The future of digital memory and cultural heritage: atti del Convegno internazionale, Firenze, 16-17 ottobre 2003.* A cura di Vittoria Tola e Cecilia Castellani. Roma : ICCU, 2004.

*Il futuro della memoria: la trasmissione del patrimonio culturale nell'era digitale. Torino, 10-11 novembre 2004.* Atti del convegno a cura di Agata Spaziante. Torino: CSI-Piemonte, 2005. M. Guercio ha considerato questo convegno una delle più autentiche occasioni di approfondimento e riflessione di tipo interdisciplinare:<sup>100</sup> in effetti è stata occasione di scambio informativo e dibattito, ma anche di applicazione concreta, tra aree disciplinari apparentemente molto distanti.

Sul problema dei supporti per la conservazione anche la pubblicazione degli atti del convegno: *Gli attuali supporti per la conservazione ottici e magnetici per i documenti digitali.* Convegno 21-23 novembre 2005. Roma: IRTEM, 2006.

Devono, invece, ancora esserne pubblicati gli Atti, ma sono disponibili le presentazioni o i pre-print delle relazioni, dei seguenti incontri internazionali:

Convegno internazionale “Conservare Il digitale. Un Confronto Internazionale. Asolo, 29-30 Settembre 2006”. Gli atti del convegno sono disponibili on-line nella loro prima stesura e nella pagina Web del convegno è dato avviso che si prevede di aggiornarli non appena sarà pronta la stesura definitiva. La documentazione si trova a partire dalla pagina:

<<http://www.ulssasolo.ven.it/webpublic/ddw.aspx?n = 2261&h = -2147477295&Tipod = Conv>>.

Convegno “Il digitale e la memoria: dal presente al futuro”, Milano 12-16 dicembre 2006 : Seminario Internazionale InterPARES2, 12-13 dicembre : I dibattiti di ArchiExpo, 14-15 dicembre. Associazione Nazionale Archivistica Italia – ANAI. Le presentazioni di sono disponibili on-line all’URL: <<http://www.anai.org/archiexpo/02Inter pares2.asp> >.

Conferenza internazionale “Cultural Heritage on-line. The challenge of accessibility and preservation”, Firenze, 14-16 dicembre 2006. Le presentazioni di sono disponibili on-line all’URL: <<http://www.rinascimento-digitale.it/indexEN.php?SEZ = 396>>.

---

<sup>100</sup> Guercio (2006)

## **2. Alcuni esempi di progetti e iniziative per la conservazione dei documenti digitali**

### **2.1. Iniziative e progetti internazionali.**

#### 2.1.1. The Internet Archive

L'Internet Archive è un'organizzazione fondata da Brewster Kahle, con sede a San Francisco, negli Stati Uniti, che dal 1996 raccoglie materiali presenti su Internet e siti Web con lo scopo di costruire una biblioteca digitale per tutti,<sup>101</sup> come dice il logo nell'intestazione del sito «Universal access to human knowledge».

Nel principio guida di quest'impresa si può riconoscere un proposito di libera e generalizzata accessibilità ai materiali conservati, come nella tradizione della biblioteca pubblica anglosassone. Questo progetto si caratterizza, rispetto a tutti i progetti di archiviazione che lo seguiranno, oltre che per la tempestività dell'iniziativa, per il fatto che la preoccupazione di archiviare il più possibile i prodotti del fenomeno in atto, sin dall'inizio, ha impedito di attendere esiti di ricerche su possibili criteri di ordinamento e senza applicare neanche una pur minima schedatura, identificando il sito Web e gli oggetti digitali solo in base all'URL, ed in questo modo ad oggi, salvando anche molti dei primi siti che lo hanno popolato, ha già ottenuto l'enorme risultato di conservare testimonianza di Internet come fenomeno storico complessivo, a beneficio dei contemporanei e delle generazioni future.

L'impegno di Internet Archive è stato da subito volto esclusivamente ad impedire che Internet, che si configura come un nuovo medium di fondamentale importanza storica, ed altri materiali "support-digital" potessero sparire con il passare del tempo, e perciò ha presto ottenuto la collaborazione della Library of Congress e lo Smithsonian Institute, al fine di conservare a lungo termine le sue registrazioni; riceve da vari enti pagine Web e altri materiali digitali, immagini etc.

Le collezioni conservate, riprese dall'intero World Wide Web o almeno della sua parte ad accesso libero e diretto, si dividono in: Web, full Web, metadati, mailing list, FTP, raccolte Usenet e arriva a riprendere fino a due miliardi di pagine nel mondo.

---

<sup>101</sup> «The Internet Archive is building a digital library of Internet sites and other cultural artifacts in digital form. Like a paper library, we provide free access to researchers, historians, scholars, and the general public».

Dalla pagina principale, si può accedere anche a collezioni speciali di Internet Archive, quali Hurricanes Katrina and Rita, le elezioni 2000 e 2005, e il National Archives, il Web Archive dei siti governativi della Gran Bretagna, che ha avuto avvio nel 2003, raccolti dall'Internet Archive per conto del Archives del Regno Unito.

Un altro progetto con il quale collabora Internet Archive è quello con la nuova Biblioteca d'Alessandria d'Egitto.<sup>102</sup>

Dalle pagine dedicate al Web<sup>103</sup> si può accedere ad una speciale raccolta retrospettiva dei siti che hanno contribuito alla "the Internet revolution". Tra questi siti, per i quali è possibile vedere la prima pagina conservata (queste prime sono dal settembre 1996 e gennaio 1997) e la pagina con le memorizzazioni in ordine per anno, si trovano i seguenti:

The Internet Movie Database, generato nel 1989 e tra i primi ad arrivare sul Web nel 1993, è la base di dati di film,<sup>104</sup> il primo Yahoo! "La guida di Jerry al World Wide Web", iniziata nel 1994 con una coppia di PC al Campus universitario di Stanford, che dopo sette anni divenne il fenomeno "Yahoo!";<sup>105</sup> il primo sito registrato è dell'ottobre 1996 ed è Amazon.com, nato nel luglio del 1995 come bookstore in linea, ad opera di Jeff Bezos, premiato nel 1999 'uomo dell'anno' dal Time Magazine's per la sua idea.<sup>106</sup>

Della propria attività di preservazione fornisce queste indicazioni: la produzione della collezione digitale avviene con due interventi, l'immagazzinamento e la preservazione vera e propria.

Immagazzinare le collezioni dell'archivio contempla l'analizzare, indicizzare e fisicamente codificare i dati. L'hardware consiste di PC con serie di dischi fissi IDE. I dati sono memorizzati sul nastro DLT e sui dischi fissi nei diversi formati, in base all'accumulo; i dati Web sono ricevuti e memorizzati nell'archivio di files ARC da 100 MB. ARC è proposto da Alexa Internet come standard per l'archiviazione di oggetti digitali.

Per la preservazione si provvede a mantenere l'accessibilità dei dati mentre i formati diventano obsoleti, anche se il nastro di DLT si è valutato possa vivere 30 anni, la regola pratica è migrare i dati ogni 10 anni; per i continui sviluppi nell'hardware dei PC la migrazione è sempre più frequente. Vengono raccolti anche i software e gli emulatori, per fornire queste conoscenze a chi fa ricerca. La raccolta dei siti Web avviene tramite un crawler di Alexa Internet; i robot di Alexa passano circa ogni due mesi in rassegna i siti accessibili sull'intero Web e ne memorizzano il contenuto, producendo una snapshot di

---

<sup>102</sup> <<http://archive.bibalex.org>>.

<sup>103</sup> <<http://www.archive.org/web/web.php>>.

<sup>104</sup> <[www.imdb.com](http://www.imdb.com)>.

<sup>105</sup> <[www.yahoo.com](http://www.yahoo.com)>.

<sup>106</sup> <[www.amazon.com](http://www.amazon.com)>.

ciascun sito al momento della sua cattura; ogni sei mesi il materiale memorizzato è ceduto gratuitamente all'Internet Archive; con questa tempistica, mediamente un sito non viene ripreso più di minimo sei, massimo dieci, volte in un anno. Alla base dell'archivio ci sono cinque diversi sistemi di Peer-to-Peer che permettono la distribuzione di file audio e video senza incorrere in problemi legati alla capacità di banda, una capacità che solo una grande azienda dovrebbe garantire.

Un'accorta trattazione dei termini per l'uso, dell'attento rispetto delle leggi sulla privacy e di copyright, sono pubblicate su una pagina dedicata a questi argomenti.<sup>107</sup> Già dall'inizio della messa on-line dell'archivio di Internet Archive è stato stipulato un accordo che ne governa l'uso; di questo accordo viene dato avviso con richiesta di accettazione – determinante per l'accesso - agli utenti prima della visualizzazione delle collezioni protette.

Dal 2001 l'archivio è consultabile on-line e la ricerca è affidata alla Wayback Machine: si tratta di un motore di ricerca che consente il recupero e la visualizzazione dei siti Web. La Wayback Machine («surf the Web as it wass»), uno strumento che consente di studiare i mutamenti delle pagine nel corso del tempo, ha creato alcune sezioni tematiche di pagine Web che documentano aspetti della vita sociale americana.

La ricerca si fa in base alle URL e riporta la lista di date di memorizzazione, che linkano alle snapshots. Vi è anche una ricerca avanzata che permette di inserire, dopo l'URL del sito oggetto della ricerca, opzioni di scrematura: lasso di tempo del contesto tra due date, tipi di file e altre differenze rispetto al comportamento dell'URL (redirect, duplicates, alias).

Tra gli aspetti non riusciti di questo archivio si riportano alcuni caratteri: talvolta non funzionano correttamente i link fra le pagine di uno stesso sito, mentre in altri casi rinviano automaticamente alle pagine on-line e non a quelle presenti nell'archivio; possono essere prive di immagini; non sono comprese le pagine che si formano dinamicamente o attraverso compilazione di maschere di ricerca o parti il cui accesso è subordinato alla digitazione di password; nell'archivio, al di là della data di memorizzazione, non esiste alcun dispositivo che permetta di contestualizzare i siti in relazione agli enti o alle persone responsabili; la presenza dei siti può considerarsi esaustiva ma è possibile, data l'estensione del Web e i meccanismi della raccolta con il robot, possono esservi delle assenze.<sup>108</sup> Pur con i limiti

---

<sup>107</sup> <<http://www.archive.org/about/terms.php>>.

<sup>108</sup> Vitali riporta il significativo esempio della presenza nell'archivio del sito ufficiale del Genoa Social Forum, <<http://www.genoa-g8.org>>, ma l'assenza delle pagine relative ai giorni del luglio 2001 con le relative iniziative di contestazione. Si nota però la presenza di siti, poi chiusi dall'autorità giudiziaria, come il <<http://www.brigaterosse.org>>, con materiali per lo studio di alcuni aspetti del terrorismo, oppure il sito della Presidenza del Consiglio italiano <<http://www.governo.it>>, di cui sono conservate una quarantina di istantanee tra il 1998 e il 2001 che rendono interessante il confronto tra la gestione ad opera dei governi diversi. E anche

segnalati, l'Internet Archive costituisce una risorsa probabilmente unica, non solo per tracciare la storia del Web nel suo complesso o di suoi singoli segmenti, ma anche per attingere materiali capaci di illustrare gli infiniti aspetti della nostra contemporaneità documentati da siti Internet attualmente non più accessibili o profondamente trasformati.

Suggestivo è il paragone che Vitali<sup>109</sup> fa della Wayback Machine dell'Internet Archive con una scatola di vecchie fotografie disordinate e spesso inutile di particolari, che, con queste sue attestazioni, ribalterebbe la logica effimera del Web. Si potrebbe anche notare che questo 'disordine' delle foto, che si possono collocare solo attraverso l'URL originario e delle date, in fondo è molto indicativo di quello che è Internet e riporta esattamente l'affastagliarsi, che più che indicare solidi contenuti, sottolinea proprio solo il temporale avvicinarsi continuo pagine Web, spesso nate solo per un consumo immediato e vecchie quasi subito la loro messa on-line, e implacabilmente coperte con quelle che vanno a sostituirle.

Va comunque sottolineato che, come ha notato lo stesso Vitali, ad oggi, probabilmente, quanto si è salvato delle prime pagine Web è recuperabile solo attraverso l'Internet Archive e sui personal computer di singoli produttori e (coscienziosi) operatori di siti Web, che li hanno conservati nei propri computer, fino a quando questi saranno utilizzabili.

Parallelamente all'Internet Archive, e poi ad esso associato, è stato condotto il "September 11 archive": si tratta di un progetto che raccoglie migliaia di siti Web sull'attacco alle torri gemelle e di tutti gli altri avvenimenti dell'11 settembre: ha quindi un carattere tematico e cronologicamente limitato. È nato per iniziativa di Kirsten Foot, professoressa di scienza delle comunicazioni dell'Università di Washington e di Steven M. Schneider, professore di scienze politiche presso l'Università dello Stato di New York in collaborazione con l'Internet Archive e con la Library of Congress. Questa esperienza, come le altre cui si è fatto cenno inizialmente (come altre raccolte di siti Web tematiche o legate ad eventi specifici come le elezioni presidenziali) sono state promosse dall'Internet Archive con la Library of Congress; ad esse si può accedere dall'home page del progetto MINERVA della stessa Library of Congress.<sup>110</sup> Queste esperienze sono il risultato della consapevolezza del ruolo di documentazione del Web su avvenimenti e fenomeni della società e del mondo contemporanei, grazie al quale vanno sedimentandosi fonti potenziali di estremo interesse per scienziati sociali e storici; in questo senso l'11 settembre ha certamente segnato una

---

della fragilità del Web poiché al cambio di governo vennero rimosse dall'archivio dei comunicati tutti i documenti della presidenza precedente.

<sup>109</sup> Vitali (2004 a)

<sup>110</sup> <<http://www.loc.gov/minerva/index.html>>.

svolta; l'archivio che ne raccoglie le testimonianze si presta efficacemente ad una verifica di quali e quanti elementi di novità e di specificità le reti telematiche introducano nei processi di costruzione degli eventi e nelle forme della loro ricezione da parte degli individui e molti approfondimenti possono essere fatti in merito. Contiene circa 30.000 siti, per 5 terabyte di materiali Web; su questi materiali è in corso la catalogazione per ottenere descrizioni accurate e fornire informazioni di contesto sui responsabili, con una classificazione dei soggetti produttori sulla base di una griglia predefinita, quindi un'insieme di documentazione dotato di una propria organicità e di strumenti esplicativi.

Ultimamente è stato messo a disposizione "ARCHIVE.it" (web 2.0 ), un servizio di archiviazione basato su abbonamento, che ha già coinvolto una vasta gamma di istituzioni americane. Archive.it permette agli abbonati di bloccare, catalogare e conservare il materiale messo in linea dalle loro proprie istituzioni, così come proveniente dal World Wide Web. Gli utenti possono generare ed accedere alle loro collezioni reperibili, senza avere bisogno del supporto tecnico o dell'infrastruttura. Permette che le istituzioni costruiscano, controllino e cerchino il loro proprio archivio di Web con un'applicazione facile da usare, senza richiedere alcuna perizia tecnica o unità centrali. Gli abbonati possono bloccare, catalogare ed archiviare il proprio Web site, il sito della loro istituzione o sviluppare le collezioni estratte dal Web e cercano e passano in rassegna la raccolta accumulata, una volta completi. Archivio-it è stato disegnato, stando alla dichiarazione d'intenti, su misura per individui e per i bisogni di molti tipi di organizzazioni, come sono gli archivi, le biblioteche, le istituzioni accademiche, le organizzazioni no profit, i musei, gli storici ed i ricercatori indipendenti. Tra partners ci sono: la Library of Congress, Library of Virginia, University of Toronto, University of Southern California, Indiana University and North Carolina State Archives, oltre ad altre Università e istituzioni accademiche e di conservazione culturale.

### 2.1.2. The Nordic Web Archive

Il progetto Nordic Web Archive (NWA)<sup>111</sup> delle biblioteche nazionali nordiche ha sviluppato, dal giugno 2000, la cooperazione tra gli istituti che svolgono progetti simili nei Paesi aderenti all'iniziativa: Finlandia, Danimarca, Islanda, Norvegia e Svezia. In questi Paesi sono in vigore legislazioni per il deposito legale dei documenti dello spazio Web nazionale e la cooperazione NWA ha permesso di instaurare una ricerca comune ed avanzata per affrontare la sfida della preservazione; tutte le informazioni sull'attività attuale del NWA vengono messe a disposizione sul sito Web. Il NWA prevede un forum delle biblioteche nazionali nordiche per il coordinamento e lo scambio di esperienze sull'harvesting e l'archiviazione dei documenti Web; nel 2000 la cooperazione del NWA (sovvenzionata da Nordunet2, NORDINFO e le Nordic National Libraries) ha cominciato a sviluppare il NWA Toolset, un pacchetto di programmi per l'accesso ai documenti Web archiviati. Dal 2003 NWA ha iniziato la sua collaborazione con l'International Internet Preservation Consortium (IIPC).<sup>112</sup>

Archivi delle risorse digitali esistono in tutti i Paesi partecipanti, che hanno trasferito i propri contenuti di Web archiving al NWA centrale. Nei paesi del NWA l'approccio alla conservazione dei siti Web è stato tendenzialmente di tipo esaustivo, una scelta basata su considerazioni di economicità, ma anche di imprevedibilità rispetto ai futuri interessi informativi del pubblico sul patrimonio di/in Internet nel tempo. Il progetto prevedeva dal suo inizio di creare strumenti per l'analisi dello sviluppo storico dei siti Web ed escludeva la possibilità di catalogare in modo tradizionale le risorse Internet; anche per i sistemi di accesso agli archivi, basati sui metadati, si è rimandato a quando potessero essere disponibili metodi, sia per la generazione dei metadati, sia per la catalogazione, di tipo automatici.

Un primo pacchetto di programmi NWA (NWAToolset) è stato sviluppato usando l'esperienza portata da ogni biblioteca associata, per permettere la navigazione e la ricerca del contenuto archiviato, sia nel tempo, sia nello spazio. La seconda fase del progetto di NWA è cominciata verso la fine del 2002: era concentrata sulle ricerche per la stabilità del software, sulle esigenze di funzionalità supplementari da parte dell'utente e sul miglioramento dell'interfaccia dell'utente.

Inizialmente, non essendo il software ancora disponibile per uso su larga scala, il progetto ha adottato un file descrittivo (un file "multi-part MIME") associabile al documento elettronico descritto e diviso in tre parti:

---

<sup>111</sup> <<http://nwa.nb.no/>>.

<sup>112</sup> <<http://www.netpreserve.org/>>.

- a - informazioni sulle modalità di raccolta,
- b - dati forniti dal Webserver,
- c - informazioni sul contenuto della risorsa.

Dal gennaio 2006 ha avuto avvio WERA (Web aRchive Access), un'applicazione di visualizzazione dell'archivio che fornisce un accesso all'archivio delle collezioni Web, del tipo della Wayback Machine di Internet Archive, con la possibilità di fare anche una ricerca per full text e con una facile navigazione tra le diverse versioni di una pagina Web. WERA sostituisce NwaToolset, sul quale, comunque, si è basato; è stato costruito usando PHP e Java ed utilizza standard aperti (come il protocollo HTTP e l'XML). WERA è una soluzione liberamente disponibile per la ricerca e la navigazione delle collezioni Web archiviate. La ricerca nel contenuto dei documenti è possibile grazie al motore di ricerca Search Engine di NutchWAX.

L'archivio Web, reso direttamente disponibile per questo servizio di ricerca, è stato implementato con l'harvester di Heritrix; i documenti Web archiviati sono memorizzati nei files di ARC e WERA.

È stata installata una demo come servizio per dimostrare la funzionalità di WERA (v. 0.4.1) e NutchWAX (v. 0.4.2) al pubblico e come banco di prova dall'IIPC, da cui è stato patrocinato lo sviluppo di WERA e sul cui sito si vede l'ultima release, oltre a poter accedere alla documentazione informativa, con una descrizione sull'architettura di base, le funzionalità, i requisiti per l'installazione e il manuale per le istruzioni d'installazione di WERA, tutto scaricabile dal sito Web.<sup>113</sup>

Le biblioteche nazionali di Svezia, Finlandia, Norvegia, Danimarca e Islanda portano nel consorzio l'esperienza dell'harvesting dello spazio Web nazionale del Nordic Web Archive, infatti i singoli Paesi aderenti al progetto NWA hanno avviato da alcuni anni propri archivi di raccolta, cui qui di seguito si dà una breve presentazione (eccezion fatta per la Norvegia, cui viene dedicato il cap. 2.4).

In Svezia, The Royal and National Library of Sweden (Kungliga Biblioteket), che dal 1661 ha il compito di raccogliere la produzione nazionale svedese e oggi anche le pubblicazioni in formato digitale, ha inaugurato nel 1996 il KULTURARW<sup>3</sup> project (The Royal Swedish Archiw<sup>3</sup>e).<sup>114</sup>

---

<sup>113</sup> <<http://archive-access.sourceforge.net/projects/wera/>>. Per la demo: <<http://nwa.nb.no/wera/>>.

<sup>114</sup> The Kulturaw<sup>3</sup> Project. URL: <<http://kulturaw3.kb.se/>>. Si vedano: Johan Mannerheim. *The WWW and our digital heritage - the new preservation tasks of the library community*, Allan Arvidson, Krister Persson. *The Kulturaw<sup>3</sup> Project - The Royal Swedish Web Archiw<sup>3</sup>e - An example of "complete" collection of web pages*. Disponibili all'URL: <<http://www.ifla.org/IV/ifla66/papers/158-157e.htm>> e <<http://www.ifla.org/IV/ifla66/papers/154-157e.htm>>.



Il KULTURARW3 project già nella primavera del 1997, secondo le stime di due salvataggi annuali, aveva archiviato circa 70 milioni di documenti dei siti svedesi, 1500GB, metà dei quali testuali: periodici e giornali, pagine HTML in aggiornamento, archivi FTP e basi dati (con l'eccezione dei gruppi di discussione). Il sito riporta tre sezioni cronologicamente distinte degli archivi consultabili in linea e una statistica dettagliata delle tipologie dei file. Il sistema informativo prevedeva dall'inizio solo la ricerca per URL nella collezione, che comprendeva il patrimonio digitalizzato, i siti Web e i documenti svedesi in Internet, e solo all'interno della rete della Royal Library, dalle postazioni della biblioteca.

In Svezia è stato implementato anche DiVa (Digitala Vetenskapliga Arkivet in Swedish) – l'Archivio svedese sviluppato dal 2000 per le pubblicazioni on-line di ambito accademico e scientifico, delle Università di Svezia, Norvegia e Danimarca.<sup>115</sup>

In Danimarca, la Royal Library's, The State and University Library si sono concentrate su una raccolta con harvesting e cooperano per costituire un archivio unico, collaborando a livello di strategie e di software, al fine di raccogliere, archiviare, preservare e garantire accesso alla collezione digitale: dal 2001 hanno avviato l'“Internet archiving project Netarkivet.dk”,<sup>116</sup> per un archivio nazionale pubblico, legato anche ad una nuova legge per il deposito legale, che contiene anche progetti di raccolte basate su selezioni per eventi e altre finalità di documentazione. Si ricorda, inoltre, che nel catalogo nazionale on line della *Bibliotek.dk* sono presenti documenti Internet ed è possibile la ricerca per i “Net documents”, per parole del testo, autore o titolo, titolo del sito, soggetto, lingua, anno di pubblicazione.

In Finlandia l'harvesting viene effettuato dal 1999 dalla Biblioteca Nazionale ed è attivo il progetto cooperativo EVA<sup>117</sup> (per le pubblicazioni scientifiche ed accademiche in formato digitale), di biblioteche, editori, esperti, voluto dal Ministero dell'Educazione all'interno di un programma strategico per la formazione e l'istruzione, il cui scopo centrale è creare metodi, strumenti, registrazioni ed archivi in Internet e cercare condizioni per la preservazione a lungo termine nelle biblioteche.

---

<sup>115</sup> < <http://epc.uu.se/diva-app-profile/mainpage.jsp> >. Eva Müller - Uwe Klosa - Stefan Andersson - Peter Hansson. *The DiVA Project - Development of an Electronic Publishing System*. “D-Lib Magazine”. November 2003. URL: <<http://www.dlib.org/dlib/november03/muller/11muller.html>>.

<sup>116</sup> <[www.netarkivet.dk](http://www.netarkivet.dk)>. Alcune fonti: Christensen-Dalsgaard, Birte. *Web Archive Activities in Denmark*. “RLG DigiNews”, 8, 3 (June 15 2004), Department of Research, Cornell University Library. URL:<[http://www.rlg.org/en/page.php?Page\\_ID=17661](http://www.rlg.org/en/page.php?Page_ID=17661)"\l "article0">. Finnemann, Niels Ole. *Netarkivet.dk Internet — a cultural heritage of our time*. in *Preserving the Present for the Future. Conference on strategies for the Internet. Proceedings*. Copenhagen: The Royal Library/Denmark's Electronic Research Library, 2001. P. 31-42. URL: <[http://www.deflink.dk/upload/doc\\_filer/doc\\_alle/846\\_Trykt\\_proceeding.Pdf](http://www.deflink.dk/upload/doc_filer/doc_alle/846_Trykt_proceeding.Pdf)>. Finnemann, Niels Ole. *Arkivering af internettet*. “Kritik”, 34, 151 (2001), Copenhagen. P. 13-18.

<sup>117</sup> <<http://www.lib.helsinki.fi/eva/english.html>>; Lounamaa, Kirsti - Salonhariju, Inkeri. *EVA: the acquisition and archiving of electronic network publications in Finland*. <<http://hul.helsinki.fi/tictolinja/0199/evaart.html>>.

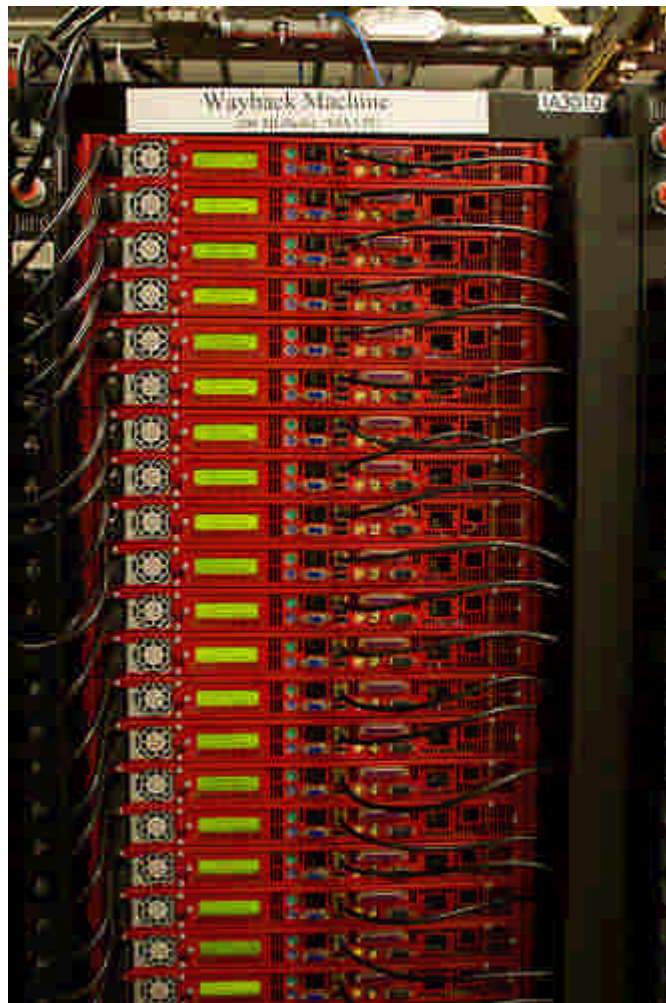
EVA coopera con altri progetti gestiti dalla Helsinki University Library, quali, per esempio, ELEKTRA,<sup>118</sup> presso le Oulu and Jyväskylä University Libraries, per lo spoglio della letteratura scientifica periodica, per il deposito di tesi e altre pubblicazioni elettroniche.

Per quanto riguarda l'Islanda si segnala il VESTNORD project<sup>119</sup> con digitalizzazioni di periodici e stampa pubblicati tra il 1773 e il 2001, progetto cooperativo tra le biblioteche nazionali Faeroe Islands, Greenland and Iceland.

---

### **Large Scale Data Repository: Petabox di INTERNET ARCHIVE.**

Un'immagine del PetaBox(tm), progettato dallo staff di Internet Archive per memorizzare in sicurezza un petabyte (milion di gigabytes) d'informazioni.<sup>120</sup>



<sup>118</sup> <<http://www.lib.helsinki.fi/elekra/english.html>>.

<sup>119</sup> <<http://www.timarit.is/umverkefni.jsp?lang=4>>.

<sup>120</sup> L'immagine è stata ripresa dalla pagina di Internet Archive <<http://www.archive.org/web/petabox.php>>.

### 2.1.3. In Gran Bretagna: CEDARS, UKWAC, UK Government Web Archive

Con queste brevissime descrizioni sono presentati tre progetti cooperativi in ambito britannico che coprono settori della produzione digitale: l'ambito più propriamente accademico, i siti Web britannici e la documentazione amministrativa (l'e-gouvernement); per finire si ricordano iniziative come la Digital Preservation Coalition, volta alla ricerca, anche con collaborazioni internazionali e il Digital Curation Centre.

CEDARS (Exemplars in Digital Archives):<sup>121</sup> come iniziativa eLib e sotto l'egida del CURL (Consortium of University Research Libraries), il Joint Information Systems Committee britannico ha avviato il progetto CEDARS nel 1998, progetto che fa capo alle tre università di Oxford, Leeds e Cambridge, con l'obiettivo di individuare alcune questioni strategiche, metodologiche e pratiche relative alla preservazione dei documenti elettronici, e fornire, quindi, una serie di linee guida per le biblioteche, attraverso le occasioni fornite da focus group, seminari e liste di discussione. CEDARS ha recepito le indicazioni contenute nel modello OAIS, adattandole e ampliandole con riferimento alla strategia della conversione dei formati e preferendola a quella dell'emulazione.

Per iniziativa del National Archives britannico, l'amministrazione archivistica inglese, nel giugno 2004 è partito un progetto pilota della durata di due anni al quale hanno aderito la British Library, il Wellcome Trust, l'Higher Education Funding Council of England, la National Library of Scotland, e la National Library of Wales, il Joint Information Systems Committee (JISC), tutti riuniti nel UK Web Archive Consortium per l'archiviazione, il progetto UKWAC.<sup>122</sup> Il progetto si è mosso dalla constatazione che il Web è diventato il luogo ove si trovano tutte le notizie e le informazioni che riguardano la vita culturale del paese, dalla ricerca scientifica, all'arte, allo sport, ma, nonostante un uso diffuso del mezzo Internet, si è molto trascurato l'aspetto della conservazione dei contenuti, aspetto che in breve tempo potrebbe, invece, diventare una vera emergenza. Questa iniziativa vuole ovviare a questo pericolo di perdita delle conoscenze contemporanee, che spesso ormai hanno sul Web il loro principale contenitore, costituendo un consorzio di sei istituzioni guida per lo sviluppo di un archivio selettivo dei siti Web britannici. Il consorzio britannico parte da una base già testata come il sistema PANDORA, sviluppato dalla National Library of Australia e da questa base intende intraprendere propri ulteriori sviluppi tecnici; utilizza il software e il sistema di archiviazione digitale di PANDORA, PANDAS (PANDORA Digital Archiving

---

<sup>121</sup> <<http://www.leeds.ac.uk/cedars/>>.

<sup>122</sup> <<http://www.webarchive.org.uk/>>.

System) e per l'harvesting HTTrack, già sperimentato appunto per PANDORA. Anche sui principi e le modalità di archiviazione si è seguito quanto stabilito nell'esperienza australiana di PANDORA. La forma del consorzio tra istituzioni consente la condivisione di costi, esperienze, selezione mirata dei siti, facilita l'archiviazione per un accesso futuro e la distribuzione ottimale dei carichi, tale da creare un ambiente operativo dei processi d'archiviazione Web; tra i membri vengono ripartite anche le incombenze tecnologiche, l'infrastruttura, la raccolta e il deposito, la ricerca di standard e tecniche migliorative. Il consorzio è una struttura organizzativa basata sulla condivisione degli strumenti hardware e software, ma anche degli altri aspetti: ogni membro del consorzio seleziona e cattura determinati siti, occupandosi del copyright, delle relative autorizzazioni e licenze dei possessori dei diritti, editori, autori, distributori etc., sul materiale di propria competenza. Per esempio, la biblioteca nazionale del Galles raccoglie i siti che riflettono la vita sociale nel Galles contemporaneo, mentre la British Library punta l'attenzione sui siti d'importanza culturale, storica e politica. Una volta che i siti sono stati archiviati, catalogati e controllati per vedere se c'è completezza, saranno resi direttamente accessibili dal sito Web appositamente allestito. La ricerca può essere fatta per soggetto, URL o titolo e l'accesso è possibile anche a partire dai singoli cataloghi delle biblioteche partecipanti al consorzio.

Oltre al UKWAC, il National Archives ha avvertito la responsabilità di raccogliere e conservare i siti Web dell'amministrazione governativa, quali fonti per la storia e come testimonianza del cambiamento nella relazioni tra Stato e cittadino. Anche per questo progetto si è partiti dal calcolo che in Inghilterra oltre la metà di tutte le relazioni tra lo Stato e i cittadini avviene ormai on-line e che attraverso le pagine Web istituzionali lo Stato informa, lasciando tracce di sé e della propria attività nelle risorse elettroniche remote. Da questa consapevolezza nasce l'UK Government Web Archive<sup>123</sup> che, per il periodo iniziale di sperimentazione, con un contratto di harvesting a partire dal giugno 2003, si serve di Internet Archive (v. paragrafo 2.1.1.) e del suo Web crawler in una versione appositamente modificata; la scelta di Internet Archive è maturata dalla considerazione che si tratta di uno dei servizi di archiviazione Web più potenti nel mondo. Già nel 2004 i siti archiviati erano 50 (80 nel marzo 2007), tutti appartenenti ad organi istituzionali, quali i Dipartimenti, le Agenzie esecutive, etc.; si tratta delle pagine che meglio rappresentano l'intero dominio Web dello Stato. I siti sono raccolti usando il crawler software che recupera e registra ogni pagina del sito (per questo si rimanda al paragrafo su Internet Archive), programmato ogni 6 mesi per alcuni siti, settimanalmente per altri, quotidianamente per altri ancora; i siti vengono

---

<sup>123</sup> <<http://www.nationalarchives.gov.uk/preservation/webarchive/>>.

archiviati e conservati in un particolare formato sviluppato da Internet Archive al quale si può accedere grazie alla Wayback Machine. La ricerca avviene per URL: gli utenti inseriscono l'indirizzo del sito che desiderano vedere e poi selezionano attraverso la data la versione disponibile in archivio. Si stanno cercando ulteriori approcci tecnici per la conservazione dei siti Web e il sistema di Internet Archive verrà sottoposto a valutazione al fine di sviluppare una propria strategia a lungo termine. Ultimamente l'UK Government Web Archive coopera anche con l'European Archive,<sup>124</sup> l'iniziativa affiliata di Internet Archive dal cui portale si possono visualizzare queste collezioni, con raccolta settimanale e mensile.

La Digital Preservation Coalition (DPC)<sup>125</sup> nasce nel 2001 sotto l'urgenza di garantire la preservazione delle risorse digitali nel Regno Unito e di collaborare con le altre agenzie internazionali sulla sicurezza globale della memoria digitale attraverso una base di condivisione delle conoscenze; la sua azione si esplica apportando il proprio bagaglio informativo e di competenze in ogni progetto britannico del settore. La DPC ha un suo sito Web con due accessi, Public e Members, quest'ultimo riservato ai membri dello staff e i partecipanti all'organizzazione. Le pagine Web danno conto del background della DPC, con informazioni sulla mission, i programmi, l'amministrazione e i partecipanti.

Nel 2004 il JICS<sup>126</sup> ha istituito il Digital Curation Centre (DCC)<sup>127</sup> come fondamento per i programmi di patrimonio informativo scientifico elettronico. Lo scopo è quello di focalizzare ricerca e sviluppi della preservazione sulla questione della digital curation, che già nel 2005 si è espressa in un *Digital Curation Manual* redatto da esperti per costituire una risorsa di determinazione comune e fornire dati utili ad autori, curatori e operatori che devono affrontare nei vari ruoli la sfida della preservazione dell'informazione digitale nel tempo. Tra l'altro, ha condotto un'analisi sui meccanismi di audit e certificazione dei depositi digitali britannici, a partire dalla valutazione critica a fini applicativi dell'*Audit checklist for certifying digital repositories* elaborata da RLG-NARA e col specifico tentativo di misurare e testimoniare la qualità degli attributi previsti.<sup>128</sup>

---

<sup>124</sup> <<http://www.europarchive.org/>>.

<sup>125</sup> <<http://www.dpconline.org/graphics/>>.

<sup>126</sup> Si ricorda che il Joint Information Systems Committee (JISC) è volto a fornire una leadership di livello internazionale per l'uso innovativo dell'informazione e della tecnologia delle comunicazioni a supporto dell'educazione e della ricerca; finanzia un portfolio di servizi nazionali e una vasta gamma di programmi e progetti. URL: <[http://www.jisc.ac.uk/publications/publications/pub\\_digipreservationbp.aspx](http://www.jisc.ac.uk/publications/publications/pub_digipreservationbp.aspx)>.

<sup>127</sup> <<http://www.ac.uk/about/what/>>, partners sono: University of Edinburgh, HATII at the University of Glasgow, UKOLN at University of Bath, Council for the Central Laboratory of the Research Councils (CLRC).

<sup>128</sup> Seamus Ross – Andrew Mchugh. *The role of evidence in establishing trust in repositories*. "Archivi & Computer" 1 (2006). P. 46-59.

#### 2.1.4. International Internet Preservation Consortium (IIPC)

È un Consorzio internazionale per la preservazione di Internet ufficialmente sottoscritto nel luglio 2003 alla Bibliothèque Nationale de France, che ne è coordinatrice, da dodici istituzioni partecipanti che hanno raggiunto un accordo finanziario e di collaborazione per progetti e gruppi di lavoro al fine di conseguire gli scopi che il consorzio si prefigge.

Mission dell'International Internet Preservation Consortium<sup>129</sup> è acquisire, preservare e mantenere accessibile per le future generazioni, la conoscenza e l'informazione veicolata da Internet, universalmente, promuovendo lo scambio globale e le relazioni internazionali. Fanno parte del Consorzio: le biblioteche nazionali di Svezia, Finlandia, Norvegia, Danimarca e Islanda (che portano nel consorzio l'esperienza dell'harvesting dello spazio Web nazionale del Nordic Web Archive), Francia, Australia, Canada, la Biblioteca Nazionale Centrale di Firenze, la Library of Congress, la British Library, Internet Archive.

Gli scopi sono:

- ≈ rendere possibile la preservazione della ricchezza dei contenuti di Internet provenienti da tutto il mondo, intendendo la preservazione come salvaguardia e possibilità di accesso nel tempo;
- ≈ favorire lo sviluppo e l'uso di strumenti, tecniche e standard condivisi per la creazione di archivi internazionali;
- ≈ sostenere le biblioteche nazionali che intendono occuparsi dell'archiviazione e preservazione di Internet.

Azioni:<sup>130</sup>

- ≈ forum per lo scambio di conoscenza sull'archiviazione di Internet;
- ≈ sviluppo di standards e raccomandazioni;
- ≈ sviluppo di strumenti interoperabili e tecniche che permettano acquisizione, archiviazione e accessibilità ai siti Web;
- ≈ stimolare la consapevolezza dei problemi della preservazione di Internet anche tramite conferenze e iniziative diverse, workshops, training events, publications, etc.

---

<sup>129</sup> <<http://netpreserve.org/>>.

<sup>130</sup> IIPC International Internet Preservation Consortium. *Press Release*, May 5 (2004).

Uno degli obiettivi principali dell'IIPC è sviluppare un pacchetto software di alto livello qualitativo, facile all'uso ed open source, per creare una "catena di archiviazione", un processo di Web archiving efficace e alla portata universale.

I gruppi di lavoro specifici hanno focalizzato le necessità relative ai problemi di deep Web, accessibilità, content management, producendo, al 2006, applicazioni scaricabili a partire dalle pagine del sito dell'IIPC:

<i>Problematiche</i>	<i>Esigenze</i>	<i>Strumenti trovati al 2006</i>
Acquisizione	Archiviazione di qualità per un harvesting graduabile, nella raccolta al livello di dominio	Heritrix, Web crawl estensibile, open source, Web-scale. Sviluppato da Internet Archive e NWA. Deep Arc, editor grafico, usabile per mappe, data model relazionali, XML Schema. Sviluppato dalla B. N. de France.
Selezione e verifiche	Strumenti che permettano di programmare il crawl in modo analitico e per priorità. Un gestore che offra un'interfaccia semplice per la selezione, il monitoraggio e la verifica dei siti Web archiviati	Web Curator Tool (WCT), Apache Public License, sviluppato da National Library of New Zealand e la British Library.
Immagazzinamento e manutenzione della collezione	Strumenti adeguati per la manutenzione	BAT (BnFArcTools), processore API per files ARC, DAT o CDX. Sviluppato dalla B. N. de France.
Strumenti per l'accesso e la ricerca	Strumenti per la visualizzazione degli archivi Web, con possibilità di navigazione temporale, indicizzazione; interfaccia per la ricerca nella banca dati	NutchWAX (Nutch with Web Archive eXtensions), sviluppato da Internet Archive e NWA. WERA (Web aRchive Access), sviluppato da Internet Archive e la Nat. Library of Norway. Xinq (XML INquire), sviluppato dalla Nat. Library of Australia.

## 2.2. Iniziative e progetti nell'Unione Europea.

Nel presentare le iniziative cardine sulla digital preservation in Europa, non si può prescindere dal sottolineare l'azione che l'Unione Europea ha svolto e continua a svolgere nello sviluppo di un unico, variegato ma organico, sistema di produzione e conservazione digitale, sempre ricercando l'attuazione di un approccio congiunto, con lo scambio di esperienze su politiche, programmi, regolamentazioni, esame delle strutture, valutazione dei risultati. L'azione politica dell'Unione si è focalizzata proprio sulla progressiva 'digitalizzazione' dei servizi, del patrimonio e della produzione culturale, dando origine e sostenendo la maggior parte dell'attuale realtà del patrimonio culturale digitale europeo.

Il Consiglio d'Europa ha guardato allo sviluppo del mondo digitale dal punto di vista politico e con lo scopo di promuovere la proliferazione di prodotti digitali; l'interesse verso il digitale è stato di tipo strumentale, considerandolo tramite di sviluppo economico e industriale, ma soprattutto veicolo di promozione, di consapevolezza e cittadinanza nella costruzione dell'identità europea: ciò in quanto le tecnologie digitali rappresentano il più promettente mezzo di comunicazione, trasmissione dell'informazione, nonché di promozione e conservazione del patrimonio culturale. Ovviamente, ben presto i prodotti digitali hanno indotto all'interessamento della loro conservazione nel tempo.

Negli ultimi vent'anni gli interventi europei per lo sviluppo dei servizi e l'impatto tecnologico e organizzativo hanno avuto come focus le biblioteche, in quanto ad esse è stato assegnato un ruolo decisivo per lo sviluppo della società dell'informazione in Europa, quali mediatori per la diffusione della conoscenza e della cultura e per la conservazione del patrimonio europeo.<sup>131</sup> Per avviare l'utilizzazione diffusa delle tecnologie elettroniche iniziano i finanziamenti di programmi specifici per la ricerca e lo sviluppo tecnologico e telematico nelle biblioteche (*Programmes for Research and Technological Development, Telematics for Libraries*, Framework 1-5, dal 1985 al 1998) per la costruzione dell'IST; si assiste così all'introduzione di infrastrutture, tecnologie informatiche, standard, interconnessione dei sistemi e reti per la condivisione delle risorse, nuovi servizi e prodotti basati sulle Information Technologies (IT); dai primi anni Novanta cominciano i progetti di digitalizzazione di libri antichi e/o rari, per la creazione di banche dati accessibili tramite

---

<sup>131</sup> Ris. 27 settembre 1985 del Consiglio in cui si raccomandava una stretta collaborazione tra le biblioteche e il settore dell'informatica e invitava la Commissione a porre in essere, in stretta collaborazione con le principali biblioteche della CEE, azioni coordinate per favorire lo sviluppo dell'informatizzazione.



Internet (è del 1994 uno dei primi progetti finalizzati alla creazione di banche dati di opere accessibili tramite Internet, *Deborà*, dedicato alla digitalizzazione di libri rinascimentali).<sup>132</sup>

Con le *Linee guida del Consiglio d'Europa e di Eblida sulla legislazione bibliotecaria in Europa*<sup>133</sup> si presenta il documento che più compiutamente compendia la figura della biblioteca nella società contemporanea, con suggerimenti tecnici e proposte operative. Si afferma che le biblioteche sono elemento essenziale e insostituibile dell'infrastruttura culturale, educativa e informativa della società, e che tra i loro compiti c'è

«...lo sviluppo e la conservazione delle collezioni nazionali, in qualunque forma e di qualsiasi contenuto».<sup>134</sup>

Una serie di Piani di azione e risoluzioni sono alla base dell'attuale situazione di programmi e progetti di sviluppo e ricerca specifici sul digitale:

≈ *eEurope – An information Society for all* (C.E. 8/12/1999)<sup>135</sup>: iniziativa finalizzata ad accelerare l'adozione delle tecnologie digitali in tutta l'Europa e a garantire la massima diffusione delle conoscenze e delle competenze necessarie per l'impiego di tali tecnologie. I piani eEurope avranno diverse riedizioni tra il 2000 e il 2005.

≈ Ris. del Consiglio, d. 21-01-2001, *Cultura nella società della conoscenza* : sostiene «... la cooperazione e lo scambio di informazioni e buone pratiche in Europa, per l'accessibilità dei contenuti e la qualità dei siti web culturali, su cultura e società della conoscenza».<sup>136</sup>

≈ Ris. del Consiglio, d. 21-01-2002, *Ruolo della cultura nello sviluppo dell'Unione europea*: invita alla digitalizzazione dei programmi di contenuto culturale con l'interoperabilità dei relativi sistemi e a sostenere le organizzazioni preposte alla conservazione nel raccogliere e mantenere disponibili nel tempo i contenuti digitali.<sup>137</sup>

≈ Ris. del Consiglio, d. 25-06-2002, *Conservazione della memoria del domani - conservazione dei contenuti digitali per le generazioni future*, che ribadirà la necessità di monitorare le iniziative di conservazione realizzate in Europa e le cooperazioni risultate efficaci per scambi di esperienze su politiche,

---

<sup>132</sup> < <http://cordis.europa.eu/libraries/> >; < <http://cordis.europa.eu/ist/digicult/index.html> >.

<sup>133</sup> Adottate dalla Commissione Cultura del Parlamento Europeo (ottobre 1999) e Eblida (European Bureau of Library Information and documentation Associations), sulla scorta delle indicazioni discusse alla conferenza "Libraries and democracy: the responsibility of the State, local authorities and professionals", Strasburgo, 23-25 novembre 1998. Commissione Cultura del Parlamento Europeo (19a seduta, 12-14 ottobre 1999), Council for Cultural Co-operation (68a seduta, 19-21 gennaio 2000). "Digital content, Books and Archives". Strasburgo, 20-01/21-02-2000 (DECS/CULT/POL/book(2000)).

<sup>134</sup> Per ogni ulteriore approfondimento sul documento si rimanda all'articolo *Intervista a Giuseppe Vitiello*. "AIB Notizie", 12 (2000), n. 9.

<sup>135</sup> Informazioni si possono trovare anche all'URL < <http://www.w3.org/WAI/References/eEurope> > e < <http://www.e-eurostandards.org/> >.

<sup>136</sup> G.U.C.E. n. C 32/01.

<sup>137</sup> G.U.C.E. n. C 32/02.

programmi e problematiche, di sostenere le istituzioni deputate alla conservazione, con infrastrutture organizzative (Reti) e norme tecniche necessarie a creare e mantenere nel tempo le raccolte digitali (anticipando la *Carta per la conservazione del patrimonio digitale* dell'UNESCO, del 2003). Con questa risoluzione l'U. E. ha individuato nella conservazione di lungo periodo una delle linee d'azione prioritarie: l'importanza di giungere alla consapevolezza degli operatori, la centralità del ruolo delle istituzioni della memoria, l'urgenza per il pericolo della perdita di questo patrimonio, la necessità della pianificazione delle attività, la centralità della formazione, gli aspetti legati al copyright e ai metadati, il bisogno di creare strumenti in grado di identificare i costi della conservazione.

Queste risoluzioni hanno avuto come conseguenze dirette due iniziative fondamentali per l'odierno sviluppo della salvaguardia del patrimonio culturale europeo digitale: si tratta dei *Principi di Lund*, del 4-4-2001, che evidenziano l'impegno sul tema della conservazione nel lungo periodo del patrimonio culturale digitale e della sostenibilità tecnica con i software aperti, e della costituzione di MINERVA e-Europe (*MINisterial NETwork for Valorising Activities in digitisation*),<sup>138</sup> che nasce il 1 marzo 2002.

Il progetto Minerva, al cui interno è stato istituito un gruppo di lavoro sulla memoria digitale, è volto a creare un network di ministeri della cultura europea per analizzare, raffrontare e armonizzare le attività intraprese nel campo della digitalizzazione dei contenuti scientifici e culturali, fornire linee di indirizzo con particolare riferimento all'interoperabilità, alla qualità dei siti web, ai metadati, ai repertori per i fondi digitalizzati e alla preservazione. Hanno fatto seguito i progetti delle 'biblioteche del digitale' europeo: Michael (Multilingual Inventory of Cultural Heritage in Europe)<sup>139</sup> e la EDL (European Digital Library)<sup>140</sup> che dal 2008 convoglierà in TEL (The European Library).<sup>141</sup>

Si ricordano anche, dal 2001, la promozione del Document Lifecycle Management (DLM), che con un Forum<sup>142</sup> periodico affronta il problema della preservazione e dell'accessibilità nel tempo, manifesta l'intento di trovare buone pratiche e soluzioni

---

<sup>138</sup> <<http://www.minervaeurope.org/>>: avviato il 1 marzo 2002 per creare un network dei Ministeri della cultura degli Stati membri, per armonizzare le attività di digitalizzazione dei contenuti culturali e scientifici, per creare una piattaforma europea condivisa (e multilingue) attraverso l'elaborazione di linee guida, raccomandazioni, manuali e strumenti sulla digitalizzazione, sui metadati, sull'accessibilità a lungo termine e sulla conservazione, con un WP apposito per la realizzazione del piano di lavoro collegato alla Ris. C 162/02, sulla conservazione a lungo termine della memoria digitale.

<sup>139</sup> <<http://www.michael-culture.org/>>: progetto volto alla definizione di un approccio comune e di un modello di servizi per la fruizione del patrimonio culturale europeo e la creazione di un portale di accesso comune online, in modalità multilingue, per la consultazione e la ricerca dei portali culturali nazionali; inizialmente ha adottato la piattaforma tecnologica del progetto francese per il *Catalogues des fonds culturels numérisés*.

<sup>140</sup> European Digital Library project, sett. 2006-dic.2007, <<http://www.edlproject.eu/about.php>>.

<sup>141</sup> TEL (The European Library), <<http://www.theeuropeanlibrary.org/>>, progetto europeo che costituisce il punto d'arrivo dei progetti TEL-ME-MOR (The European Library: Modular Extensions for Mediating Online Resources, febr.2005-genn.2007) e EDL (European Digital Library project, sett. 2006-dic.2007): è l'infrastruttura multilingue che permette attualmente di consultare i cataloghi online di diverse biblioteche nazionali e di accedere anche ad una serie di risorse digitalizzate delle biblioteche aderenti.

<sup>142</sup> <<http://dlmforum.typepad.com/>>.

tecniche, promuove la cooperazione fra le pubbliche amministrazioni e l'industria dell'informazione e della comunicazione; nonché l'iniziativa eTERM (European Training programme in Electronic Records Management for administrators and archivists)<sup>143</sup> che ha lo scopo di individuare un percorso formativo rivolto agli operatori del settore.

≈ Comunicazione "i2010 - digital libraries" (30-09-2005): afferma che le biblioteche digitali fanno parte del nuovo Piano d'azione "i2010 - a European Information Society for growth and employment", che dal 01-06-2005 sostituisce *e-Europe* e che nella sua azione per favorire l'inclusione e il miglioramento della qualità della vita dei cittadini, avvia iniziative per la preservazione del patrimonio digitale, estremamente importante per la salvaguardia della cultura europea, ma vitale nel settore dei dati scientifici, con implicazioni di sicurezza per le grandi tematiche che riguardano la società e la vita del pianeta, essendo la volatilità dell'informazione digitale fonte di rischio per la conservazione di percorsi e traguardi delle conoscenze e dei contenuti stessi, che devono essere accessibili e utilizzabili nel futuro.<sup>144</sup>

≈ Raccomandazione della Commissione Europea, d. 24-08-2006, *Digitisation and online accessibility of cultural material and digital preservation*.<sup>145</sup> le digital libraries («... - Our Digital Memory») vanno garantite sotto i profili dell'accessibilità dei contenuti, dei diritti, della conservazione a lungo termine.

Di seguito vengono presentate le iniziative fondamentali, condotte a partire dal 1997, con il compito di raccordo e di coordinamento per la ricerca, le strategie e i diversi progetti sulla digital preservation.

---

<sup>143</sup> <<http://www.ucl.ac.uk/e-term/index.htm>>.

<sup>144</sup> < [http://ec.europa.eu/information\\_society/activities/digital\\_libraries/index\\_en.htm](http://ec.europa.eu/information_society/activities/digital_libraries/index_en.htm) >;

<[http://ec.europa.eu/information\\_society/eeurope/i2010/index\\_en.htm](http://ec.europa.eu/information_society/eeurope/i2010/index_en.htm) >. Sulle problematiche d'accesso sulla diffusione e preservazione dell'informazione scientifica: Conferenza "Pubblicazione scientifica nello Spazio europeo della ricerca: accesso, divulgazione e conservazione nell'era digitale", Bruxelles 15/16-02-2007; con lo studio dei servizi di stoccaggio ("Towards a European eInfrastructure for eScience digital repositories") Digital Repositories for eScience. Al proposito si citano i progetti affini: CASPAR (per la gestione dell'accesso futuro e la preservazione dei dati scientifici, di cui si parlerà in una nota del Cap. 2.2.3), ORIEL (*Online Research Information Environment for the Life Sciences*, <<http://www.oriel.org/aboutpartners.html>>), DRYVER (Digital Repository Infrastructure Vision for European Research, ricerca come connettere i repertori informativi scientifici, < <http://www.driver-repository.eu/>>, <<http://www.sherpa.ac.uk/projects/driver.html>>), SEADATANET (volto a sviluppare un'infrastruttura paneuropea di gestione di dati marittimi al fine d'integrare i repertori nazionali, <<http://www.seadatanet.org/>>), EURO-VO-DCA (European Virtual Observatory Data Centre Alliance, <<http://www.euro-vo.org/>>), BELIEF (Bringing Europe's eElectronic Infrastructures to Expanding Frontiers , <<http://www.beliefproject.org/>>), TNT (The Neanderthal Tools, <<http://www.the-neanderthal-tools.org/>>).

<sup>145</sup> < [http://ec.europa.eu/information\\_society/newsroom/cf/itemlongdetail.cfm?item\\_id=2782](http://ec.europa.eu/information_society/newsroom/cf/itemlongdetail.cfm?item_id=2782) >.

### 2.2.1. Networked European Deposit Library (NEDLIB)

NEDLIB (Networked European Deposit Library)<sup>146</sup>, promossa dalla Conference of European National Libraries e finanziata dalla Commissione europea, è stata un'iniziativa nata nel 1997 e terminata nel 2000, cui hanno partecipato:

la Koninklijke Bibliotheek Agentschap Rijksarchiefdienst (con il ruolo di coordinatore), la Bibliothèque Nationale de France, la National Library of Norway, sede di Mo I Rana, la Helsinki University Library, Die Deutsche Bibliothek, la Biblioteca Nacional de Portugal, la National Library of Switzerland, la Biblioteca Nazionale Centrale di Firenze, l'Instituto De Engenharia de Sistemas e Computadores, l'INESC CSC Ploenzke AG (Germany). Alle biblioteche nazionali si sono affiancati alcuni editori (Springer, Elsevier, Kluwer) in un'ottica di collaborazione nella definizione di standard, metodologie e difesa degli interessi commerciali legati al copyright attraverso un'implementazione nel controllo degli accessi.

L'obiettivo del Network era quello di progettare un'infrastruttura di base su cui edificare una sorta di "sistema" per garantire il deposito delle pubblicazioni elettroniche e il loro accesso nel tempo. Tecnicamente, si è progettato che il meccanismo di deposito dovesse basarsi su un sistema di archiviazione di informazione secondo il modello OAIS (Open Archival Information System).

Il progetto NEDLIB ha studiato le procedure di livello superiore di un sistema di deposito per le pubblicazioni elettroniche, proponendosi di seguire il modello e la terminologia OAIS per la gestione e la preservazione dei dati; così descrive i suoi principi guida:

≈ Immissione, attuata attraverso una struttura d'interfaccia (che fornisce e cattura una pubblicazione elettronica) e la sua archiviazione informatica;

≈ Archiviazione, che consiste nel recupero e nell'immagazzinamento della pubblicazione elettronica, salvaguardandone l'integrità;

≈ Gestione dei dati, in cui si procede al recupero e all'immagazzinamento di metadati relativi alla pubblicazione, grazie al suo sistema d'amministrazione;

≈ Accesso, con il quale la pubblicazione elettronica e i metadati ad essa associati sono resi disponibili attraverso una procedura di interfaccia.

≈ Amministrazione, responsabile per l'operatività del sistema.

---

<sup>146</sup> <<http://nedlib.kb.nl/index.html>>. Alcune delle notizie qui riportate sono state attinte da Faiella (2000).

≈ Conservazione, procedura grazie alla quale l'accesso e la leggibilità della pubblicazione elettronica può continuare sul lungo periodo.

Il modello è strutturato in maniera da seguire il ciclo di vita del documento digitale dalla sua creazione alla sua fruizione, attraverso le fasi dell'acquisizione, archiviazione, gestione dei dati, amministrazione, accesso e divulgazione. L'informazione viene corredata di metadata e quindi preparata come information package (IP), (che di volta in volta potrà essere un submission information package (SIP), piuttosto che archival o dissemination) affinché possa essere gestita nelle varie fasi del suo percorso.

Il pacchetto di informazioni per l'archiviazione è una ipotesi di lavoro, che si propone di dare una risposta a quali siano i metadata da apporre ad un documento per la sua preservazione e come essi vadano conservati.

Le procedure appena descritte sono state studiate nel dettaglio della loro attuazione in un minuzioso studio compiuto nell'ambito del progetto NEDLIB e sono da considerarsi il migliore fondamento di ogni sistema di deposito delle pubblicazioni elettroniche, secondo la ricerca europea.<sup>147</sup>

Il processo di formazione delle collezioni in ambiente elettronico è stato studiato approfonditamente dal progetto NEDLIB. Questo progetto ha mostrato la necessità di elevati investimenti per la realizzazione dell'infrastruttura tecnologica adeguata, con procedure tecniche uniformi di immagazzinamento e di conservazione delle pubblicazioni elettroniche, ma secondo Giuseppe Vitiello,<sup>148</sup> ha anche centrato il futuro delle collezioni nelle biblioteche nazionali, che risiede proprio in questo "scaffale digitale permanente" che si riuscirà ad implementare nelle biblioteche.

---

<sup>147</sup> NEDLIB ha fornito una proposta per un core set di metadata dedicato alla conservazione, URL: <<http://www.kb.nl/coop/nedlib/results/D4.2/D4.2.htm>> e anche un contributo - accolto - di revisione del modello OAIS, URL: <<http://www.kb.nl/coop/nedlib/results/OAISreviewbyNEDLIB.html>>. La documentazione è reperibile all'URL: <<http://www.kb.nl/coop/nedlib>>.

<sup>148</sup> Vitiello (2002) .

### 2.2.2. Electronic Resource Preservation and Access Network (ERPANET)

ERPANET (Electronic Resource Preservation and Access Network),<sup>149</sup> progetto avviato nel 2001 e terminato nel 2005, è stato finanziato dalla Commissione Europea nell'ambito del V Programma Quadro di Ricerca e Sviluppo.

Si proponeva di creare un Consorzio Europeo per far conoscere ed aiutare a mettere in pratica informazione, “best practice” e sviluppo di capacità nel settore della conservazione digitale del patrimonio culturale. ERPANET è diventata una rete di eccellenza di formazione, discussione e diffusione di materiali e soluzioni nel campo della conservazione di oggetti culturali digitali, tra cui gli archivi. È un progetto presentato all'Unione Europea dalle Università di Glasgow e di Urbino e dalla Scuola Archivistica di Amsterdam, nella stessa direzione del National Archives di Washington, che ha avviato la creazione di un laboratorio virtuale che coinvolga partner di ricerca a livello nazionale e internazionale nell'analisi di studi di casi concreti. ERPANET ha riunito organizzazioni culturali (musei, biblioteche e archivi), l'industria della tecnologia dell'informazione e della comunicazione (ICT), istituti di ricerca, organizzazioni governative (anche a livello locale), industrie dello spettacolo e della creatività, settori commerciali (compresi quelli farmaceutico, petrolchimico e finanziario). L'intento principale di ERPANET consisteva nel mettere a disposizione dell'utenza un deposito virtuale ed una base di conoscenza degli ultimi traguardi raggiunti nel settore della conservazione digitale.<sup>150</sup>

Nei risultati raggiunti da ERPANET si possono sicuramente ricordare l'obiettivo lavoro a favore della preservazione degli oggetti culturali e scientifici attraverso l'aumento della consapevolezza, fornendo l'accesso all'esperienza, condividendo politiche e strategie, migliorando la pratica, studiando i comportamenti in alcuni settori e svolgendo della ricerca. Tutti i materiali e le conoscenze acquisite in questo progetto sono divenuti il bagaglio più prezioso per le iniziative successive, come la DPE (cfr. il paragrafo successivo).

ERPANET ha lavorato per:

1. aumentare la consapevolezza delle necessità e strategie della preservazione digitale;
2. valutare le sorgenti di informazione e sviluppi nella preservazione digitale ed aumentare l'accessibilità di questo lavoro;
3. offrire un servizio a livello europeo di orientamento sulle questioni, pratiche e tecnologie della preservazione;

---

<sup>149</sup> <<http://www.erpanet.org/>>.

<sup>150</sup> Fonte di notizie è stata soprattutto Guercio (2002), p. 102 e Guercio (2004).

4. stimolare la ricerca nella preservazione digitale;
5. promuovere la consapevolezza tra i produttori di software delle necessità di preservazione della comunità utente;
6. migliorare i metodi e le pratiche attuali;
7. ha organizzato 21 seminari e workshop nell'Unione Europea.

L'Attività di ERPANET nel triennio 2001-2004 è stata:

1. affrontare casi di studio nel pubblico e privato;
2. con lo scopo principale di diffondere messaggio ampiamente;
3. affrontare il problema pur nella diversità nel tipo di organizzazione, vincoli organizzativi, e culturali con i quali si confrontava;
4. promuovere una consapevolezza di distinzione fra i records e le risorse;
5. accumulare e rendere accessibili informazioni sugli approcci alla longevità digitale:
6. identificare i temi di ulteriori ricerche.

Metodologie approntate per i casi di Studio di ERPANET

- ≈ Percezione e consapevolezza dei rischi posti dai materiali digitali
- ≈ Capire come la preservazione digitale influisce sulle organizzazioni
- ≈ Identificare quali azioni sono state prese per prevenire la perdita di materiali digitali
- ≈ Capire come le attività di preservazione sono monitorate (politiche o programmi)
- ≈ Capire come le organizzazioni pianificano i requisiti futuri per la preservazione (per es., i costi)
- ≈ Raccogliere le tecnologie disponibili

Sul sito Web di ERPANET, in Erpaguidance, sono stati messi a disposizione una prima serie di strumenti guida, utili alle istituzioni per valutare l'iniziativa della conservazione digitale: *Ingest Strategies, Orientamento sui costi, Scegliere la tecnologia, Linee guida per la conservazione digitale, Gestione dei rischi.*

### 2.2.3. DigitalPreservationEurope ( DPE)

“DigitalPreservationEurope” (DPE)<sup>151</sup> è il progetto che promuove attivamente collaborazioni e sinergie fra le iniziative nazionali esistenti nella European Research Area: tratta ed affronta le necessità per incrementare il coordinamento, collaborazione, e consistenza nelle attuali attività per assicurare la longevità dei materiali digitali; si basa sui risultati e l’esperienza di ERPANET (del 5° Programma Quadro della Comunità Europea, 2001-2004) utilizzando e valorizzando gli strumenti e la base di conoscenza prodotta e validata da un’estesa comunità di utenti

La DPE è una Coordination Action finanziata dalla Commissione Europea nell’ambito del Sesto Programma Quadro, della durata di tre anni, con inizio dal 1° Aprile 2006, che mira a coinvolgere tutti i paesi dell’Unione nell’attività di sensibilizzazione, promozione e ricerca sulle problematiche tecniche ed organizzative legate alla ‘digital preservation’, in una “rete di ricerca e sviluppo” nel settore.

Avvalendosi della collaborazione di esperti provenienti dall’ambito della ricerca accademica, degli istituti culturali, dell’amministrazione pubblica e dell’industria, la rete mira a sviluppare collaborazioni e sinergie tra le tante iniziative in corso nell’Unione Europea. DPE intende indirizzare le necessità e le sfide emergenti nella preservazione digitale, operando per trasformarle in risorse ed opportunità per tutta la comunità degli utenti.

Si può riassumere la DPE come la spiega Maurizio Lunghi<sup>152</sup>:

«Una azione di concertazione e coordinamento a livello europeo, che possa innanzitutto stimolare l’attenzione e la consapevolezza su tale problema spesso nascosto dietro le formidabili aspettative delle nuove tecnologie e poi per creare un’occasione di scambio delle esperienze dei vari paesi membri della Unione Europea al fine di unire gli sforzi verso obiettivi comuni».

Quest’azione di coordinamento parte dalla constatazione che la conservazione delle informazioni digitali e la possibilità di fornire ad esse un accesso continuato nel tempo sono temi troppo complessi perché un singolo soggetto possa sviluppare delle soluzioni ampiamente condivisibili ed attualmente lo sviluppo di soluzioni è affrontato in maniera frammentaria dal settore pubblico, dalle organizzazioni commerciali e da individui privati; diviene, quindi, molto sentita l’esigenza di costruire sinergie nell’area della preservazione digitale, in quanto, solo attraverso il coordinamento delle esperienze, delle risorse e degli

---

<sup>151</sup> <[www.Digitalpresevation.org](http://www.Digitalpresevation.org)>.

<sup>152</sup> Lunghi (2006); da questo contributo di M. Lunghi sono tratte le informazioni qui di seguito riportate.



sforzi di organizzazioni che operano in ambiti diversi e multidisciplinari, è possibile giungere alla produzione di risposte efficaci. L'attività della Coordination Action DPE permetterebbe anche di evitare la duplicazione degli sforzi da parte di organizzazioni individuali e di settori disciplinari diversi e potrebbe assicurare che le risorse sia umane sia finanziarie siano utilizzate nella maniera più efficiente ed efficace possibile.

La durata è 36 mesi a partire dal 1 Aprile 2006, nei quali:

≈ Creare una piattaforma coerente per una collaborazione proattiva, scambio e diffusione dei risultati di ricerca ed esperienze nella preservazione degli oggetti digitali.

≈ Incrementare la importanza dei servizi della preservazione e loro fattibilità e ripartizione dei costi. Agenda

≈ Aumentare la consapevolezza, competenze, e risorse disponibili.

Gli obiettivi della Coordination Action sono:

1. Sviluppare ed Attivare la Ricerca
2. Coordinamento di iniziative sui depositi digitali nella EU
3. Censimento delle iniziative sui depositi digitali e proposta di federazione
4. Facilitare la raccolta delle esperienze sui depositi
5. Repository Planning Tool
6. Servizi di Audit e Certificazione
7. Creare dei tool Self-Audit
8. Promuovere lo schema DPE per la Audit e Certificazione
9. Integrazione della Ricerca e Pratiche
10. Ricerca Pilota e Programma per lo Scambio Industriale
11. Creazione di un Database sulla Ricerca

Il progetto persegue due macro obiettivi:

1 - promuovere lo scambio di esperienze, la collaborazione e le sinergie fra esperti di settori diversi che operano nello studiare e fronteggiare le problematiche della preservazione delle memorie digitali, stimolando anche la creazione di agende comuni tra progetti di ricerca ed iniziative esistenti nell'area della ricerca europea;

2 - incoraggiare l'acquisizione di consapevolezza da parte di tutti sull'importanza e la centralità della preservazione delle memorie digitali per la futura società dell'informazione, sostenendo i decisionmakers, le autorità di settore, le agenzie di finanziamento, i

professionisti nel riconoscere i rischi ma anche le opportunità e comunque la necessità di attivare azioni di sviluppo nel settore.

Gli obiettivi sono pertanto diretti alla creazione di una rete di coordinamento che faciliti la diffusione del tema della preservazione delle memorie digitali attraverso l'esortazione a coordinare, condividere e disseminare informazioni, esperienze e risultati. Per realizzare questi obiettivi la parte centrale del progetto è identificata nella creazione di un'ampia 'community' che raggruppa diverse tipologie di partners ed utenti, in modo da organizzare gruppi di esperti di riferimento, documentazione e raccomandazioni, tutorial, programmi di formazione e creare sinergie operative.

Gli obiettivi specifici sono tre, a loro volta descrivibili attraverso ulteriori sotto obiettivi:

1. Creare una piattaforma polifunzionale per una cooperazione operativa e per la disseminazione di risultati, ricerche ed esperienze:

a. identificare e sviluppare la consapevolezza sul tema della preservazione degli oggetti digitali nell'ampio scenario del patrimonio culturale e scientifico europeo;

b. contribuire all'eliminazione della duplicazione degli sforzi nelle attività di ricerca da parte delle istituzioni e permettere l'identificazione, la raccolta e la condivisione di conoscenze ed esperienze;

c. creare un canale di comunicazione tra la comunità di ricerca e le comunità di utenti per incoraggiare approcci collaborativi alle necessità della preservazione;

d. stimolare ulteriori ricerche sulla preservazione digitale ed incoraggiare lo sviluppo e l'adozione diffusa di standard.

2. Stimolare la diffusione di servizi per la preservazione dei contenuti digitali:

a. sostenere lo sviluppo di un approccio condiviso a livello europeo per la verifica e la certificazione di 'trusted digital repositories' come passo essenziale nel costituire le basi per la realizzabilità e l'implementazione di servizi di deposito digitale;

b. stimolare le compagnie ICT e gli sviluppatori di software ad approfondire problematiche tecniche legate alla preservazione nelle nuove generazioni di software;

c. correlare più direttamente le agende sulla preservazione digitale con il settore industriale.

3. Incrementare l'acquisizione di consapevolezza, sviluppare le capacità e sfruttare le risorse disponibili:

a. valutare le questioni centrali che determineranno lo sviluppo di linee guida, metodi e strumenti per permettere la costruzione di funzionalità di preservazione dei contenuti digitali in Europa;

- b. avviare un insieme di seminari di formazione basati sulle best practice e identificare iniziative di formazione per i professionisti e di sviluppo del personale.

Per quanto riguarda i sott’obiettivi 2.a e 2.b, alla complessità della conservazione digitale e dei contenuti della risorsa digitale, una prima risposta può essere data da uno studio di comparazione tra esperienze diverse di Repository, ma anche da modelli già applicati, come OAIS e PREMIS, per la creazione di “Magazzini digitali”.<sup>153</sup>

Gli strumenti operativi con cui il progetto intende raggiungere gli obiettivi prefissati sono di diverso genere, in relazione agli ambiti di attività cui fanno riferimento: Integrazione, Formazione, Diffusione.

Integrazione delle attività di ricerca	Condivisione di esperienze e risultati, creazione di gruppi di lavoro di esperti su temi specifici, promozione di buone pratiche e linee guida condivise, organizzazione di un osservatorio sulle tecnologie, sviluppo di agende di ricerca comuni.
Formazione didattica e professionale	Tutorials, briefing papers, bollettini informativi, seminari e workshop.
Diffusione di consapevolezza nella comunità di utenti	Pubblicazioni, newsletter e report annuali, strumenti Web e servizi d’assistenza, materiale divulgativo e comunicati stampa, conferenze annuali. <sup>154</sup>

Il progetto DPE, inoltre, utilizza e valorizza gli strumenti e la base di conoscenza prodotta nell’ambito del progetto ERPANET, di cui come già detto si avvale per le esperienze e i risultati ottenuti dal 1999 al 2004.

Il Consorzio Digital Preservation Europe è stato formato per raccogliere diverse esperienze, interessi e capacità nel campo della preservazione delle memorie digitali. Tutti i membri sono ricercatori o professionisti nell’area della preservazione digitale, coinvolti in progetti ritenuti cruciali a livello nazionale, europeo ed internazionale. Posseggono approfondita esperienza di carattere internazionale nei propri settori disciplinari: ricercatori nell’area della preservazione digitale, attori nel settore culturale e scientifico, fornitori di nuove tecnologie, centri di comunicazione e formazione.

A luglio 2006 erano questi i membri del DPE:

<sup>153</sup> Cfr. Giovanni Bergamin, anche nella sua relazione presentata in Cultural Heritage, Firenze (2006).

<sup>154</sup> FP6 Preservation Showcase Annual Conference (Glasgow, 23 November 06).

- .. Direzione Generale per i Beni librari e gli Istituti culturali (Min. B.A.C.) – <http://www.librari.beniculturali.it/>
- .. FernUniversität Hagen – <http://www.fernuni-hagen.de/>
- .. Fondazione Rinascimento Digitale – <http://www.rinascimento-digitale.it/>
- .. HATII, University of Glasgow – <http://www.hatii.arts.gla.ac.uk/>
- .. Národní knihovna České republiky – <http://www.nkp.cz/>
- .. Nationaal Archief van Nederland – <http://www.nationaalarchief.nl/>
- .. Statsbiblioteket – <http://www.statsbiblioteket.dk/>
- .. Technische Universität Wien – <http://www.tuwien.ac.at/>
- .. Vilniaus universiteto Komunikacijos fakultetas – <http://www.kf.vu.lt/?lang=en>

Tra gli obiettivi chiave del progetto DPE, un elemento di grande importanza è dato dalla necessità di costruire una solida rete di collaborazione tra tutte le tipologie di utenti interessati alla preservazione delle memorie digitali; assunto che un aspetto critico, che riduce l'impatto delle attività di preservazione digitale, è determinato dai diversi livelli di consapevolezza riguardo a questo tema, la rete DPE deve, pertanto, cercare di costruire alleanze strategiche con gli attori consapevoli del valore della preservazione, identificabili in partners di altri progetti, organizzazioni internazionali e professionali, istituzioni culturali, che già operano nel settore, creando sinergie, condividendo risorse e consolidando canali di comunicazione. Come Seamus Ross ha indicato nella sua riflessione sulle metodologie di approccio delle istituzioni al tema della preservazione del digitale e quindi sul ruolo delle reti di Collaborazione/Cooperazione,<sup>155</sup> tra le iniziative che applicano il metodo della cooperazione interrelata (esempi come InterPARES, PADI, Digital Preservation Coalition, Digital Curation Centre, DELOS), anche Digital Preservation Europe (DPE), che segue a ERPANET, si pone proprio come una rete e si impegna anche in strategie di “costruzione di una comunità” dei soggetti operanti, quella “metacommunity” ideale già definita da Feliciano Faiella<sup>156</sup> indispensabile per promuovere gli scambi, ma anche la crescita della consapevolezza e di un'informazione distribuita.

Gli strumenti prodotti attraverso questa collaborazione (pubblicazioni, raccomandazioni, tutorial, linee guida, etc.) dovranno diventare strumentali per il coinvolgimento dei soggetti che non lavorano in questo settore ma che, comunque, hanno la necessità di preservare i propri contenuti digitali. Fine della rete DPE è di innescare un circolo virtuoso necessario per lo sviluppo e la capitalizzazione di diverse esperienze e background, derivanti sia dagli utenti

---

<sup>155</sup> Ross (2006); v. anche nel cap. 1.2.2.

<sup>156</sup> Faiella (2000).

‘consapevoli’, sia da quelli ‘non consapevoli’. La strategia della Coordination Action è quella dunque di ottenere un consenso ampio e trasversale, attraverso l’identificazione degli ‘Associate Partners’ che devono operare per implementare gli obiettivi del progetto.

I partners sono così identificabili:

1. Organizzazioni internazionali o professionali. Questa categoria di utenti è principalmente composta da partners istituzionali internazionali (per esempio UNESCO, NRG, PADI, RLG, DPC, DCC), con i quali stabilire alleanze strategiche per la disseminazione di informazioni sulle attività e le ricerche in corso.

I partners selezionati hanno specifici compiti e benefici derivanti dalla collaborazione con DPE: alcuni possibili termini di cooperazione possono essere identificati nel patrocinare eventi o condividere canali di comunicazione come newsletter, pubblicazioni etc.<sup>157</sup>

2. Partners che operano in altri progetti. Questi partners (per es. NESTOR, CASPAR PLANETS)<sup>158</sup> sono impegnati in progetti di ricerca europea. Il loro coinvolgimento nella rete DPE è decisivo per la condivisione e disseminazione dei risultati delle ricerche in Europa e nel mondo. Questa collaborazione aiuterà a superare la frammentazione delle attività in corso in Europa e per sviluppare un comune punto di vista sulla preservazione digitale. I partners devono stimolare lo scambio e la disseminazione di buone pratiche, promuovere l’adozione di raccomandazioni, linee guida e costituire gruppi di lavoro di esperti; possono attivare nuove metodologie di trasferimento delle informazioni e produrre materiale formativo diretto

---

<sup>157</sup> Si segnala anche l’iniziativa permanente, già esistente, delle conferenze annuali dell’IWAW, collegate con la European Conference on Digital Libraries (ECDL), iniziate nel 2001. URL: < <http://www.iwaw.net/> >.

<sup>158</sup> <<http://nestor.sub.uni-goettingen.de/index.php>>; <<http://www.casparpreserves.eu/>>; <<http://www.planets-project.eu>>. NESTOR (Network of Expertise in Long-Term Storage of Digital Resources) in Germania coopera con RLG e NARA per un repertorio di criteri per la valutazione di depositi digitali e la comparazione di sistemi di archiviazione a lungo termine, digitalizzazione e preservazione dei musei tedeschi; nel suo sito ha attivato un servizio di ‘mutua integrazione’ con l’australiano PADI. Sempre in Germania, KOPAL (Kooperativer Aufbau eines Langzeitarchivs digitaler Informationen, <<http://kopal.langzeitarchivierung.de/index.php>>) è il nuovo progetto della D. Nationalbibliothek, Staats und Univers.Bibliothek Göttingen, Gesellschaft f. wissenschaft. Datenverarbeitung mbH Göttingen (GWDG) e IBM Deutschland GmbH, per lo sviluppo di strategie di archiviazione digitale permanente con relativi strumenti e servizi, come lo sviluppo di un SW per la memorizzazione con controllo amministrativo. per istituzioni di ogni livello, adottando standard e metadati per l’archiviazione a lungo termine. CASPAR (Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval) è un consorzio tra leaders nella preservazione, partner competenti dei diversi settori culturali e scientifici e partner commerciali, che mira a stabilire una metodologia per la conservazione digitale di risorse eterogenee, prodotte in ambito scientifico, culturale e d’arte contemporanea (performing arts), con le linee Guida stabilite dallo standard OAIS, componenti per la ricerca e l’integrazione avanzata per un’ampia gamma di attività, una struttura e una piattaforma SW che offra una varietà di applicazioni innovative. Il progetto quadriennale FP6 IST5 PLANETS (Preservation and Long-term Access through NETworked Services) riunisce biblioteche ed archivi nazionali, centri di ricerca ed aziende di tecnologie dell’informazione europei (British Library, N. L. of the Netherlands, Austrian N. L., le Royal Library e State and University L. of Denmark, gli Archivi nazionali in Netherlands, Regno Unito e Confederazione Svizzera, le Università di Colonia, Friburgo, HATII di Glasgow, University of Technology di Vienna, ARC GmbH, IBM, Microsoft Research Lim.), uniti nella ricerca per la conservazione permanente dell’accesso alla conoscenza culturale e scientifica digitale, con lo sviluppo di servizi per la pianificazione di metodologie, strumenti etc., ed un ambiente integrato di produzione per la conservazione delle informazioni digitali, rivolto soprattutto alle organizzazioni tenute alla conservazione della memoria sociale e culturale.

a piccole organizzazioni interessate alle memorie digitali (biblioteche, musei, archivi, imprese). I partners, ovviamente, traggono vantaggio dall'adesione alla rete DPE in quanto possono condividere esperienze, acquisire visibilità e promuovere i risultati.

3. Comunità degli utenti. Questi partners sono rappresentanti da tutti gli utenti che fruiscono dei contenuti digitali, sia coloro che li producono sia gli utenti finali. Possono essere policy-makers di diversi settori, professionisti coinvolti in istituzioni culturali, comunità di ricerca, industria o semplici cittadini interessati ad acquisire competenze nel settore della preservazione digitale. In particolare possono essere suddivisi in tre sottogruppi:

- a. i data creators producono i contenuti digitali: data providers, editori, organizzazioni governative, istituzioni di ricerca, produttori di contenuti aperti, singoli cittadini;
- b. i data curators devono progettare, selezionare e implementare un'appropriata metodologia di gestione del deposito digitale, sufficientemente flessibile per soddisfare esigenze che mutano nel tempo;
- c. i data users hanno la necessità di utilizzare i contenuti digitali per le proprie applicazioni, ad esempio service providers, istituzioni di formazione ed educative, servizi di e-learning, e-government and e-health.

La Coordination Action intende includere questa ampia comunità per diffondere i temi della preservazione ed accogliere i contributi degli utenti, in termini di esperienze svolte o problematiche ancora non risolte. I membri possono beneficiare della rete DPE per scambiare e condividere con gli altri partners esperienze, best/bad practice, necessità e opportunità ed accedere agli strumenti operativi forniti dal progetto.

4. Associate partners. I membri della comunità degli utenti interessati al progetto DPE possono richiedere al consorzio di diventare membri del gruppo. Questa particolare categoria è considerata cruciale per permettere la diffusione del tema a livello nazionale, in quanto questi partners figureranno come leader della nazione cui appartengono e dovranno promuovere a livello nazionale le azioni della rete DPE, disseminare best practice, raccomandazioni e consapevolezza, anche presso nuovi ambiti professionali e diversi settori di ricerca, creare il consenso dell'utenza. Per diventare Associate Partner occorre avere un mandato da parte di autorità nazionali, che ne garantisca credibilità e visibilità in settori scientifici e/o culturali (ad es. le Biblioteche Nazionali, gli Archivi di Stato, associazioni culturali e professionali), ma anche in settori quali l'e-government, l'e-health o l'industria.

#### 2.2.4. DELOS (Network of Excellence on Digital Libraries)

Nell'ambito del V programma quadro della Commissione europea, è nato il progetto DELOS (Network of Excellence on Digital Libraries)<sup>159</sup> di durata triennale, partito dal 1 gennaio 2000 e rinnovato nel 2004. Lo scopo di DELOS è fornire un contesto per lo sviluppo continuato di una agenda di ricerca internazionale e costituire un punto di riferimento per progetti di Digital Library (DL) finanziati dal programma IST, stimolando lo scambio di esperienze in questo settore multidisciplinare, mettendo a disposizione banche di prova, per facilitare l'interoperabilità e fornire meccanismi per la valutazione di modelli, tecniche ed approcci e lo scambio di componenti software da fonti accessibili. Vuole contribuire alla definizione di standards rilevanti, allo sviluppo di modelli idonei per agevolare la diffusione di tecnologie Digital Library da parte delle industrie, alla creazione di una rete di collegamenti con la comunità internazionale dei ricercatori. Si propone di favorire il confronto tra le diverse professionalità che partecipano alla complessa e articolata società dell'informazione (informatici, documentalisti, bibliotecari, archivisti, operatori museali ed esperti di organizzazione). Il progetto individua una serie di priorità e di obiettivi misurabili per ciò che riguarda le attività di conservazione. I piani sviluppati all'interno dei progetti sono economicamente rilevanti e strettamente correlati alle iniziative nazionali, che tra l'altro essi hanno lo scopo di alimentare e sostenere. Le attività sono state suddivise in cinque Forum: DL Ricerca, Valutazione, Standardizzazione, Disseminazione e Trasferimento delle Tecnologie, Cooperazione Internazionale.

In questo contesto è nata l'idea di organizzare una Scuola estiva, una serie di corsi dove favorire l'incontro tra professionisti di discipline diverse e stimolare una comune attività di studio e ricerca nei diversi ambiti delle biblioteche digitali. In particolare la Scuola di DELOS intende sostenere la comunità dei ricercatori e dei professionisti che operano nel settore, con una proposta formativa che includa i temi di maggiore complessità tecnologica e organizzativa, tuttora oggetto di progetti di ricerca, ma anche al centro di sperimentazioni e di concrete applicazioni negli ambienti più avanzati della gestione e conservazione documentale in forma elettronica. Aperta a studenti di diversa provenienza geografica e professionale, nonché di diverso grado di conoscenza di base delle problematiche inerenti la conservazione digitale.

Il corso, intensivo e per la durata di una settimana, previsto per gli anni 2005-2006-2007, ha come obiettivo specifico quello di fornire un quadro generale dei metodi finora elaborati e degli strumenti più promettenti tra quelli in fase di sviluppo nel difficile lavoro di indagine

---

<sup>159</sup> (IST-1999-12262 EU FP5). URL: < <http://www.delos.info>>.

che le comunità professionali specifiche, tra cui in particolare i conservatori delle memorie digitali in campo archivistico e biblioteconomica, hanno avviato. Alla fine della settimana si intende di aver raggiunto progressi in termini di consapevolezza acquisita, di contenuti specifici appresi e di capacità di orientamento.

Più specificatamente le competenze che l'iniziativa si è proposta di sviluppare riguardano:

- una generale capacità di conoscenza delle criticità degli aspetti conservativi in ambiente digitale, sia per chi progetta biblioteche e archivi digitali, sia per chi ha il compito specifico della custodia a lungo termine;
- l'approfondimento dei principali interrogativi che sono oggi alla base dei più importanti progetti di ricerca;
- la coerente comprensione – non solo teorica – del complesso di attività che costituisce il nucleo centrale della funzione conservativa.

Può essere interessante segnalare alcuni degli argomenti trattati nel corso:

- o tecniche di modelling nel campo specifico della conservazione digitale;
- o definizione e valutazione dei metadati di settore;
- o i processi concreti di acquisizione e gestione delle risorse digitali nei depositi;
- o analisi del modello OAIS (Open Archival Information System) e della sua funzione per la creazione di informazioni di rappresentazione delle collezioni e degli archivi informatici;
- o tecniche di selezione e valutazione nel campo biblioteconomico e archivistico;
- o i processi di audit e certificazione dei depositi digitali;
- o creazione di test di sperimentazione e lo sviluppo di procedure di analisi per la progettazione di interventi mirati alla conservazione;
- o definizione di principi e tecnologie finalizzate a sostenere l'autenticità e l'affidabilità di oggetti digitali persistenti mediante l'utilizzo di formati standard.

L'iniziativa in sé e anche la riuscita effettiva di questi anni hanno suscitato ottimi giudizi:

«La Scuola estiva di Delos, [...] , può essere considerata la più significativa iniziativa europea di alta formazione specialistica nel campo/ della conservazione digitale, non solo per i contenuti avanzati resi disponibili e la qualità dei docenti, ma anche per le modalità di erogazione che hanno reso possibile l'attivo coinvolgimento di tutti i partecipanti».<sup>160</sup>

---

<sup>160</sup> Maria Guercio in Guercio (2005 b)



## 2.3. Australia: PANDORA.

PANDORA (Preserving and Accessing Networked Documentary Resources of Australia),<sup>161</sup> avviato nel giugno del 1996, è il progetto della National Library of Australia in collaborazione con altre nove biblioteche e istituzioni culturali australiane, che si è proposto di istituire un archivio digitale delle pubblicazioni elettroniche a livello nazionale. Ha realizzato un'archiviazione selettiva di fonti in rete (periodici elettronici, siti, pubblicazioni governative), in base ad una politica delle collezioni e a linee guida redatte della biblioteca stessa, ed ha sviluppato politiche e procedure per la preservazione a lungo termine delle pubblicazioni nazionali su Internet, implementando un servizio di agenzia per indicizzazioni, redazioni di abstract e allocazione di PURL (Persistent URL) per assicurare che i link nell'archivio PANDORA persistano nel tempo e che i documenti rimangano sempre accessibili, anche nel caso in cui l'editore smetta di esistere.

PANDORA fa riferimento all'applicazione dello standard Z39.50, al modello OAIS e agli standard internazionali per la catalogazione delle risorse e i metadati.

La National Library of Australia, per garantire una efficace politica di preservazione digitale sulle collezioni di PANDORA, continua la propria attività di ricerca sulle problematiche della preservazione, anche in collaborazione con altre organizzazioni, come RLG (Research Libraries Group), OCLC (Online Computer Library Center) e con l'IIPC (International Internet Preservation Consortium). Ha dato un contributo significativo alla campagna sulla salvaguardia del patrimonio digitale dell'UNESCO, nella preparazione della Carta del 2003.<sup>162</sup>

Inoltre, la National Library of Australia partecipa all'IFLA-CDNL Alliance for Bibliographic Standards (ICABS),<sup>163</sup> collaborando come partner responsabile, in particolar

---

<sup>161</sup> <<http://pandora.nla.gov.au/index.html>>

<sup>162</sup> V. già al § 1.2.1. e al Cap. 3, introduzione.

<sup>163</sup> L'IFLA-CDNL ICABS (Alliance for Bibliographic Standards), vede la National Library of Australia, la Library of Congress, la British Library, la Koninklijke Bibliotheek, la Deutsche Bibliothek, la Biblioteca Nacional de Portugal, IFLA Headquarters (Governing Board, Professional Committee, Divisions, & Sections) il CDNL (Conference of Directors of National Libraries) e i gruppi regionali di CDNL e infine l'UNESCO, partecipare a questo accordo per assicurare coordinamento, comunicazione e supporto per le attività fondamentali nel settore del controllo bibliografico, per tutti i tipi di risorse, relativi formati e protocolli standard. Il fulcro dell'alleanza è offrire strumenti concreti per migliorare organizzazione e sviluppi a livello globale; gli obiettivi sono: sostenere, promuovere e armonizzare gli standard esistenti, sviluppare strategie per il controllo bibliografico e l'avanzamento delle conoscenze sulle problematiche relative all'archiviazione a lungo termine delle risorse digitali, includendo la promozione di pratiche innovative e raccomandazioni per questo tipo di archiviazione. Le attività sono di coordinamento, di supporto, di garanzia, di diffusione; sono condotte attraverso obiettivi mirati ed azioni, che vengono rinnovati nei piani d'azione annuali (il primo risale al 2003); ogni partner dell'alleanza concorda il proprio supporto per una o più azioni, di cui diviene responsabile. URL: <<http://www.ifla.org/VI/7/icabs.htm#icabs>>.

modo, per quanto riguarda l'avanzamento delle conoscenze relative alla conservazione delle risorse elettroniche (*quarto obiettivo*: "Advance understanding of issue related to long-term archiving of electronic resources") e per quanto riguarda il ruolo del catalogo, per il miglioramento delle infrastrutture per la ricerca e le funzionalità per gli utilizzatori, delle conoscenze sui meccanismi della ricerca, gli standard, regole di catalogazione ecc. (nel ambito del *terzo obiettivo*, per il controllo bibliografico, "Develop strategies for bibliographic and resource control and ensure the promotion of new and recommended conventions"). Nell'ambito dell'archiviazione a lungo termine delle risorse elettroniche, cui ha collaborato al fianco della Koninklijke Bibliotheek, con le priorità di "Promoting standards, guidelines and best practices" e "Providing unrestricted Access to information", ha coordinato azioni per:

- ≈ il Web harvesting, con la ricerca e la promozione di metodiche per l'archiviazione di pubblicazioni Web raccolte con le tecniche di raccolta automatica,
- ≈ lo sviluppo di strumenti per migliorare la pianificazione della preservazione con la classificazione del contenuto (nell'azione 2005-'06) e implementazioni di soluzioni, come l'allarme di rischio per le risorse di cui si prevede l'inaccessibilità per obsolescenza in un breve lasso di tempo (nell'azione 2005-'07),
- ≈ far progredire la condivisione e la disseminazione delle conoscenze sui punti emergenti della preservazione a lungo termine e l'accesso alle risorse digitali.

In questi piani d'azione, la N. L. of Australia ha portato tutta la sua ormai decennale esperienza nel settore del Web archiving, con PANDORA e con il portale PADI, nonché i progetti nati per affrontare alcune particolari problematiche, come APSR (Australian Partnership for Sustainable Repositories) e AONS (open source Automatic Obsolescence Notification Service).<sup>164</sup>

Nel 2004 la N. L. of Australia, nell'ambito del suo contributo di collaborazione con l'ICABS, ha organizzato la conferenza *Archiving web resources: international conference*, tenuta presso la National Library of Australia, a Canberra, dal 9 all'11 novembre 2004.<sup>165</sup>

Nel 2005, in concomitanza con l'indagine della Koninklijke Bibliotheek, sempre su incarico dell'IFLA-CDNL ICABS, sullo stato dell'arte e le prospettive dell'archiviazione digitale nel mondo delle biblioteche,<sup>166</sup> la N. L. of Australia, per l'ICABS, ha condotto una ricerca per una possibile e adeguata guida di fonti per la preservazione dei materiali digitali,

---

<sup>164</sup> <<http://www.apsr.edu.au/>>.

<sup>165</sup> <<http://www.nla.gov.au/webarchiving/about.html>>.

<sup>166</sup> Di cui è stato già detto nel capitolo 1.3.

pubblicata nel *Report to ICABS on guidance for digital preservation: Report on a survey of Sources*.<sup>167</sup>

### 2.3.1. Politica di raccolta e Linee guida

Il nome, PANDORA, è una sigla che indica la mission dell'archivio: conservazione e accesso alle risorse documentarie in rete dell'Australia; a questi due obiettivi la N. L. of Australia dedica un quadro completo di azioni, volte a permettere il miglior accesso possibile a collezioni selezionate in base ad una pianificazione della raccolta e sottoposte ad una processo strutturato di preservazione. Con una politica di cooperazione con le biblioteche statali ed altre agenzie di raccolte culturali del Paese, si è fronteggiato al meglio il problema della quantità delle risorse on-line da archiviare e conservare e degli alti costi. La N. L. of Australia ha implementato l'infrastruttura, deciso la politica e le procedure tecniche per l'archiviazione delle pubblicazioni e dei siti Web ed ha invitato le Università e gli Enti Governativi e altre agenzie di raccolta a supportarla nella costruzione dell'archivio di PANDORA anche con progetti specifici come, per esempio, con i progetti ARROW Project (Australian Research Repositories Online to the World) e APSR. La N. L. of Australia sta collaborando con altre agenzie del Commonwealth che sono responsabili della creazione e dell'amministrazione delle informazioni on-line di governo, con GeoScience, con l'Australian Bureau of Statistics and the National Archives of Australia, per definire le responsabilità di conservazione e di accesso di lunga durata a questi dati. Collabora anche con le agenzie di tutto il mondo per condividere risultati esperienze e sviluppo delle tecnologie, mettendo a disposizione su un apposito portale Web (PADI) informative e confronti sullo stato dell'arte a livello internazionale, mentre la vasta documentazione archiviata sulla propria attività è messa sul sito Web di PANDORA ed anche sul sito Web della N. L. of Australia.

Le scelta di strategie selettive sono dovute anche al fatto che si è ritenuto più opportuno identificare, tra le moltissime pubblicazioni on-line che si moltiplicano nel Web, quei titoli che con maggiore probabilità sono destinati ad avere un interesse futuro per la ricerca e applicare tutte le risorse a disposizione al trattamento di questa selezione con standard di alta qualità.

I siti sono scelti sulla base di alcuni criteri generali predefiniti, privilegiando quelli che riguardano la nazione, un argomento di significato e rilevanza sociale, politica, culturale,

---

<sup>167</sup> Disponibile all'URL: <<http://www.nla.gov.au/nla/staffpaper/2005/documents/webb1.pdf>>.

religiosa, scientifica o economica, oppure scritti da un autore o una autorità australiana riconosciuta; c'è poi una selezione anche tra siti considerati più significativi per la popolarità o l'interesse che hanno suscitato nella società, sempre in ambito australiano.

Per quanto riguarda il metodo di selezione e l'archiviazione delle pubblicazioni del Web, la N. L. of Australia ha adottato un metodo selettivo perché, allo stadio attuale di sviluppo tecnologico dell'archiviazione Web, ne ha valutato i seguenti vantaggi:

≈ ogni articolo di pubblicazione Web inserito nell'archivio è della qualità maggiore possibile, compatibilmente alle attuali possibilità tecniche;

≈ con l'uso dei programmi che attualmente permettono di evidenziare gli aggiornamenti di un sito Web che cambi frequentemente, vengono tenuti il più possibile sotto controllo le risorse inserite;

≈ ogni articolo nell'archivio può completamente essere catalogato e quindi può rientrare nella bibliografia nazionale ed i dati bibliografici possono essere recuperati in diversi repertori bibliografici e OPAC;

≈ gli utenti necessitano di risposte rapide esaustive e non ridondanti e, quindi, le risorse nel catalogo devono essere integrate tra loro;

≈ ogni articolo nell'archivio può essere accessibile immediatamente ai lettori via Web, perché è stata negoziata l'autorizzazione con gli editori;

≈ i responsabili dell'archiviazione conoscono le proprietà di ogni risorsa e delle classi di risorse nell'archivio, e ciò aumenta la loro capacità di affinare i metodi e gli strumenti per raccogliere, immagazzinare, rendere accessibile i dati e, anche, la conoscenza di strategie di conservazione che saranno richieste per mantenerli accessibili per il lungo termine;

≈ i siti Web inaccessibili possono essere identificati e raccolti usando altri metodi, tramite accordi stipulati con l'editore; così pure per i siti e le Banche Dati commerciali, che possono richiedere una password fornita dell'editore.

Tuttavia, nella consapevolezza che il metodo selettivo inevitabilmente non coglie tutte le risorse importanti e le prende dal contesto del Web a cui sono collegate, la N. L. of Australia ha intrapreso raccolte dal dominio “.au”, per i periodi da giugno-luglio del 2005 ed agosto-settembre del 2006 come prova; continuerà una grande raccolta annuale di dominio almeno per i prossimi anni, al fine di sviluppare esperienza in questo metodo di archiviazione. La

biblioteca sta studiando gli ostacoli legali e tecnici a fornire l'accesso al contenuto delle raccolte di interi domini.

Per il prossimo futuro è probabile che il programma d'archiviazione del Web di PANDORA unisca il "metodo selettivo" ed il metodo di "dominio intero", realizzando i vantaggi di entrambi.

Ad ogni agenzia che contribuisce all'archivio di PANDORA è richiesto di effettuare la Guida di riferimento di selezione delle risorse australiane pubblicate on-line aggiornata e delle quali si assumerà la responsabilità. Questa Guida di riferimento è resa disponibile sul sito Web di PANDORA. Una pubblicazione, indipendentemente dal suo formato, è un'informazione che viene messa a disposizione del grande pubblico, o ad un pubblico identificato, gratuitamente o a pagamento. In teoria, questo include tutto su ogni sito Web pubblicamente disponibile all'interno del dominio Web australiano, in pratica, le agenzie che contribuiscono a PANDORA archiviano soltanto determinati tipi di informazioni on-line, come dettagliato nella Guida di riferimento di selezione.<sup>168</sup> L'archivio è una collezione di risorse pubblicate e non sono inclusi i materiali tecnici di carattere archivistico, pertanto, generalmente, non è ritenuto opportuno archiviare gli interi siti Web un ente governativo, per concentrarsi sui materiali pubblicati attraverso questi siti. I titoli sono selezionati per tema d'archiviazione; se un argomento è contenuto nell'intero sito, quello sarà il punto di accesso per la descrizione del catalogo e per gli indici di soggetto e di titolo sulla Home Page di PANDORA; le parti sussidiarie del sito Web, quali i rapporti o i bollettini, che non sono segnalate nel titolo saranno rintracciabili con il Search Engine.

Questa politica di selezione in una precisa area di interesse è indicativa della volontà di elevare il livello d'accesso e quindi la qualità del servizio: le fonti vengono selezionate in base al valore del contenuto della risorsa, ma vengono valutati anche la disposizione o considerazioni tecniche che appaiono stimolanti. L'archivio include formati dinamici e la N. L. of Australia ha messo a punto un metodo per includere le informazioni pubblicate nei database.<sup>169</sup>

---

168 Molta documentazione è disponibile sul sito Web di Pandora, come la Collecting Australian Online Publications, <<http://pandora.nla.gov.au/bsc49.doc>>.

169 Il metodo coinvolge l'uso di Xinq(<<http://www.nla.gov.au/xinq/>>), che permette ad un'interfaccia generica al contenuto delle basi di dati archiviate di essere installato; è stato sviluppato dalla biblioteca come componente per il trattamento del deep Web. Il software è disponibile con SourceForge: <<http://sourceforge.net/projects/xinq/>>.

### 2.3.2. Archiviazione - Catalogazione e garanzia di identificatori persistenti - Accesso per gli utenti.

Sull'archiviazione è stato redatto il *manuale PANDORA Archive Electronic Cataloguing Manual*,<sup>170</sup> pubblicato sul sito Web di PANDORA e ad uso di tutti i partecipanti.<sup>171</sup> Vi è una responsabilizzazione del personale di PANDORA nell'accertarsi che si tratti esclusivamente il materiale di cui si ha l'autorizzazione di utilizzo e che vengano aggiunti all'archivio quei filtri per garantire la qualità e il rispetto dei diritti.

I siti archiviati in PANDORA sono stati catalogati secondo i tradizionali standard bibliografici a livello massimo e sono inseriti nel catalogo on-line della N. L. of Australia, in condivisione con l'Australian National Bibliographic Database, disponibile con il servizio Bibliotecario dell'Australia:<sup>172</sup> ciò significa che le risorse sono registrate nella bibliografia nazionale e che, quindi, sono conformi con gli standard di PANDORA. Le Biblioteche australiane partecipanti inviano i propri inserimenti in formato MARC alla N. L. of Australia, che importa le registrazioni per i titoli archiviati dagli altri partecipanti di PANDORA nel proprio catalogo on-line. Sul sito del progetto sono presenti due semplici liste, l'una alfabetica, l'altra articolata in grandi aree tematiche; selezionando la voce si accede ad una scheda che fornisce alcuni elementi di descrizione e di contestualizzazione.<sup>173</sup>

La biblioteca ha elaborato un sistema d'assegnazione di PURL adatto a tutte le risorse Web archiviate in PANDORA e mantiene costantemente il controllo perchè questi identificatori persistenti non si deteriorino e affinché sia rispettata l'autorizzazione degli autori all'utilizzo.

---

<sup>170</sup> La N. L. of Australia ha redatto un manuale d'uso sulla catalogazione delle risorse nell'archivio PANDORA: *PANDORA Archive Electronic Cataloguing Manual*. I principali codici di catalogazione utilizzati sono: AACR2, 2ed., 2002 rev., *Chapter 9, "Computer files"*; CONSER *Cataloguing Manual, Module 31, Remote Access Electronic Serials (Online Serials)*; OCLC *Cataloguing Internet Resources: A Manual and Practical Guide*; OCLC *Cataloguing Electronic Resources: OCLC-MARC Coding Guidelines*, rev. 1999; Library of Congress *Guidelines for the Use of Field 856*, rev. 2003; Library of Congress *Guidelines for Coding Electronic Resources in Leader/06*, rev. 2003; Library of Congress *Use of Fixed Fields 006/007/008 and Leader Codes in CONSER Records*; Library of Congress *MARC21 Format for Bibliographic Data : Field List*; National Library of Australia *Kinetica Client Cataloguing*; National Library of Australia *Kinetica Guidelines for the Cataloguing of Electronic Resources*.

<sup>171</sup> I Partners di PANDORA sono: Australian Institute of Aboriginal and Torres Strait Islander Studies ; Australian War Memorial; National Film and Sound Archive; National Library of Australia; Northern Territory Library; State Library of New South Wales; State Library of Queensland; State Library of South Australia; State Library of Victoria; State Library of Western Australia.

<sup>172</sup> N. L. of Australia e Libraries of Australia. URL: <<http://librariesaustralia.nla.gov.au/apps/kss>>.

<sup>173</sup> Al marzo 2007 gli argomenti presentati sul Browser Subjects sono: Arts & Humanities; Health; News & Media; Business & Economy; History & Geography; Politics & Government; Computers & Internet; Indigenous Peoples; Science & Technology; Education; Juvenile; Society & Culture Environment; Law & Criminology; Sports & Recreation.

L'accesso all'archivio di PANDORA è semplice per chiunque, tramite Internet. Una piccola proporzione dei titoli (circa 2%) è limitata per i motivi commerciali, ma la maggior parte sono disponibili, anche gratuitamente, su una postazione PC nella sala lettura della N. L. of Australia. La biblioteca permette un certo numero di accessi contemporanei al contenuto dell'archivio. Sull'Home Page del sito Web di PANDORA sono presentati i temi-soggetti ed organizzate le liste per la ricerca; è inoltre disponibile sul sito Web un Search Engine nell'archivio di PANDORA, per la ricerca full text, ma esiste anche un'opzione avanzata di ricerca che consente le ricerche anche soltanto per data, oggetto, dominio e titoli.

È possibile trovare l'indirizzo di PANDORA attraverso motori di ricerca commerciali quali *Google* e *Yahoo!*, ma, in questa fase, non è disponibile l'accesso pubblico alle raccolte dell'intero dominio, benché sia possibile per alcune di esse.

### 2.3.3. Copyright e Deposito legale

L'archiviazione in PANDORA è effettuata in conformità con la legislazione, le direttive di governo e gli accordi con gli editori. La maggior parte delle giurisdizioni coperte dall'archivio di PANDORA ancora non hanno disposizioni legali del deposito che includano le pubblicazioni ed i siti Web on-line, tranne quelle nel territorio a nord dell'Australia: per queste la legislazione include le risorse del Web, senza alcuna ambiguità. Tre partecipanti, la Library of New South Wales, la State Library of Victoria e la State Library of Western Australia, hanno disposizioni governative che consentono di inserire le notizie nell'archivio e di fornire loro l'accesso. La N. L. of Australia, con l'assistenza dell'agenzia del copyright del Commonwealth, ha aderito alle disposizioni generali, assieme ad un numero crescente di enti governativi del Commonwealth. Al di fuori di queste disposizioni, coloro che contribuiscono a PANDORA sono obbligati a richiedere l'autorizzazione degli editori all'utilizzo di un titolo nell'archivio e all'accesso perpetuo. Questo permesso viene richiesto con una lettera preimpostata dalla N. L. of Australia e la risposta dall'editore a questa costituisce il permesso del copyright (la "patente" per l'IPR) e la corrispondenza con gli editori è conservata dalle segreterie amministrative dell'agenzia contribuente; la N. L. of Australia e gli altri partecipanti di PANDORA, nel rispetto del copyright, acconsentono ad un periodo di limitazione all'accesso in rete, se la pubblicazione è ancora in commercio e il software di PANDAS, infatti, è stato progettato con impostazioni che permettono delle limitazioni di accesso.

Nel negoziare sui diritti, i partecipanti di PANDORA mirano a realizzare il più breve periodo di limitazione compatibile con gli interessi commerciali dell'editore. La N. L. of Australia ha richiesto la correzione della Legge di copyright del 1968 per estendere le disposizioni legali del deposito fino le risorse on-line. Questa legge di copyright richiede agli editori di depositare una copia di ogni edizione presso la biblioteca di riferimento per ogni Stato, ma queste disposizioni ancora non riguardano le pubblicazioni elettroniche remote, solo in alcuni casi sono previste le risorse su supporto fisico, quali il CD-ROM e il DVD.

L'archivio nazionale ScreenSound si sta adoperando per ottenere un emendamento alla legge per includere tutti i tipi di pubblicazioni elettroniche, come pure la pellicola ed il suono, e sono stati presentati due progetti di legge al comitato per la revisione della legge sul copyright nel Deposito legale. Tutte le biblioteche del territorio e statali che contribuiscono a PANDORA, tranne nell'Australia occidentale, hanno una legislazione legale del deposito per le pubblicazioni della stampa e in alcuni casi questo è stato modificato per includere le pubblicazioni elettroniche. Nessuna delle biblioteche è però completamente sicura che la sua legislazione riguardi le pubblicazioni on-line, e, a causa della mancanza di disposizioni legali del deposito precise in merito, ciascun partners di PANDORA richiede l'autorizzazione direttamente agli editori. Caso speciale è la Tasmania, che ha una legislazione legale del deposito che riguarda le pubblicazioni on-line e non collabora con PANDORA; ciò ha contribuito alla costituzione, senza la necessità di possedere l'autorizzazione degli editori, di una sorta di "isola del Digitale": è STORS, l'Archivio a lungo termine dei Siti Web e documenti elettronici della Tasmania.<sup>174</sup>

Per la maggior parte delle informazioni sul Deposito legale e la posizione delle Biblioteche Nazionali su questo argomento, si veda il documento<sup>175</sup> scritto da Pam Gatenby nel 2002. PADI fornisce nelle sue pagine Web molte informazioni sul Deposito legale delle pubblicazioni elettroniche sia australiane sia mondiali.

---

<sup>174</sup> <<http://www.stors.tas.gov.au/logon.do>>.

<sup>175</sup> <<http://pandora.nla.gov.au/legaldeposit.html>>.



### 2.3.4. Infrastruttura PANDAS e formazione del personale.

Gi obiettivi nel sistema di archiviazione PANDORA sono i seguenti:

1. preservation
2. acquisition
3. accessibility
4. record management

A supporto dei flussi di lavoro e delle esigenze dell'archivio è stato implementato PANDAS, il PANDORA Digital Archiving System.

La gestione dell'archivio prevede un flusso delle risorse (si veda fig. PANDAS: Archive Management, pag. seg.) tra sotto-componenti dell'Archivio (l'archivio "Working", l'archivio "Preservation" e l'archivio "Display") e sotto-componenti di PANDAS (gestione dell'archivio, gestione dei diritti, acquisizione, gestione di garanzia di qualità, gestione del display per gli utilizzatori), come illustrato nel diagramma (v. fig. successiva):

1. Rights management: si assicura che i produttori delle risorse abbiano rilasciato la loro autorizzazione per la conservazione nell'archivio;
2. Acquisition: la maggior parte del materiale è raccolto con harvesting (metodo "pull"), mentre alcune risorse vengono depositate nell'archivio direttamente dagli editori (metodo "push");
3. Quality Assurance (QA) management: permette di garantire la qualità dell'archivio di conservazione, con due opzioni possibili: "clean up" (conservation, nei casi compromessi l'unico trattamento possibile) o "leave alone" (preservation, nel senso vero, cioè nel modo ottimale) e l'analisi avviene manualmente;
4. Public Display management: gestisce l'interfaccia dell'archivio per gli utenti; in questa fase, paradossalmente dopo tanti sforzi per acquisire e mantenere i documenti nel modo migliore possibile, si trova a dover, alcune volte, limitare l'accessibilità e restringere le possibilità di fruizione di tanti documenti perfettamente conservati, per osservare le restrizioni legali che possono intervenire.

PANDAS è un sistema composto da diversi sottosistemi (si veda fig. PANDAS: System substructure, pag. seg.) che svolgono le diverse funzioni:

1. Record Management: provvede a un interfaccia d'uso per il manager della collezione, per gestire l'archivio di PANDORA;
2. Acquisitions: il sottosistema può anche acquisire gli inserimenti in maniera programmata;
3. Display Restriction: può azionare restrizioni al pubblico accesso, disabilitare o riabilitare l'accesso;
4. QA Processing: per agevolare la preservazione, esegue richieste di controllo sull'archivio di lavoro;
5. Preservation: gestisce le operazioni di preservazione ;
6. Notification: grazie ad un servizio di messaggistica interno, ogni sottosistema può inviare messaggi, nell'ambito del proprio ruolo, con informazioni utili a completare o proseguire il processo.

L'Archivio di preservazione è composto di tre archivi:

- Working: ove sono lavorati i dati e vengono elaborate anche le richieste del QA Processing;
- Preservation: la vera e propria "camera della preservazione";
- Display: dove vengono alloggiati i risultati per la ricerca.

PANDAS usa una piattaforma Solaris e Linux, il software è stato sviluppato sui questi due sistemi operativi; il webserver è Apache. È stato usato un DB relazionale Oracle per l'archivio.

PANDAS è stato sviluppato con tecnologia Java in ambiente Apple WebObjects Application.

Per l'acquisizione è utilizzato il robot di harvesting HTTrack.

È stato anche sviluppato un sistema di valutazione (PANDAS Evaluation System).

L'IIPC produce degli aggiornamenti sulla preservazione e questi vengono, appena possibile, tradotti in adeguamenti del software, che adatta le ultime indicazioni ai bisogni di PANDORA (come è avvenuto per lo sviluppo dell'applicativo Xinq, open source, per il deposito da parte degli editori). La biblioteca ha installato quest'infrastruttura e continua a svilupparla arricchendola, senza costi aggiuntivi, grazie alle collaborazioni dei partners.

PANDAS è stato usato per la prima volta nel mese di giugno del 2001 e una seconda versione molto avanzata è stata pubblicata nel mese di agosto del 2002; nel 2006 si è giunti alla versione 3.0.

PANDAS fornisce un identificatore persistente al livello di titolo e genera una volta per tutti gli elementi, per esempio, per un articolo all'interno di un'edizione di un periodico elettronico, o per un'immagine o una tabella all'interno di un sito Web. L'identificatore è persistente e può essere tranquillamente citato perché garantito; è collocato nella parte inferiore di ogni pagina dell'entrata di titolo, appena sotto l'identificatore per il titolo.

I documenti tecnici per le informazioni più dettagliate sul sistema sono recuperabili sempre sul sito di PANDORA. La biblioteca inoltre intende mettere liberamente a disposizione il software PANDAS per chiunque desideri usarlo. Accedere al sistema di valutazione (versione 2) non è attualmente più possibile a causa del rilascio di PANDAS 3.0. (2006).

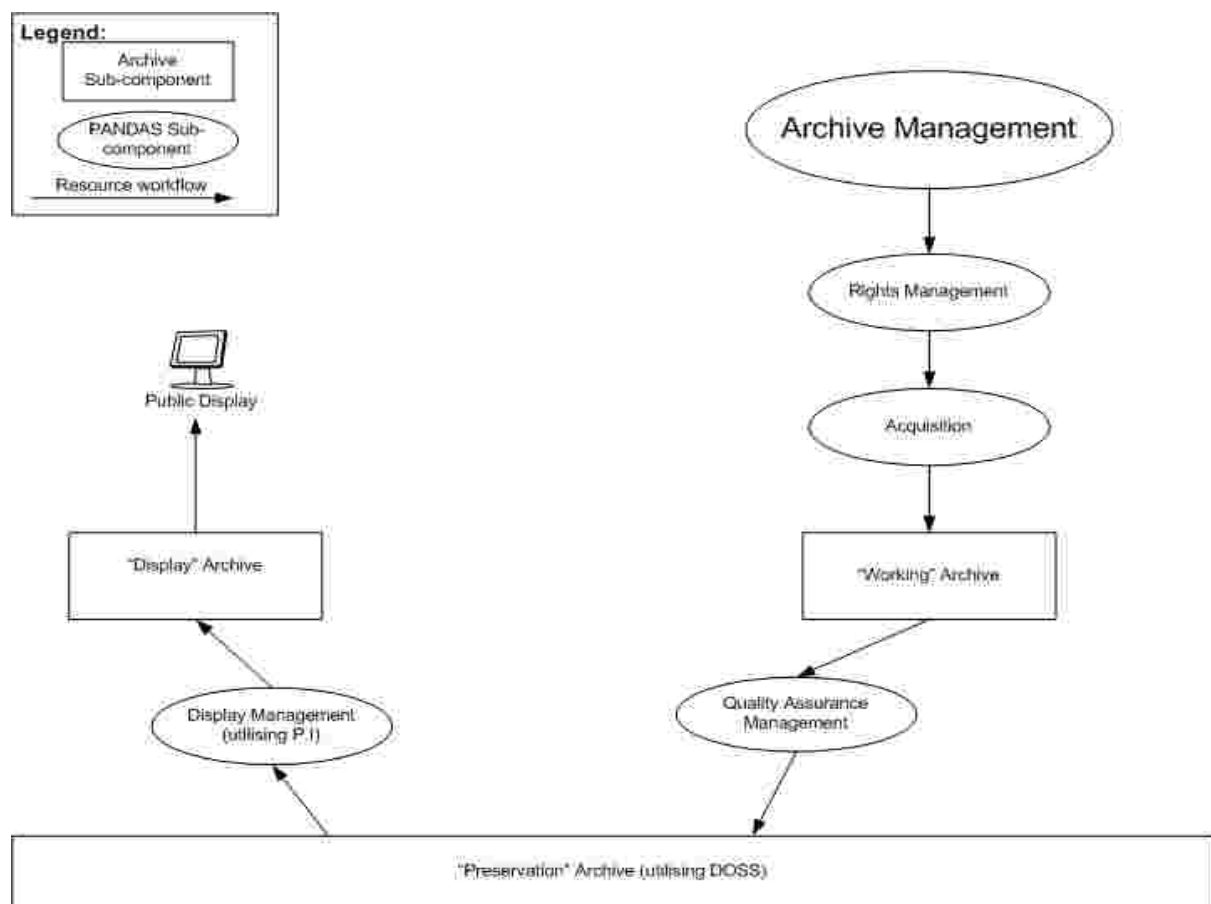
Al personale è richiesta formazione continua, necessaria per la continua evoluzione in questo settore, abilità specifiche, attitudine ma anche un impegno autodidattico; alla formazione collaborano, anche economicamente, le agenzie che partecipano a PANDORA.

Nel corso del lavoro d'archiviazione Web, la N. L. of Australia ha sviluppato e messo a disposizione un certo numero di manuali e di guide per aiutare il proprio personale ed altri partecipanti ad intraprendere le procedure in questione, pensando che possa essere di interesse per altre agenzie che stanno studiando la possibilità di installare un archivio selettivo, PANDAS fornisce:

- un manuale di archiviazione, che dettaglia i processi ed i workflows seguiti nella costruzione dell'archivio di PANDORA;
- una guida veloce, con indicatori-chiave, per l'uso di PANDAS, una lettura necessaria per chiunque intendesse studiare la possibilità di chiedere l'accesso al sistema per scopi di valutazione;
- un manuale di PANDAS completo, con le procedure graduali per usare tutte la funzionalità del sistema, spiega workflows e mansioni previste;
- il manuale di catalogazione elettronico, che dettaglia il metodo di catalogazione delle risorse on-line;
- FAQ, le domande frequentemente fatte, che PANDORA archivia, sono offerte, con le loro risposte, come informazioni per la soluzione di quesiti e problemi.

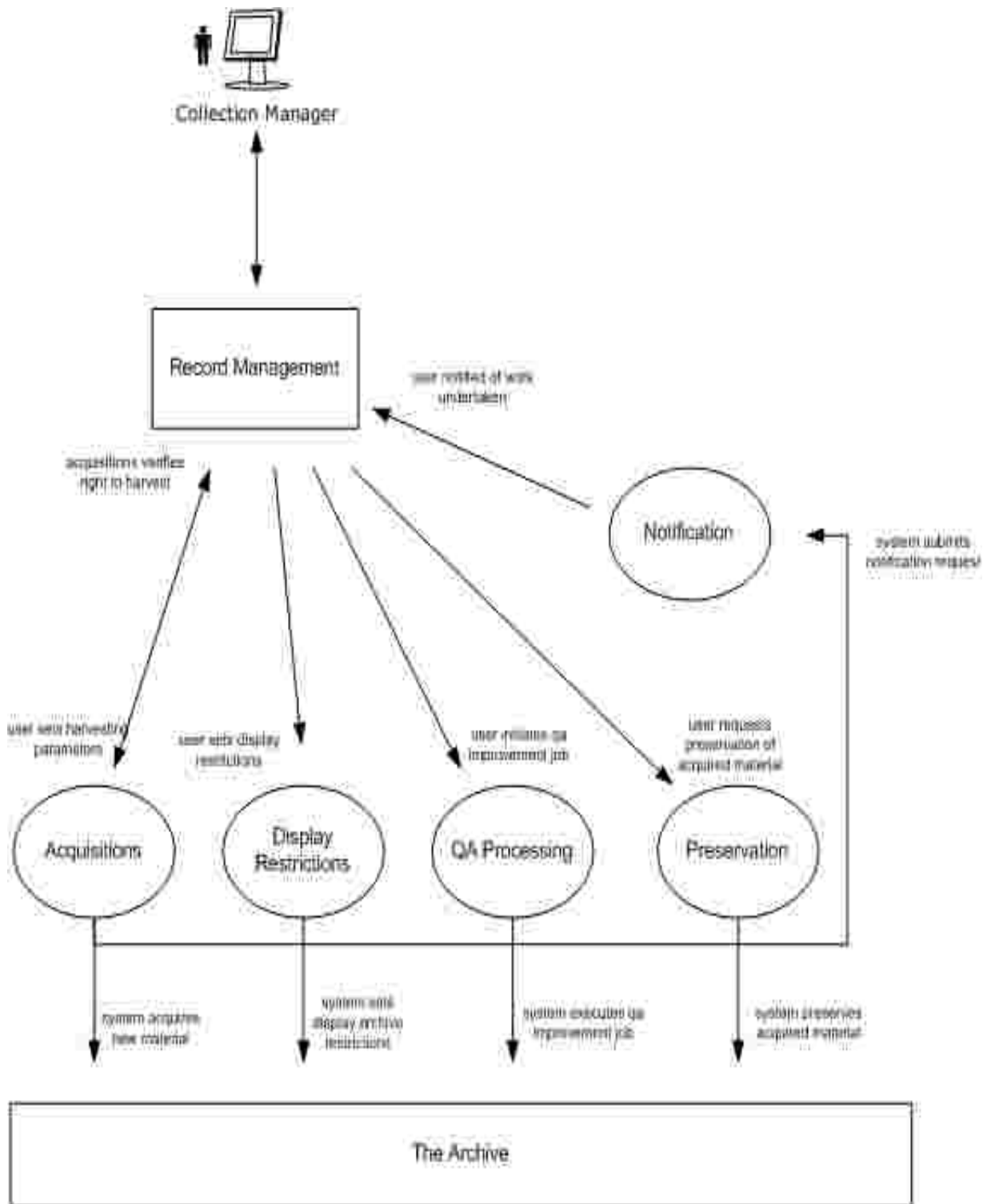
La N. L. of Australia mette a disposizione sul sito di PANDORA manuali che dettagliano le procedure di archiviazione e l'uso del software di PANDAS. Vengono organizzate periodicamente esercitazioni d'addestramento e PANDAS ha un sistema di training per il programma. La Biblioteca Nazionale, inoltre, fornisce i criteri di selezione per reclutare il personale.

Dalla documentazione disponibile a partire dal sito Web di Pandora sono tratte le due illustrazioni seguenti, utili a comprendere la struttura di PANDAS (il PANDORA Digital Archiving System):<sup>176</sup>



PANDAS: Archive Management

<sup>176</sup> <<http://pandora.nla.gov.au/pandoratech.html>>.



PANDAS: System substructure

### 2.3.5. La preservazione digitale e il ruolo di PADI

La N. L. of Australia definisce la conservazione digitale come i processi che portano a realizzare un determinato livello di accessibilità degli oggetti digitali nel tempo; ciò coinvolgendo un certo numero di elementi critici riguardo a come i materiali digitali sono generati, selezionati, trasferiti, descritti e controllati. Ha sviluppato una vera “politica di preservazione del digitale”, con la precisazione dei principi che guidano le relative attività digitali di conservazione e ha intrapreso una valutazione di rischio dettagliata per le relative collezioni digitali. Usa un piano d'azione di preservazione “in evoluzione”, capace cioè di adeguarsi alle esigenze che possono presentarsi e alle valutazioni sui risultati raggiunti; collaborando con l’Australian Partnership for Sustainable Repositories (APSR), identifica eventuali priorità e indirizza al relativo lavoro digitale di conservazione.

La Biblioteca nazionale, anche tramite il sito Web, offre un'informazione completa su tutta la propria attività di preservazione del digitale, dal punto di vista teorico, metodologico e di sistema informativo.

Nell’ambito delle attività di PANDORA rientra anche la realizzazione di PADI (Preserving Access to Digital Information), un portale tematico delle fonti relative alla preservazione delle memorie digitali. Si tratta di uno strumento di supporto on-line che si è proposto come punto di raccordo per informazioni, iniziative, forum specializzato, con l’obiettivo di facilitare lo sviluppo di strategie e linee guida per la preservazione dell’accesso alle informazioni digitali e ai Web Archives delle biblioteche nazionali. Ad oggi vi sono presentate informazioni sui programmi di sedici paesi: Australia, Austria, Canada, Repubblica Ceca, Danimarca, Finlandia, Francia, Germania, Giappone, Lituania, Olanda, Nuova Zelanda, Norvegia, Svezia, Regno Unito, USA. In questi Paesi, con programmi d’archiviazione di risorse Web portati avanti dalle biblioteche nazionali, si prospettano ricerche e soluzioni di vario genere, tra cui l’applicazione del Deposito legale relativo alle collezioni elettroniche e la possibilità di stabilire partnership internazionali attraverso le quali esplorare e testare un più ricco ventaglio d’ipotesi per la progettazione e la costruzione di archivi Web affidabili.

Per quanto riguarda PANDORA, che è uno dei primi archivi Web costituiti al mondo, data la sua riuscita esperienza, spesse volte, è stato oggetto di studio da parte di quelle biblioteche che intendevano programmare la propria archiviazione del Web.

## 2.4. Norvegia: il progetto PARADIGMA

Il progetto d'archiviazione del Web avviato dalla Nasjonalbiblioteket norvegese (la Biblioteca Nazionale, d'ora in poi anche B.N. No.) nel 2001 nasce su un terreno particolarmente predisposto a tale impegno, in quanto la Norvegia è probabilmente stata la prima Nazione a considerare con attenzione, anche dal punto di vista legislativo, il patrimonio disponibile sui materiali prima analogici e poi digitali, partendo quindi molto presto con un atteggiamento di considerazione nei confronti del patrimonio bibliografico e culturale veicolato dai supporti non tradizionali forniti dalle nuove tecnologie.

Il caso norvegese può essere considerato un esempio sotto diversi aspetti:

- per la gestione dei supporti d'informazione e di conservazione di beni culturali è stata sviluppata una progettualità con nuove strutture fisiche e concettuali;
- si è ottemperato ad un'idea politica di salvaguardia e documentazione del proprio patrimonio culturale a tutto tondo;
- sono state seguite le proprie tradizioni catalografiche e di servizio bibliografico, di disponibilità alla cooperazione a livello territoriale<sup>177</sup> (si pensi ai vari progetti che legano il mondo scandinavo, come il già citato Nordic Web Archive), ma anche di confronto e attenzione rivolti alle progettualità di altri paesi, nell'Unione Europea e nel resto del mondo (con riferimento, per esempio, alla partecipazione esterna, in veste di "osservatori",<sup>178</sup> al progetto europeo MINERVA, nonché ai legami di confronto e collaborazione con i progetti australiani MAVIS e PADI).

La Nasjonalbiblioteket della Norvegia, ha costituito l'Archivio Digitale Nazionale (A. D. N.), collocato fisicamente nel Vault<sup>179</sup> della sede di Mo i Rana e ha avviato nel 2001 il

---

<sup>177</sup> Al proposito, Unni Knusten chiarisce alcuni aspetti che caratterizzano questa volontà di cooperazione dei paesi scandinavi (Danimarca, Islanda, Norvegia, Svezia e Finlandia): pur senza dimenticare le loro diversità, comunque ne sortisce un affresco che dimostra un forte tentativo di costruzione di un comune sistema di servizio sul patrimonio bibliografico. Sia che si tratti di deposito legale, sia del problema delle risorse elettroniche remote, sia della bibliografia nazionale, in queste regioni del Nord si vede lo sforzo comune per ottenere un alto livello di controllo bibliografico e la cooperazione instauratasi nel tempo ha fatto sì che si realizzassero efficaci reti di cataloghi regionali, che da tempo sono quasi tutti in Internet, e di sistemi che conducono al progetto chiamato SVUC (Scandinavian Virtual Union Catalogue). Kunsten (2006).

<sup>178</sup> Si ricorda che ad oggi la Norvegia non fa parte dell'Unione Europea, ma segue spesso programmi europei con il ruolo di 'osservatore'; per esempio, ha accolto e attuato per prima l'invito dell'Unione Europea di raggruppare Musei, Biblioteche ed Archivi in un unico settore ("MBA") ed oggi questo settore è il più sviluppato d'Europa, al fianco della Gran Bretagna. Si veda, on-line: The Norwegian Archive, Library and Museum Authority. URL : <<http://www.abm-utvikling.no/om/english.html>>.

<sup>179</sup> Il vault, o *caveau*, è costituito dalla struttura sotterranea dove sono stati collocati i depositi che, grazie all'ottimale condizionamento climatico naturale della sede, oltre alle propizie condizioni latitudinali della stessa Mo i Rana, rappresentano una situazione ideale per la conservazione del materiale digitale.

progetto PARADIGMA, con il quale ha pianificato l'immagazzinamento dei materiali digitali depositati nell'A. D. N., con l'aspirazione che questo archivio possa un giorno diventare una fonte vitale per le ricerche degli studiosi e la documentazione sulla cultura e la società norvegese. La piattaforma tecnica per l'A. D. N., il "Vault digitale della Memoria a lungo termine", è uno degli investimenti tecnologici più onerosi fatti dal Governo della Norvegia nel settore culturale: la sua struttura e gli sviluppi continui hanno goduto della massima priorità di fondi della B. N. della Norvegia in questi ultimi tre anni. La capacità d'immagazzinamento del Vault al momento non annuncia limiti quantitativi sul numero di documenti digitali che possono essere raccolti dal dominio norvegese del WWW, quindi è stato considerato che la sfida da affrontare non sta in un deficit di capacità di memorizzazione, ma piuttosto nel "come" gestire questo gigantesco aumentare di produzione documentale.

Nel panorama mondiale di implementazioni per la conservazione del Web, questo progetto si caratterizza per due aspetti di fondo: la legge nazionale sul deposito legale e la ricerca bibliografica. Innanzitutto si fregia di basarsi su una delle legislazioni più avanzate sul Deposito legale, che, già dal 1989, impone di considerare documenti su qualsiasi supporto, e, pertanto, ha alle spalle una Legge nazionale che ha portato alla salvaguardia della produzione culturale norvegese esistente in Internet. Nel suo secondo aspetto c'è la volontà dei responsabili del progetto di realizzare un approccio completo alle esigenze che sorgono dagli utenti e alle loro aspettative, attraverso la progettazione di sistemi informativi a misura del bisogno dei fruitori di recuperare l'informazione più completa, nonché il documento nella sua unità, a fronte della grande complessità da cui è tipicamente caratterizzato il documento digitale.

#### 2.4.1. Il contesto

La Nasjonalbiblioteket appartiene al gruppo di Biblioteche Nazionali del Nordic Web Archive (NWA)<sup>180</sup>, che in Europa stanno realizzando l'harvesting e l'archiviazione dei documenti digitali nell'ambito dei propri domini Internet nazionali, basandosi sulla normativa concernente il Deposito legale.

La B.N. No. ha partecipato ai progetti internazionali relativi al Deposito legale di documenti digitali NEDLIB e BIBLINK; ha attivato dal 2003 la Norwegian Digital Library

---

<sup>180</sup> NWA, URL: <<http://nwa.nb.no/>>. V. cap. 2.1.2.



Initiative (NDLI) per l'accesso pubblico e facile alle fonti dell'informazione e della conoscenza.<sup>181</sup>

Relativamente all'offerta su Web delle risorse digitali, si segnala anche la recente creazione del portale Kulturnett, il cui scopo è quello di presentare le risorse elettroniche norvegesi al grande pubblico.<sup>182</sup>

#### 2.4.2. La legge sul Deposito legale.

Il progetto che qui viene presentato è legato sia alla storia della B.N. della Norvegia, sia alla storia del Deposito legale in quel paese e su questi legami può risultare interessante dare qualche ragguaglio.

La B.N. norvegese nasce come biblioteca di tipo universitario, con un patrimonio costituito, sin dal XVII sec., sostanzialmente attraverso l'istituto del Deposito legale, cioè con un fondo già selezionato in base ad una 'identità nazionale', più che ad una specializzazione, ma con la contraddizione di un'identità non esattamente propria, norvegese o sami, e, nei fatti, rivolta ad un pubblico specializzato, come quello universitario.<sup>183</sup>

Ha iniziato a realizzare pienamente le funzioni nazionali, come la missione di documentazione e conservazione, alla fine degli anni Ottanta del Novecento, epoca che coincide anche con l'insediamento della nuova Nasjonalbiblioteket norvegese nella sede di Mo i Rana, città che si trova a più di mille chilometri a nord di Oslo, a ridosso del Circolo polare artico, in una scuola dismessa ristrutturata per le nuove funzioni. La nuova biblioteca viene a costituire l'istituto di conservazione della memoria scritta della Norvegia, curando l'archivio cartaceo ed elettronico delle pubblicazioni norvegesi, mentre nella sede di Oslo restano la raccolta storica, l'agenzia bibliografica e il catalogo collettivo.

---

<sup>181</sup> Per il progetto BIBLINK, relativo ad una possibile alleanza tra Editori e Biblioteche per il controllo bibliografico, progetto dell'U. E. degli anni 1996-2000, si veda all'URL: <<http://hosted.ukoln.ac.uk/biblink/>>.

Per gli altri progetti in corso della B. N. Norvegese: Van Nuys (2005 a).

<sup>182</sup> URL: <<http://www.kulturnett.no>> e <<http://norskdigitalbibliotek.no/englishindex.html>>.

<sup>183</sup> La Norvegia rimase unita alla Danimarca fino al 1814 e, dunque, priva di proprie raccolte librerie di origine reale o nobiliare; la raccolta maggiore è in origine quella della Biblioteca Universitaria, anche se il nucleo era formato dai duplicati della biblioteca danese, che, secondo il dettato del deposito legale del 1697, aveva il ruolo di punta della raccolta del patrimonio culturale nazionale danese e riceveva cinque esemplari, che venivano poi distribuiti; queste collezioni passarono, verso la fine dell'Ottocento, alle rispettive Biblioteche Nazionali: la Det Kongelige Bibliotek di Copenaghen e la Nasjonalbiblioteket di Oslo. La prima legge norvegese sul deposito legale nasce per motivi di censura nel 1815; poi, nel 1882, fu emanata una legge più adeguata per la costituzione del patrimonio culturale e dette immediatamente impulso alla pubblicazione della bibliografia nazionale. (Vitiello, già cit., p. 62, 64).

Alla costituzione della seconda sede<sup>184</sup> della Biblioteca nazionale, che diventa così la sede della ‘memoria virtuale’ della Biblioteca nazionale stessa, si accompagna l’innovativa legge sul Deposito legale: la legge di riforma che, per la prima volta al mondo, estende il deposito anche alle pubblicazioni elettroniche. Dopo la legge del 1882, che obbligava alla bibliografia nazionale, cui fece seguito il Legal Deposit Act del 1939, sempre relativo alle sole opere a stampa, arriva la proposta di legge di riforma sul Deposito legale, elaborata nel 1988: è emanata il 9 giugno 1989<sup>185</sup> e nel 1990 viene adottata dal Parlamento, proprio mentre si dà vita alla sede di Mo i Rana.<sup>186</sup>

Lo scopo del Legal Deposit Act del 1989 è:

«...garantire che i documenti che contengono informazione pubblicamente disponibile siano depositati in collezioni nazionali, affinché queste registrazioni della vita sociale e culturale norvegese possano essere conservati e resi disponibili in quanto fonti finalizzate alla ricerca e alla documentazione» (§ 1).

Ritenuto estremamente moderno quando apparve, l’attuale Legal Deposit Act considera tutti i documenti norvegesi “pubblicamente disponibili” memorizzati su “qualsiasi supporto”, ad esempio carta, microforme, fotografie, documenti su più supporti, registrazioni sonore, film, video, documenti digitali e programmi radio o teletrasmessi; sono rappresentati anche i documenti pubblicati all’estero da o per editori norvegesi e quelli destinati in special modo ad un pubblico norvegese. In più, il deposito dichiara la volontà di proseguire quel “pubblicamente disponibile”, quindi l’accesso, indipendentemente dal loro supporto.

Per la raccolta e l’archiviazione dei documenti digitali coperti dalla legge sul Deposito legale, la B. N. Norvegese ha dato l’avvio nell’agosto 2001, con termine il 31 dicembre 2004, al progetto PARADigMa (Preservation, Arrangement and Retrieval of Assorted Digital Materials)<sup>187</sup>, volto alla conservazione, sistemazione e recupero informativo di ogni tipologia di materiale digitale. Lo scopo del progetto è proprio quello di sviluppare e stabilizzare procedure per la raccolta e la preservazione dei documenti digitali e permetterne la fruibilità, ricollegandosi alla legge che, dal 1990, ha sancito il deposito dei documenti

---

<sup>184</sup>La coraggiosa scelta di divisione della Biblioteca Nazionale è opera, sempre dei primi anni ’90, del Bibliotecario Nazionale (Bendik Rugaas), una figura istituita sempre in quel periodo, incaricata di coordinare il lavoro delle, peraltro autonome, biblioteche di ricerca norvegesi, che così assecondava l’aspirazione della comunità bibliotecaria. (Per questa parte Vitiello (2002), p. 64-66).

<sup>185</sup> [Legge sul Deposito legale (1989)] Norway. *Act relating to the legal deposit of generally available documents : no. 32 of 9 June 1989 : with regulations*. [Published by the Ministry of Church and Cultural Affairs ; unofficial English translation published by the National Library of Norway. [Oslo]: National Library of Norway, 1997.

<sup>186</sup> Vitiello (2002); Van Nuys (2003), ove sono riportati alcuni passi fondamentali del Legal Deposit Act, tra cui quello riproposto qui di seguito.

<sup>187</sup> URL: <[http://www.nb.no/paradigma/eng\\_index.html](http://www.nb.no/paradigma/eng_index.html)>, da cui sono tratte le informazioni generali qui riportate. Il sito non è più attivo dal febbraio 2007. È possibile però trovare, nella banca dati di ricerca sul sito della B. N., la raccolta del Forum di discussione sul progetto aperto dall’agosto 2002 all’ottobre 2004, Archives of PARADIGMA-FORUM@NB.NO - *Paradigma diskusjonsforum* (ult. visita il 12 marzo 2007).

digitali, ma che, fino a questo progetto, vedeva una ridotta possibilità di accessibilità ai documenti stessi: al momento del varo della legge, infatti, gli strumenti e le infrastrutture erano adeguati ad accogliere solo digitali su supporto fisico, come le pubblicazioni su CD-Rom e le pubblicazioni elettroniche raccolte manualmente.

Secondo la legge sul Deposito legale, tutti i documenti devono essere registrati e resi accessibili per la ricerca e la documentazione, mentre il Copyright Norvegese regola e limita l'accessibilità di quanto conservato, proteggendo il Diritto d'autore sui documenti digitali: così, la capacità di garantire un alto livello di protezione alla proprietà intellettuale, restringe fortemente i modi e le possibilità in cui le biblioteche possono rendere accessibili i documenti digitali per il loro pubblico. La Direttiva dell'Unione Europea sulla Società dell'Informazione, emessa nel 2001, armonizza alcuni aspetti della legge Europea sul Diritto d'autore, e da quando questa è entrata anche nella legislazione norvegese, alla fine del dicembre 2002, ha anche influenzato il modo in cui PARADIGMA tratta i documenti digitali sotto questo aspetto. Infatti questo progetto, poiché si poneva come primo obiettivo il garantire un Deposito legale soddisfacente dei documenti digitali norvegesi, ha disegnato un sistema complesso, che prevede lo sviluppo di tecnologie, metodologie e routine per la selezione, la raccolta, la descrizione e l'identificazione di tutti i tipi di documenti digitali - compresi quelli pubblicamente disponibili in Internet - e, pertanto, ha dovuto affrontare anche l'aspetto legale già nella strutturazione dell'accesso al proprio archivio Internet: ci si propone di condurre ulteriori approfondimenti e studi per chiarire l'aspetto legale per la raccolta, l'archiviazione e la fornitura d'accesso ai materiali digitali depositati per legge. PARADIGMA sostiene la necessità per la B.N. No. di trattare accordi con i possessori dei diritti d'autore e le loro associazioni, così da assicurare l'accesso dei documenti digitali all'utenza.

Con questo impegno, il progetto PARADIGMA spera di attuare completamente le intenzioni della legge sul Deposito legale relativamente a questo materiale.

### 2.4.3. La raccolta dei documenti digitali.

Come viene presentato nell'introduzione al PARADIGMA Project della Biblioteca nazionale Norvegese al Congresso internazionale dell'IFLA, che si è tenuto a Buenos Aires nell'agosto 2004,<sup>188</sup> la raccolta, sulla base del Legal Deposit Act e delle raccomandazioni di PARADIGMA, è condotta dalla B.N. No. iniziando con l'harvesting generale di tutti i documenti digitali pubblicamente disponibili nello spazio Web norvegese “.no”, ma prevedendo che, in futuro, anche i documenti reperiti su domini come “.com”, “.org” e “.net” saranno oggetto di harvesting.<sup>189</sup>

Poiché il primo scopo di Paradigma è la collezione di tutti i documenti regolati dal Deposito legale norvegese, la base dati informativa è originata da molte fonti: la raccolta Internet, la ricezione di e-mail pubblicate e le news in rete dei gruppi di discussione, i pacchetti di interi database depositati.

L'approccio all'harvesting di carattere generale si basa anche sulle considerazioni che non può essere stabilito pregiudizievolemente il valore culturale dei documenti e che, comunque, l'operazione di selezione risulterebbe più costosa, in quanto svolta manualmente; inoltre, il pensare a grandi quantità di documenti non preoccupa gli implementatori del sistema PARADIGMA, perché i costi (sia economici, sia logistici in termini di spazio) dell'archiviazione digitale vanno a diminuire nel tempo. Infine, la raccolta globale permette una maggior quantità di informazione all'utente, che, nell'effettuare una ricerca in un archivio Internet anche a testo libero, può passare in rassegna tutti i documenti. Comunque, sulle collezioni ottenute, sarà sempre possibile operare una selezione a posteriori, stabilendo i criteri descrittivi che risulteranno utili all'occorrenza.

Al momento la vera selezione è affidata alla fase successiva, o meglio, nel valore ritenuto bibliograficamente rilevante o no, che viene riconosciuto al materiale raccolto; sta nella definizione dei criteri per distinguere quei documenti, reperibili sulla rete Internet, che si ritiene valgano una descrizione bibliografica manuale.

A partire dal 2001 la Legal Deposit Section della B.N. Norvegese ha indicizzato in modo semiautomatico una selezione di documenti Web usando il software HTTrak, e i documenti sono stati descritti nel catalogo della biblioteca (BIBSYS). Il lavoro è continuato fino al momento in cui l'attività di harvesting generalizzato e le procedure collegate non sono

---

<sup>188</sup> Van Nuys et al. (2004)

<sup>189</sup> Comunque, nel citato art., si accenna al fatto che, dagli studi fatti in occasione degli harvesting condotti in Svezia e Finlandia, si presumeva di poter trovare tra il 45 e il 55 % dei siti norvegesi allocati al di fuori del dominio .no.

diventate pienamente operative. Tra le collezioni che si è deciso subito di predisporre ci sono risorse e siti Web di partiti politici in relazione ai periodi *pre-infra-post* elezioni, quindi con una selezione basata sugli eventi; poi, c'è la collaborazione tra la Library's Sound and Image Archive della B.N. No. e la Norwegian Broadcasting Corporation per il Deposito legale digitale relativo ai programmi radiotelevisivi nati in formato digitale. Altro impegno è quello della raccolta dei quotidiani Internet, che è stata avviata per 65 giornali (dati pubblicati nel 2005), con la volontà di effettuare il download di intere basi dati ad integrazione delle snapshots raccolte dal sistema automatico di PARADIGMA, quindi affrontando problemi amministrativi, giuridici e tecnici e la problematicità del deep web (nel contesto dell'IIPC), tutti aspetti non del tutto risolti.

Poiché sono molti i documenti reperibili sulla rete Internet considerabili di medio-alto valore, come s'è già detto, i responsabili del progetto stanno lavorando alla definizione di criteri di selezione per quei documenti che si ritiene debbano essere oggetto d'una descrizione bibliografica manuale: i criteri di selezione sono basati sulla legislazione relativa al Deposito legale, sulle politiche generali di raccolta come vengono formulate nel piano strategico delle collezioni e sui criteri più tradizionali delle selezioni normalmente condotte dalla biblioteca.

In PARADIGMA si vuole implementare un'architettura di sistema che consenta un procedimento in tre fasi:

- a - vengono recuperati e raccolti da Internet i documenti norvegesi e sami;
- b - i bibliotecari possono produrre automaticamente elenchi ordinati, basati su ricerche specifiche e sull'uso di vettori contenenti metadati che sono stati automaticamente estratti dai documenti raccolti;
- c - dagli elenchi ordinati i bibliotecari scelgono documenti specifici per una registrazione manuale, in base a quei criteri di selezione legali e di tradizione bibliotecaria cui sopra si è accennato.

Si prevede di far sì che il sistema possa presto essere in grado di monitorare le risorse integranti (*integrating resources*) che sono state catalogate manualmente, assistendo così i bibliotecari nel recupero e nella modifica di queste registrazioni bibliografiche, ad esempio a determinati intervalli di tempo, quando mutamenti testuali eccedono una data percentuale.

#### 2.4.4. Un archivio digitale progettato per la ricerca: FRBR

Per il sistema di ricerca dei dati nell'archivio Internet del Progetto PARADIGMA è stato preso come fondamento il modello dell'IFLA FRBR,<sup>190</sup> come dicono gli stessi relatori del progetto, «perché FRBR rappresenta il mondo reale, includendo nel proprio modello tutti gli oggetti, indipendentemente dal fatto che siano descritti in un record bibliografico o meno» e il sistema informativo voluto è fortemente orientato alla concreta funzione della ricerca da parte dell'utenza.<sup>191</sup>

Molta enfasi viene posta sull'aspetto della 'realtà': nel modello FRBR si rappresenta una "gerarchia di astrazioni" e non una "meta-gerarchia" intesa come composizione gerarchica. Così, tra le Entità del primo gruppo (v. fig. seguente), l'Item corrisponde al documento e le altre tre entità non sono "insiemi di..." rispetto al livello precedente, ma sono concetti

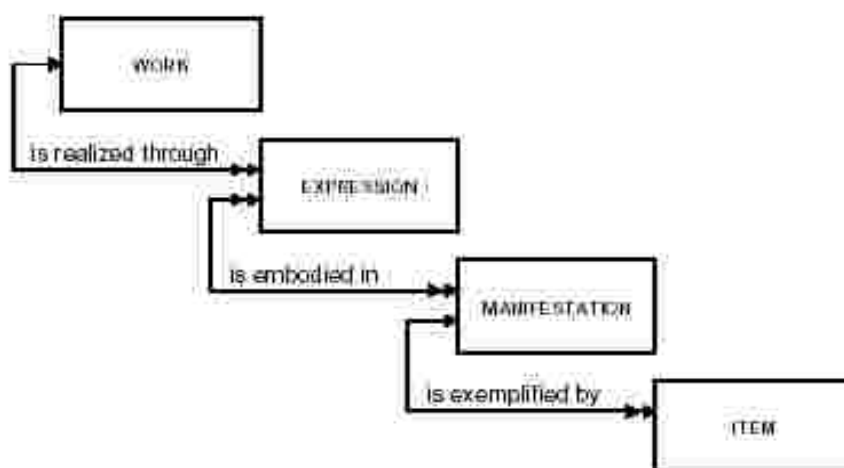
---

<sup>190</sup> Il documento dell'IFLA su FRBR è disponibile all'URL: <<http://www.ifla.org/VII/s13/frbr/frbr.pdf>>. FRBR è un modello concettuale risultato da uno studio sui requisiti minimi indispensabili per l'efficacia dei record bibliografici, che è stato pubblicato in un rapporto conclusivo nel 1998 (IFLA, *Functional Requirements...Final report*). Ha avuto origine dalla riflessione nata nell'ambito del seminario sui record bibliografici svoltosi a Stoccolma nel 1990, sotto il patrocinio dell'IFLA Universal Bibliographic Control and International MARC (UBCIM) Programme e dell'IFLA Division of Bibliographic Control, che adottò nove risoluzioni tra le quali uno studio diretto a definire i requisiti funzionali dei record bibliografici, per delineare, con l'uso di termini chiaramente definiti, le funzioni svolte da un record bibliografico rispetto ai vari mezzi di comunicazione, le varie applicazioni e i molteplici bisogni dell'utente. Lo studio aveva due obiettivi principali: il primo è quello di fornire uno schema ben definito e strutturato con chiarezza per correlare i dati che vengono registrati in record bibliografici ai bisogni degli utenti di quei record, il secondo è quello di raccomandare un livello base di funzionalità per record creati da agenzie bibliografiche nazionali. Il risultato è stata l'identificazione dell'utente come oggetto di interesse e di ricerca, e di Entità e Relazioni, univoche e distinte, come mezzi per la 'navigazione' all'interno di bibliografie, cataloghi e banche dati: da ciò è scaturita l'elaborazione di un modello di tipo E/R, quindi concettuale, per identificare i requisiti minimi della descrizione che interessano l'utente quando consulta un record in una bibliografia e in un catalogo e per definire teoricamente le finalità della registrazione e la modalità della sua strutturazione, in relazione alle tipologie dei media e alle molteplici necessità degli utenti (trovare, identificare, selezionare, ottenere accesso). Non sfugga la focalizzazione dell'interesse dello studio sulla centralità dell'utente e i meccanismi della ricerca, da cui si può ben comprendere come il modello si presta alla costruzione di un sistema bibliografico centrato sulla ricerca mediante tecnologie diversificate e avanzate. Il modello E/R è composto da entità, cui si associano degli attributi, e da relazioni tra le entità stesse. Le entità sono dieci e divise in tre gruppi: il primo gruppo comprende Opera, Espressione, Manifestazione, Item; il secondo comprende i responsabili del contenuto, della produzione, della diffusione dei prodotti: Persona, Ente; il terzo gruppo comprende i soggetti di un'opera: Concetto, Oggetto, Evento, Luogo. Ciascuna entità ha associato un insieme di attributi 'interni' (caratteristiche fisiche, aspetti formali che caratterizzano una Manifestazione) ed 'esterni' (gli identificativi della entità e le informazioni contestuali), attraverso cui gli utenti formulano richieste e, interpretando i risultati della ricerca, trovano l'entità desiderata. L'altro elemento fondante del modello FRBR sono le relazioni, che sono gli strumenti che assistono l'utente nell'esplorazione di una bibliografia, di un catalogo o di una base di dati bibliografica, con la funzione di "svelare, precisare e nominare" i rapporti tra le entità. Si potrebbe forse aggiungere anche una riflessione su una certa tradizione di attenzione ad FRBR in questa regione del Nord Europa, prendendo spunto dall'articolo di Byrum – Olivia M.A. Madison. *Riflessioni su scopi, concetti e raccomandazioni dello Studio dell'IFLA su "Functional requirements for bibliographic records"*, (in Seminario FRBR 2000), alle p. 44-45, dove si parla proprio dell'uso di FRBR nella scuola danese di biblioteconomia. Per quanto riguarda i nuovi sviluppi di FRBR si segnala anche l'applicazione Object Oriented: per es. ICOM, *CIDOC Conceptual Reference Model*. URL: <[http://cidoc.ics.forth.gr/frbr\\_inro.html](http://cidoc.ics.forth.gr/frbr_inro.html)>.

<sup>191</sup> Le citazioni di questo paragrafo sono tratte dall'articolo Van Nuys et al. (2004), che tratta specificatamente dell'uso di FRBR nell'implementazione del sistema informativo di PARADIGMA.

astratti che includono le caratteristiche comuni, e, “concretizzano – rappresentano – materializzano” l’entità di livello a loro superiore.<sup>192</sup>

Si riporta qui lo schema che illustra il Primo Gruppo di Entità, tratto da: [IFLA] IFLA Study Group on the Functional Requirements for Bibliographic Records. *Functional Requirements for Bibliographic Records: Final Report*. UBCIM Publications-New Series. Vol. 19. Munchen: K.G.Saur, 1998. [IFLA 3.1 - Figure 3.1. ]



<sup>192</sup> Sul rapporto tra FRBR e le ER si rimanda a Murtooma (1998) = Eeva Murtooma. *The impact of the functional requirements for bibliographic records recommendations on the ISBD(ER)*. In: *64th IFLA General Conference, August 16-August 21, 1998* : <<http://www.ifla.org/IV/ifla64/065-74e.htm>>. Ma anche: Byrum – Olivia M.A. Madison. *Riflessioni su scopi, concetti e raccomandazioni dello Studio sell'IFLA su "Functional requirements for bibliographic recods"*. In *Seminario FRBR (2000)*, p. 12-54, p. 36: «Le risorse elettroniche sono un tipo di pubblicazione contemporanea per la quale è particolarmente importante il problema delle versioni multiple, per via della varietà di Manifestazioni che queste risorse potrebbero assumere, compreso il caso di quelle che escono sia nel formato cartaceo sia in quello elettronico. Su questo argomento Eeva Murtooma ha scritto che FRBR «sembra sostenere il modello del record singolo per quelle risorse elettroniche che sono disponibili a distanza in formati diversi anche se il contenuto è lo stesso...». È interessante, tuttavia, come sottolinea Byrum nel suo studio *Inclusion of information covering resources in national bibliographies*, [i.e.: John Byrum, *Inclusion of Information...: Results of a Survey Conducted May-June 1998*. IFLA Conference Bangkok 1999, <<http://www.ifla.org/IV/ifla65/papers/124-153e.htm>>] che la maggior parte delle agenzie titolari di decisioni politiche in questo ambito preferisca riparare le registrazioni per le varie Manifestazioni di una pubblicazione. Negli Stati Uniti l'uso è misto e molte biblioteche seguono l'approccio del record singolo».

Continuando il distinguo tra quello che è il piano bibliografico e quello della realtà, viene considerato che l'archivio Internet è una digital library e che nel sistema informatico si trovano oggetti digitali, come il documento digitale conservato insieme ai suoi metadati; pertanto è stato creato un "agent object" che rappresenti gli oggetti di tipo non digitale, per cui esterni rispetto al sistema informatico, corrispondenti in pratica agli agenti esistenti nella realtà, come sono gli autori, i soggetti, i luoghi e gli eventi, che appunto "agendo" producono anche oggetti digitali. Sul piano strettamente bibliografico, il sistema informativo elaborato è basato sull'approfondimento dei sistemi di legame tra "Opere collegate – relativa opera" e i meccanismi per trattare raggruppamenti, o aggregazioni, di elementi.<sup>193</sup> Al concetto di raggruppamento saranno collegate poi le caratteristiche e le problematiche delle risorse elettroniche e digitali.

In particolare, in questa fase dello studio sono stati considerati come riferimento due sistemi che si sono basati su FRBR, ECHO e MAVIS.

Il progetto ECHO (del 2001) aveva come obiettivo fornire l'accesso Web di base alle collezioni di risorse audiovisive di documentazione storica ad altro valore internazionale, sviluppando un'infrastruttura SW riutilizzabile a lungo termine a supporto di archivi digitali di film e capace di supportare un servizio estensibile ed interoperabile. ECHO project ha basato i suoi metadati sul modello FRBR, estendendolo con funzioni di aggregazione per gestire relazioni fra i componenti, relativamente a nastri, pellicole dei film e gli altri materiali multimediali. In queste collezioni si trovano molti tipi di aggregati a livelli diversi, ma soprattutto a livello di Espressione.<sup>194</sup>

MAVIS è un database della società australiana Wizard Information Systems Ltd,<sup>195</sup> per la gestione di collezioni di audiovisivi; è un software adottato da molte istituzioni, tra cui lo ScreenSound Australia, the National Screen and Sound Archive dell'Australia, dalla Library of Congress Motion Picture, Broadcasting, and Recorded Sound Division e in Europa dal

---

<sup>193</sup> In G. Michetti, *Metodologie...*, in Guercio (2002), p. 231: «[ Tipi di astrazione, nella modellazione ... ] l'aggregazione è un procedimento mediante il quale si definisce una classe di oggetti integrando le caratteristiche di ognuna delle classi subordinate in una nuova classe, distinta da quelle. Ciascuna delle classi subordinate, rappresenta quindi solo una parte delle caratteristiche della classe superiore. Il concetto di documento può essere elaborato determinandone le componenti logico-strutturali (es.: data + oggetto + testo + *alia*): nessuna di esse singolarmente, bensì il loro insieme rappresenta un documento»

<sup>194</sup>ECHO (European CHronicles On-line), è un progetto finanziato dalla Commissione Europea (FP5), dal febbraio 2000 all'agosto 2005. Obiettivo del progetto era la creazione di una biblioteca digitale, accessibile via Internet/Web, contenente filmati che documentano i vari aspetti della vita sociale, economica, culturale, sportiva, etc. dei paesi europei. La biblioteca, costruita su un modello di metadati basato su IFLA-FRBR, è implementata per fornire servizi avanzati di accesso, ricerca, navigazione. Al progetto hanno partecipato: INA (Francia), TECMAT (Germania), NAA (Olanda), Istituto Luce e CNR - Istituto di Elaborazione dell'Informazione (Italia), Carnegie Mellon University (USA), Memoriav (Svizzera), EIT (Svizzera). URL: <<http://pc-erato2.iei.pi.cnr.it/echo/>>.

<sup>195</sup> URL: <<http://www.wizardis.com.au/products/mavis/whatismavis.html>>.



Norwegian Film Institute, Film Archive di Oslo. Offre un set di strumenti per gestire archivi di materiale multimediale, includendo il processo di produzione e di restauro. Il sistema MAVIS è basato su un data model che fa proprio il concetto di FRBR di una gerarchia di astrazione multi livello e fornisce funzioni di aggregazione che sono simili a quelli di ECHO, ma pone maggiore risalto al processo di gestione.

Assunto come base FRBR, del modello sono stati approfonditi i meccanismi complementari ed è stata considerata una visione 'estensiva' dei livelli d'astrazione del primo gruppo di entità:

«[Con PARADIGMA] s'introduce una nuova tassonomia, originata dalle molteplici relazioni che si formano tra i componenti aggregati, e alcune aggregazioni sono state classificate. Ciò serve per due scopi principali: possono essere introdotte nuove applicazioni senza modificazioni al modello FRBR originale e possono essere progettati meccanismi automatici per trattare ogni classe con procedure omogenee, indipendentemente dai dettagli delle relazioni semantiche».

Pensare a raggruppamenti per le Entità che si riferiscono all'opera (I Gruppo: Item, Manifestazione, Espressione, Opera) e ulteriori meccanismi aggregativi, consente a molte relazioni di essere previste in una struttura di supporto, per disporre di strumenti utili a trattare materiale che presenta particolari caratteristiche, come produzioni multimediali (come era già stato osservato nel Progetto ECHO), seriali, Opere o Espressioni composti di singole parti indipendenti, tutti elementi che si ritrovano spesso nei prodotti digitali.

Partendo dalle esigenze della ricerca, aspetto che si considera centrale nel sistema informativo, si è avvertita la forte necessità di meccanismi di aggregazione: semplificando molto i concetti, chi ricerca in un archivio digitale di risorse elettroniche remote, rischia di trovare un'eccessiva ridondanza (o anche dispersione) di oggetti e di entità, innanzitutto perché gli oggetti digitali sono per loro natura strutturale composti, e poi perché potrebbero risultare informazioni bibliografiche fortemente disaggregate su entità che, invece, sembrerebbe più logico ritrovare avvicinate da un qualche legame concettuale.

Infatti, l'utente quando ricerca in un archivio storico di Internet il più delle volte cerca i contenuti e non fa ricerche basate su altri criteri, come collocazione del documento, formato, etc.; rare volte è interessato davvero alla struttura composta del documento per ragioni tecniche o altri motivi, e, in quel caso, ne fa esplicita richiesta, quindi non avrebbe senso lasciare queste caratteristiche 'tecniche' di immediata evidenza per l'utente.

L'utente vede e considera il documento digitale come un *unicum*, invece il documento digitale è propriamente un insieme di componenti digitali diverse, che richiedono l'elaborazione di un meccanismo tecnologico complesso che le possa gestire.

Inoltre, per la ricerca sull'archivio sarebbe corretto fornire all'utente la possibilità di conoscere, per es., le diverse espressioni disponibili nelle quali si è realizzata l'opera di un'artista: ciò è possibile solo se si riflette sulla possibilità di rendere concrete aggregazioni concettuali di entità e non di carattere puramente semantico.

Sono pure necessari meccanismi di aggregazione per gestire la molteplicità degli Item che risultano dal trattamento diversificato dei componenti dei documenti digitali e riuscire, quindi, a fornire informazione esauriente sui contenuti, presentando all'utente come un'unità ricomposta (intera) documenti che in realtà consistono di diverse parti.

I casi seguenti sono riscontrabili, per lo più al livello dell'Item, dall'utente che naviga pagine Web in un archivio Internet, accedendo a documenti che sono altamente compositi:

1- Una tipica pagina Web può consistere di molti componenti separati (testo, illustrazione, fogli di stile, clips sonore, video, animazioni, etc.) e lo stesso browser Web accede ai componenti come oggetti indipendenti, e, inoltre, questi i singoli componenti possono essere inseriti in molte altre pagine Web: in questo caso gestendo questi componenti come degli insiemi, riusciamo a restituire al lettore, che richiede una copia di una pagina Web dall'archivio, un singolo documento.

2- I documenti digitali, incluse le pagine Web, sono facilmente copiabili, spesso risultando così che lo stesso documento è disponibile in molte localizzazioni, quindi occorrono procedimenti appropriati per trattare copie identiche;

3- Nel caso dei documenti Internet che hanno cambiamenti dinamici e sono aggiornati continuamente, per esempio risorse ad integrazione, continuamente modificate anche ad intervalli di pochi secondi, l'utente si rapporta al documento proprio valutandolo come "un documento dinamico", ma per la gestione occorrono meccanismi che mantengano la sequenza delle snapshots<sup>196</sup> conservate, per poterle visionare nella loro evoluzione.

4- Per le pubblicazioni digitali di tipo periodico, seriale e non ad integrazione, pubblicati tramite posta elettronica, o che compaiono come una nuova raccolta di

---

<sup>196</sup> Le istantanee, foto intermedie o 'fermo immagine', sono la prima operazione di salvataggio di un sito Web. Il termine è utilizzato anche per i programmi di grafica, a indicare una fase di lavoro che si vuole congelare nel caso in cui le variazioni operate successivamente non siano soddisfacenti, oltre alle versioni alternative di un'immagine, indica le istantanee di uno schermo o di una sequenza in movimento.

pagine sul sito Web del seriale, occorrono sempre meccanismi aggreganti che gestiscano correttamente la sequenza dei numeri.

Sono stati evidenziati dei casi nei quali meccanismi di aggregazione logica permetterebbero di districarsi anche in situazioni di eccessiva quantità di risultati della ricerca:

- o Per un utente può essere necessario una particolare tipologia di formato (HTML, Txt, JPG, TIFF, Mp3, etc.) o di risorse presenti in contesti diversi: un gruppo di Espressioni o Manifestazioni.
- o Un'Opera può essere basata su un'altra Opera.
- o Alcune Espressioni sono composite per natura: per esempio, libro con le illustrazioni di un altro artista, o una canzone, che consiste di parole e di una melodia, etc.
- o Parti di documenti, come spesso capita per le Espressioni, del tipo suono, grafica, immagini, possono partecipare come componenti di un'entità distinta, omogenea: si devono poter comprendere i due contesti di oggetti, appartenenti anche ad altri insiemi.
- o I documenti multimediali che prevedono un'associazione di diversi componenti o una sincronizzazione tra suono e immagini in movimento, i sottotitoli allo scorrere delle immagini, etc.: la sincronizzazione è molto più visibile quando è collocata su un'asse del tempo.

Le classi di entità del modello FRBR sono associate con relazioni, uno a uno (1:1) – uno a molti (1:m)– molti a molti (m:m): un'Opera è realizzata tramite una o più Espressioni, un'Espressione può essere rappresentata da una o più Manifestazioni, una Manifestazione è esemplificata da uno o più esemplare (Item) (relazione 1:n). Sono definite anche relazioni esistenti tra altre classi di entità.

Nel *FRBR Final Report* si parla di aggregazione sia per oggetti che hanno una comune astrazione, sia per oggetti che hanno una relazione in comune verso un altro oggetto (come, per esempio, i libri scritti da uno stesso autore). Per la ricerca possono essere offerti strumenti d'uso come, per esempio, delle strutture di insieme semantiche, quindi visibilmente esplicite. Nel sistema di PARADIGMA l'intenzione è quella di strutturare dei contenitori che permettano di aggregare tra loro l'enorme molteplicità degli oggetti digitali,

multimediali e Internet, sempre in vista di offrire all'utente la possibilità di una fruizione ottimale. Quindi le aggregazioni in questo sistema sono delle strutture contenitive, appunto definite 'box', che diventano delle semplici strutture che riuniscono entità senza rischiare di interferire con i valori degli attributi delle entità o di modificarle o limitarne i valori stessi.

Poiché un'aggregazione può non dipendere da uno specifico valore attribuito e non deve necessitare di valori identici, né, in nessun caso, dovrebbe richiedere modifiche alla stessa entità, visto che lo scopo è quello di preservare in un contesto archivistico le entità senza modifiche, gli aggregati restano separati e non interferiscono tra loro; per ciò è stato suggerito il paragone ad un box.

Rispetto alle aggregazioni diventa quindi utile comprenderne e definirne i caratteri: le caratteristiche delle relazioni di aggregazione e le tipologie di raggruppamenti.

Le caratteristiche dei raggruppamenti dipendono da elementi strutturali, proposti in base a quanto riscontrabile da un'analisi dei documenti Internet, come nel seguente prospetto:

Existance	l' "esistenza" degli oggetti, che si riferisce alla necessità che il gruppo ha significato in quanto è relazionato ad un oggetto esistente, per es. i diversi elementi rispetto alla pagina Web di cui sono parte
Dependence	la dipendenza, la non autonomia di taluni componenti rispetto al loro insieme
Recursion	la ripetibilità e 'ricursione', la possibilità di ricorrere in diversi livelli
Ordering	l'ordinabilità, la possibilità di ordinarli nello spazio e/o nel tempo
Referencing	la referenzialità, la possibilità di avere componenti che in realtà non fanno realmente parte del complesso oggetto digitale, ma che sono elementi esistenti altrove di cui sono solo una copia

Le aggregazioni (o raggruppamenti) sono state così classificate: *componenti, varianti, sincronici, d'aggiornamento, aggiunte, "basati su", "frammenti di"*.

Anche insiemi con caratteristiche comuni possono appartenere a classi diverse; ciascuna classe copre alcune specifiche relazioni, definite con semantiche dettagliate, anche se tutte le classi hanno identica caratteristica strutturale.

Le classi di aggregazioni identificate sono:

Component class	spesso Opere, Espressioni e Manifestazioni sono entità aggregate di entità componenti; per es. le pagine Web
Variant class	differenze di contenuti, a volte non sono molto evidenti, ma sono motivo sicuramente di selezione da parte dell'utente; per es. diverse Espressioni in lingue sono Manifestazioni di una stessa Opera, diversi formati possono rappresentare la stessa Espressione linguistica: e il fruitore sceglierà prima il linguaggio o il formato o prima il formato e poi la lingua
Synchronous class	tipiche le entità che compongono documenti multimediali e che possono essere conservate separatamente, ma che per la fruizione del documento devono essere presentate e attivate in sincronia; per es. il DVD di un film, un libro digitale con i testi a fronte in lingua originale e in traduzione
Update class	documenti ad aggiornamento, per cui l'ultimo numero sostituisce il precedente, che, quindi, potrebbe anche essere totalmente eliminato; caso tipico delle pagine Web, per le quali, a ben vedere, non c'è molta differenza tra l'archiviare le versioni precedenti e le edizioni diverse di un libro, però, soprattutto quando il numero di versioni è altissimo, per es. nel caso dei Newspapers, è più conveniente conservare un documento composito piuttosto che ogni singola versione a sé stante
Extention class	le risorse che nascono in aggiunta, come allegati; per es. i quotidiani che presentano supplementi, i blogs, i siti che ospitano il meteo, le recensioni
BasedOn	entità che svolgono il ruolo di essere basate su altre entità: essenziali in alcuni processi dei multimediali o in opere di restauro: per es., l'originale è trasferito in un film in negativo per il ritocco, un dialogo trascritto è basato sulla colonna sonora, i sottotitoli in altra lingua sono basati sulla trascrizione dei dialoghi, etc.
FragmentOn	entità-frammenti: es., in una trasmissione radiofonica, un collage sonoro può riportare solo 3 minuti di una musica di 10 minuti: se non si copia questo materiale si usa comunque un frammento che risulta parte integrante dell'originale da cui è stato estrapolato, c'è ovviamente un rapporto tra l'intero ed un frammento specificato

Tabella delle classi di aggregazione<sup>197</sup>

<i>Class</i>	<i>Cardinality</i>	<i>Existance</i>	<i>Dependence</i>	<i>Recursion</i>	<i>Ordering</i>	<i>Referencing</i>
Component	1:n	richiesta per la parte sinistra (1)	sì	sì	spazio e tempo	opzionale
Variant	.m:n	no	no	sì	nessuno	sì

<sup>197</sup> Albertsen-Van Nuys (2004), Table 1, p. 141.

Synchronous	.m:n	sì	sì	sì	tempo e spazio	sì
Update	1:n	richiesta per la parte sinistra (1)	no	no	sequenziale	no
Extention	1:n	sì	sì	no	sequenziale	sì
BasedOn	.m:n	sì	no	sì	sì	sì
FragmentOn	1:1	sì	no	sì	no	no

Tecniche e difficoltà nel raggruppare secondo i diversi livelli di FRBR (Item, Manifestation, Expression, Work):

<i>azione</i>	<i>tecnica</i>	<i>Commento</i>
Creazione degli Items (esemplari)	Quando viene analizzato un nuovo URL valido, il software di Paradigma crea un Item per rappresentare il documento presente a quell'URL. Tutti gli elementi di grafica, fogli di stile, video clip etc. sono trattati come aggregati (confronta component class)	Qualora il software rilevi, a un differente indirizzo, un contenuto del documento identico a uno già analizzato (ad esempio confrontando con un checksum MD5 <sup>198</sup> ) non viene creato un Item differente ma viene aggiunt la data e l'url all'Item precedentemente registrato
Creazione delle Manifestazioni	Appena un documento viene archiviato (cioè appena un nuovo Item viene creato) si genera una Manifestazione.	Il primo aggiornamento (soltanto per il primo aggiornamento perchè, come visto, l'update class non è ricorsiva (il secondo aggiornamento è visto come aggiornamento della risorsa già modificata e non come aggiornamento dell'originale) di una risorsa conosciuta e analizzata genera una nuova Manifestazione in relazione con la precedente (update class).
Creazione delle Espressioni	Appena viene creata una nuova Manifestazione automaticamente viene generata una nuova Espressione.	Per evitare la creazioni di una nuova estensione solo per un formato differente viene fatto un controllo estraendo il testo semplice: se il checksums è identico si tratterà di un'unica Espressione. Se il checksums non corrisponde ma vengono rispettati alcuni parametri (non specificati) non viene creata una nuova Espressione.

<sup>198</sup> Un checksum MD5 verifica l'integrità dei dati eseguendo una funzione hash sui dati stessi in seguito alla ricezione. In questo caso, il valore hash risultante viene quindi confrontato con il valore hash dei contenuti già analizzati. Se i due valori coincidono significa che il contenuto è identico.

Creazione di Work (per aggregazione delle Espressioni di un'opera)	In termini generali si può dire che per ogni nuova Espressione viene creata una nuova opera, se non viene predisposto diversamente dall'operatore (operazione manuale e quindi effettuata su una percentuale minima del totale).	L'identificazione delle espressioni di una stessa opera è piuttosto difficile e quasi impossibile da automatizzare.  Uno schiacciamento automatico è possibile nel caso in cui un documento includa metadati bibliografici dettagliati. In questo caso è possibile che il sistema proponga due espressioni come istanze di un'unica opera.
--	--	--

#### 2.4.5. I documenti Internet: dinamici e specifici

Il documento Internet è stato definito come un'unità d'informazione che può essere descritta bibliograficamente. Questa definizione non specifica intenzionalmente un insieme di componenti definite o uniche del documento, ma lascia al bibliotecario il compito d'identificare l'oggetto da descrivere: un intero sito Web può essere descritto con un record e a una particolare risorsa di quel sito può essere assegnata una descrizione. Il bibliotecario può includere o omettere suoni di sottofondo, fogli di stile etc. e può raccogliere in un documento varie pagine Web strettamente collegate, ad esempio i paragrafi di una relazione. Le procedure automatiche suggeriranno al bibliotecario definizioni del documento basate sull'analisi del contenuto, insiemi di link etc.; per default, immagini incorporate, clip audio-video direttamente collegati e fogli di stile vengono inclusi nel documento.

Ormai l'80% dei documenti Internet sono dinamici, quindi saranno i più presenti nell'archivio. Da un punto di vista formale, l'aggiornamento di una pagina Web può essere simile all'edizione di un nuovo libro: i concetti di "stampe" e di successive "edizioni" devono quindi essere ripensati nel contesto di Internet. Per i lettori la pagina di un quotidiano Internet in continua mutazione è un'entità singola, anche se variabile. Il concetto di documento dinamico è stato definito come *"l'intero ciclo di vita di una pagina Web, o di una analogo documento Internet che muta in continuazione"*. Per la catalogazione di un documento Web in aggiornamento continuo si utilizzano le AACR2, regole per le risorse integranti (CR), ossia quelle risorse bibliografiche che sono "aggiunte a" o che mutano "per mezzo di", cioè aggiornamenti che non rimangono discreti e sono integrati nel tutto, nell'insieme. I documenti quali i quotidiani Internet sono tuttavia simili ad un canale radio, un flusso continuamente mutevole di informazione transitoria, non si "integrano del tutto".

Catturare i contenuti in un documento in continuo mutamento a un dato momento è come registrare un campione di un'effimera trasmissione radiofonica; definiamo ognuno di tali campioni, o snapshots, "un documento specifico". Quando si accede ad un documento dinamico sul Web l'Item dipende da una combinazione di fattori (l'identità dell'utente, il browser Web, l'informazione relativa a precedenti accessi allo stesso documento conservati nei cookies, i parametri specificati dall'utente - per esempio in un modulo - ): spesso l'Item è generato sul momento ("on the fly") quando un utente richiede un Item, e la chiamata HTTP agisce come un servizio di "print on demand", un Item specifico derivato da un documento dinamico (che è la base di dati da dove è originato l'Item specifico), che non ha una rappresentazione permanente, ma è transitorio, a meno che non sia conservato in un archivio Internet.

Quindi un Item che esemplifica un documento dinamico come un "documento specifico" differisce da un Item tradizionale principalmente per un aspetto: è un membro di un insieme di Item che esemplificano lo stesso documento dinamico. Un documento conservato nell'archivio, o visualizzato all'utente, è ovviamente un documento specifico, ma ciò viene attenuato: una ricerca full-text ci darà per un documento dinamico tutt'al più una voce nell'elenco dei risultati. Se l'utente richiede una visualizzazione di un item, il documento dinamico viene presentato come un unità e l'utente può selezionare un Item specifico "su una linea del tempo" (timeline), ad esempio una menu line che rappresenta la durata di vita del documento. Ogni versione conservata, ossia ogni documento specifico, è indicata su questa linea del tempo con un identificatore. L'utente può avere accesso a qualsiasi documento specifico cliccando l'indicatore per una certa data o tempo, recuperando così l'Item.

Nell'archivio Internet di PARADIGMA, un documento specifico è definito nella forma in cui è stato ricevuto dal server Web; esiste un ben definito flusso di bit per ogni componente del documento (testo, immagini ecc.), la resa grafica del documento non è parte della sua definizione, questo processo è lasciato allo strumento di accesso.

Il documento specifico è identificato come il contenuto di un documento dinamico dato da componenti e metadati:

- l'origine di ciascun componente (ad esempio un URL);
- l'insieme dei parametri specificati dal client quando recupera i componenti;
- *the wall clock time* in cui ciascun componente è stato recuperato;
- il set di componenti compresi nel documento.



## 2.4.6. Metadati

Poiché si pensa di importare ed esportare i dati nell'archivio di PARADIGMA verso la B. N. Norvegese, per la descrizione di diversi tipi di materiali digitali si vorrebbero usare gli stessi metadati in uso presso la biblioteca.

Questi metadati già in uso sono: BIBSYS-MARC (il formato MARC del sistema BIBSYS) per i testi digitali, DUBLIN CORE METADATA SET, MAVIS (il sistema e formato australiano precedentemente cit.) per il materiale cinematografico, radio-televisivo, suoni e immagini, e, inoltre, altri formati usati in sistemi locali.

Questi formati di metadati sono adatti, ma non costituiscono soluzioni soddisfacenti per tutte le esigenze dell'archivio Internet, che ne richiede anche altri tipi:

- amministrativi, riguardano ad esempio la creazione e la modifica dei record di metadati;
- di gestione dei diritti e dell'accesso, per registrare informazione sul copyright e definire quali insiemi di utenti possono ottenere accesso all'archivio e quali documenti possano leggere;
- strutturali, per evidenziare le relazioni logiche tra oggetti, tra metadati, tra oggetti e metadati;
- per la conservazione a lungo termine, per la specifica, ad esempio, dei tipi di file, del SW necessario e della storia della conservazione (o migrazione) del documento;
- tecnici, per esprimere la dimensione dei documenti, script, dettagli sulla comunicazione, ecc. ecc.;
- descrittivi e analitici, per la ricerca e il recupero.

Circa il livello di descrizione che un documento digitale dovrebbe ricevere, volto a definire le caratteristiche dei metadati descrittivi e analitici, è stato scelto un modello che prevede tre livelli di catalogazione:

1. La catalogazione per inclusione nella Bibliografia nazionale e nel catalogo BIBSYS della B. N. e in altre basi dati speciali (per i pochi documenti digitali che andranno catalogati a mano, con l'uso del formato MARC, del codice AACR2 e successivi aggiornamenti, MAVIS per i materiali audiovisivi).
2. L'estrazione automatica dal documento, come da protocolli di comunicazione, etc.
3. Un livello intermedio: la catalogazione a un livello più semplice in un formato comune.

La B.N. No. progetta insieme agli editori un servizio che permetterà agli editori di produrre e distribuire metadati al momento del deposito legale; PARADIGMA sta operando per definire il formato di metadati che costituiranno le fondamenta di uno strumento amichevole. I bibliotecari potranno trattare i record di metadati forniti dagli editori utilizzandoli quali basi di registrazione bibliografiche di più alto livello.

Per specificare i requisiti dei metadati tecnici per il software del sistema d'archivio sono stati identificati i diversi fattori che possono condizionare la scelta per una catalogazione di primo livello: interoperabilità semantica con MARC; semplicità nell'uso e ricchezza di rappresentazione per un'adeguata quantità di dettaglio; facilità di conversione in altri formati, per esempio con una tabella di crosswalk di conversione con il MARC, o tra MARC21 e MODS, o tra MARC21 e ONIX, o tra MODS e Dublin Core, questi già esistenti, compatibilità con XML; estensibilità per nuovi elementi; elementi di base, come un denominatore comune che possa facilitare la ricerca e il recupero dei documenti tra i diversi tipi di materiali.

Al momento sono preferibili i formati MARC e XML compatibili, ma ancora devono essere definiti nuovi elementi per i metadati tecnici, strutturali e di gestione dei diritti e dell'accesso, forse consolidati all'interno dell'ambiente Metadata Encoding and Transmission Standard (METS); per la conservazione a lungo termine si è invece stabilito di usare Open Archival Information System (OAIS) e metadati compatibili con OAIS.

L'estrazione automatica di metadati è la soluzione dovuta anche alla considerazione realistica che «...i bibliotecari non catalogheranno mai un incredibile percentuale dei documenti interni: il 99% del nostro archivio». Quindi i metadati estratti vengono archiviati insieme agli oggetti digitali e ad altre descrizioni di metadati e resi disponibili per la ricerca strutturata nell'archivio Internet.

La tecnologia non è ancora abbastanza evoluta per decidere in modo automatico qual è il tipo di documento, ma può aiutare a ridurre la quantità di documenti cui prestare maggiore attenzione; il sistema mette a disposizione una pre-definizione automatica rispetto alle proprietà tipologiche del documento che qui di seguito sono elencate, per ottenere un risultato soddisfacente dalla ricerca nell'archivio Internet:

- lingua, vocabolario e grammatica;
- dimensione e struttura;
- fonte, editore, server Web;
- uso dei cookies;
- età e aspettative di vita di un documento;

- suono, immagini, animazioni, video e tipi evoluti di informazione;
- interazione dell'utente tramite moduli etc.;
- numero, tipo e origine dei link;
- valori dell'url, ad esempio caratteri o parole speciali;
- uso di script lato client;
- dettagli tecnici sulla comunicazione.

#### 2.4.7. Il sistema informativo

Nel 2003<sup>199</sup> è stato presentato l'ambiente di raccolta e la struttura del sistema informativo, con componenti e processi di archiviazione e gestione dei documenti.

Si tratta di una struttura completa per gestire l'archivio digitale e produrre un sistema con un controllo sull'accesso, la gestione manutenzione di formati, diverse versioni e casi tecnicamente problematici, gli strumenti per l'accesso degli utenti, etc. Il programma è stato progettato per coprire tutte le operazioni informative relative all'acquisizione, catalogazione e archiviazione a lungo termine del Deposito legale di documenti in formato digitale, perché lo scopo di PARADIGMA è allestire un sistema completo per gestire ed accedere ad una parte dei contenuti digitali e documenti nel Web e associare anche più documenti correlati o simili.

Una procedura automatica regolata alla base, estrae tutti i metadati sul documento e da altre fonti, fa una valutazione preventiva e una lista di utilità a seconda di chi richiede e di cosa si richiede. E comprende anche la possibilità di cooperare con sistemi esterni di Cataloghi di biblioteche.

Il Sistema informativo del Paradigma Project si basa su un modello UML<sup>200</sup>.

Realizza un Database centrale del Web archive, l'archivio di conservazione Digital Storage Vault (**DSM**), dove vengono archiviati tutti i documenti da parte degli utenti operatori e ricercatori, che è accessibile soltanto in lettura.

**Il DSM Web archive** riceve le risorse attraverso i seguenti componenti di Input:

---

<sup>199</sup> Van Nuys (2003). CENL, *Annual Report 2003/04 e 2005* (v. anche su *Report Coordinating digitisation in Europe*, GRN Minerva = MINERVA (2005)).

<sup>200</sup> Per UML vedi alla nota 61.

un **Harvester** (raccogliatore delle pagine Web con HTTP, FTP e altri protocolli, da Internet),

un **Email Client** (per Newsletter e Mailing list),

un **Netnews server** (server dove risiedono i Newsgroup),

un **Harvester DB**, dei Database che vengono depositati.

L'Harvester del Web accede alla rete tramite un Proxy, attraverso il protocollo HTTP o FTP.

Un crawler (o robot) parte con una serie di URL inseriti da analizzare e a partire da questi indirizzi recupera tutti i collegamenti, analizzandoli e scartando le pagine che non ricadono nell'ambito previsto dal *Deposit legal Act*.

L'Harvester DB riceve i dati dall'Harvester del Web, memorizza gli URL accettati e le eccezioni riportate nell'analisi del Web.

Sono poi previsti altri componenti del sistema: gli **Utenti**, definiti in tre tipologie: ricercatore, bibliotecario, tecnico bibliotecario o informatico.

≈ L'utente ricercatore può consultare l'**Indice** e il **DSM Web archive**, entrambi in modalità di sola lettura, tramite moduli di Browser, Archive Browser e Selector.

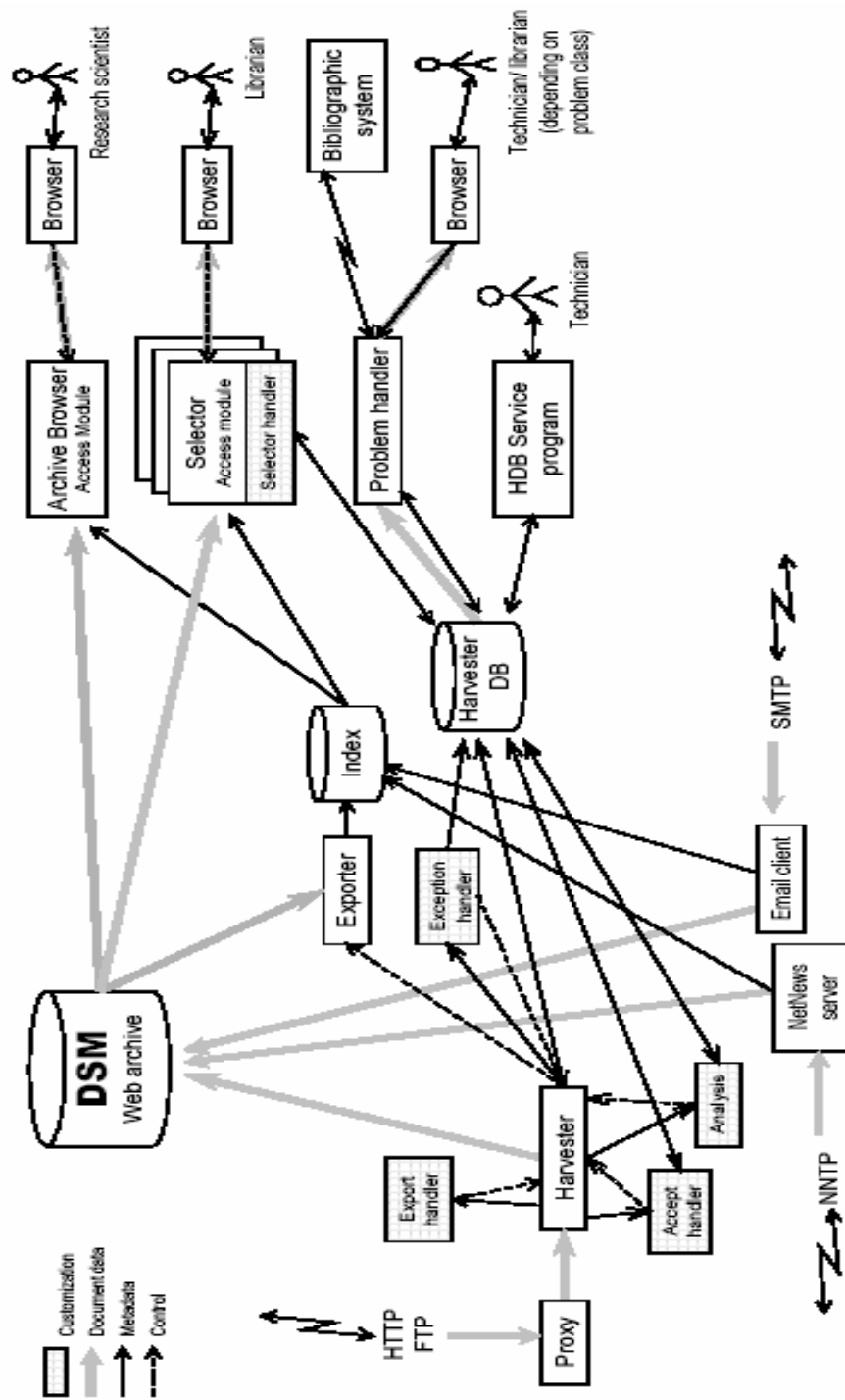
≈ L'operatore-bibliotecario può consultare l'**Indice** e il **DSM Web archive**, in modalità di sola lettura, mentre può accedere all'Harvester DB in modalità scrittura (correggere dati, inserire informazioni etc.).

≈ Il tecnico (informatico o bibliotecario) accede all'Harvester DB per la risoluzione dei problemi e delle eccezioni (Problem handler) segnalate durante l'analisi delle pagine, nonché con il Harvester DB Service program.

Il software raccogliatore è esterno; quando un oggetto viene archiviato parte dell'informazione e i metadati sono conservati nel DSM. Il database mantiene tutte le informazioni importanti per il funzionamento in Internet: l'URL, la struttura logica dei documenti, i tipi di dati riconosciuti, i codici di linguaggio, etc.; il database consta di più di 50 tavole SQL di grandezza variabile.

La raccolta Web avviene tramite crawling, con una selezione e una normalizzazione sui risultati. Un crawler inizia con una quantità iniziale di URL; gli oggetti indirizzati vengono trovati, passati in rassegna e ricercati attraverso i links che si susseguono. Poi gli oggetti vengono valutati per l'archiviazione o come base per ulteriori ricerche. Non si considerano gli URL non rientranti nella Legge sul Deposito legale. Questa procedura viene ripetuta in continuazione, scorrendo la pagina Web con tutti gli URL a cui è collegata.

Il SW proposto per il sistema: fig. tratta da PARADIGMA 2003, p. 2.



Fast è il SW di indicizzazione che opera dopo l'harvesting: nell'archivio Internet si ricerca anche a testo libero e con altri indici. Anche la piccola percentuale dei documenti catalogati manualmente diventerà disponibile tramite le registrazioni bibliografiche nel catalogo. Con l'harvesting vengono raccolti anche i metadati incorporati nella risorsa insieme ai documenti Internet che essi descrivono e la B. N. sta progettando un servizio che permetterà agli editori di generare e distribuire i metadati insieme ai relativi documenti al momento del deposito.

#### 2.4.8. Risorse raccolte e loro trattamento.

Tenendo fede al dettato della legislazione norvegese sul Deposito legale, le risorse che vengono archiviate sono le seguenti:

<i>Specifiche risorse</i>	<i>Come vengono trattate</i>
<b>Siti Web</b> (protocollo FTP e HTTP)	<p>Tutti i siti appartenenti al dominio “.no” e tutti i documenti appartenenti a domini di proprietà delle istituzioni norvegesi.</p> <p>Altri documenti (in altri domini, per esempio molti sono presenti nei domini “.fi”, “.se” e “.ru”) in norvegese o in sami (secondo alcune stime in questa categoria ricadrebbero fra il 50 e il 60% dei documenti Web).</p> <p>Alcuni documenti possono rispondere ai requisiti di selezione soltanto in parte (ad esempio una pagina con solo alcune parti o elementi in norvegese). A tal proposito è stato progettato un “algoritmo di peso” per valutare se includere o meno un documento in base ad una valutazione che prevede tre gradi (ottimo – buono – discreto) rispetto alla qualità del contenuto o per il possesso di questi fattori: dominio, grandezza, tipo di informazione. I componenti valutati possono essere conservati in un archivio a parte.</p> <p>Quando si selezionano gli oggetti per l'archiviazione e per ulteriori ricerche, l'intera pagina Web (che costituisce un documento) è considerata nel suo complesso e non come singoli oggetti Web: cioè tutti gli eventuali suoni, immagini, testi etc., vengono raccolti a priori.</p> <p>Interessante è la portata della definizione di documento: se gli URL sono in grado di reperire solo le componenti dei documenti, non i documenti interi, PARADIGMA ha la possibilità di riunire tali elementi in modo da comporre il documento intero; indicando che ciascuno di essi appartiene allo stesso insieme, è possibile disporre parti di varie pagine Web o altre unità distinte e considerarle un documento unico, se l'utente lo richiede (es. vari capitoli di un rapporto contenuti in varie pagine Web). Se il documento comprende più file, essi vengono tutti archiviati come unico documento.</p>

<b>Netnews (ovvero Usenet<sup>201</sup>) Newsgroup (basata su protocollo NNTP)</b>	<p>Vengono archiviate tutte le discussioni presenti nella gerarchia “no.” Tutti i NetNewsgroup (gruppi di Notizie in rete) sono coperti dalla legge sui depositi legali, se c’è il dominio “no.”, mentre ne sono esclusi gli altri; ad oggi non è realistico neanche farvi rientrare i contributi norvegesi contenuti in gruppi internazionali. Gli utenti norvegesi sono molto attivi sui newsgroup internazionali, ma risulta impossibile isolare i loro singoli post.</p>
<b>Mailing list / Newsletter (protocollo smtp)</b>	<p>Le mailing list e le newsletter devono essere “scoperte” manualmente (come i nuovi giornali). È necessario un controllo periodico, nell’archiviazione, per i contributi cancellati per motivi di gestione o per precisa richiesta dell’autore. Per accedere alle “mailing list” e <i>forum di discussione</i> con distribuzione un po’ più ristretta che non i NetNews, occorre una specifica richiesta-sottoscrizione, occorre il permesso dell’editore per l’archiviazione; devono essere trattati e depositati come le altre pubblicazioni digitali.</p> <p>Molte sono le Risorse ad accesso ristretto: se il livello di restrizione è alto, il documento non è considerato pubblicamente disponibile, e il Deposito legale non è applicabile; se, invece, sono a pagamento per accedere, il documento è considerato comunque legalmente archiviabile. Tutte le restrizioni vengono registrate nei metadati del DSM e ospitati nell’<i>account</i> dal sistema di controllo in accesso; viene usato un meccanismo per ogni tipo di restrizione (Pay services., X-No-Archive, Cancelled entries, Robots.txt)<sup>202</sup>, molti di questi vanno in archivi speciali.</p>
<b>Database e il deep Web</b>	<p>Occorre stipulare accordi con editori per ottenere una copia (dump) periodica del database.</p>

#### 2.4.9. Come vengono affrontati errori e casi particolari.

Nel trattamento sono considerate ‘eccezioni’ tutti i casi in cui un documento non può essere archiviato normalmente in modo automatico: tutte queste eccezioni ed errori sono riportate in maniera ordinata (ripartite per tipo: problemi di tipo legale, linguistico, tecnico) agli operatori.

Nel programma si consiglia un’ispezione giornaliera delle liste di tali errori da condursi presso la Biblioteca Nazionale.

<sup>201</sup> Usenet è una rete mondiale formata da migliaia di newsgroup; è strutturata con un ordine gerarchico.

<sup>202</sup> Pay service: servizi a pagamento; X-No-Archive è un header usato nei messaggi dei newsgroup per prevenire che un messaggio Usenet venga archiviato nei vari server, etc. Il file robots.txt contiene delle regole utilizzate dai crawler (detti anche *spider*) per applicare restrizioni di analisi sulle pagine di un sito internet.

<i>Caso specifico:</i>	<i>Come viene gestito</i>	<i>Operatori necessari per la gestione</i>
Valore di un elemento (tag) non conosciuto nell'analisi linguistica della pagina	Segnala un'eccezione e l'operatore potrà decidere se includere o escludere il valore come indicatore di linguaggio norvegese/sami	Bibliotecario e informatico
Link che dipendono da scelte dell'utente (es link dinamici)	Vakutati manualmente dall'operatore e eventualmente inseriti nella lista degli url da controllare del crawler	Bibliotecario
Cookie	Lo spider prova a visitare il sito due volte: con e senza cookie. Se il risultato è uguale il cookie viene ignorato, se è differente viene segnalata un'eccezione da controllare manualmente.	Spider Bibliotecario
Streaming media <sup>203</sup> (simile al broadcast)	Lo spider tenta di recuperare l'intero file, se non è possibile segnala un'eccezione. In quel caso si dovrà chiedere di fornire copia dei dati	Spider Ufficio legale
Tipi di file (MIME types) <sup>204</sup> non conosciuti	Registrati 400 tipi di file, nuovi file vengono riportati come eccezioni. Un tecnico deve accertarsi del plugin o del programma necessario alla lettura del file e garantire la conservazione di conseguenza	Informatico
Gestione uncini o deep url, è possibile si tratti di link di una pagina creata dinamicamente da un database	Un operatore valuta se si tratta di pagina creata dinamicamente da un database. In tal caso si attuano tutte le procedure per ottenere copia del database	Informatico Ufficio legale
Frames	Le pagine con i frame sono trattati come un unico elemento	- basta il sistema automatico
Elementi condivisi (ad esempio fogli di stile)	Vengono salvati una sola volta. Ma 'inclusi' nella definizione delle pagine collegate in modo, ad esempio, che se viene rilevato un aggiornamento dell'elemento condiviso vengono controllate anche le altre pagine	- basta il sistema automatico
Servizi, pagine a pagamento	Archivate in specifiche collezioni con limitati permessi di accesso.	- basta il sistema automatico
robots.txt	Le indicazioni dei robots.txt vengono ignorati, ma i documenti segnalati come da non indicizzare sono archiviati in specifiche collezioni con limitati permessi di accesso	- basta il sistema automatico

<sup>203</sup> Tecnologie che permettono la trasmissione continua e progressiva di flussi di informazione audio (e video), consentendo la realizzazione di vere e proprie 'stazioni trasmittenti' in tempo reale su Internet. L'utente ascolta o visiona il file contemporaneamente al suo "passaggio": le immagini o i suoni, compressi, sono inviati dal server in successione e immediatamente decompressi e visualizzati sul client: questa tecnologia permette la visione in diretta di eventi. Ciotti – Roncaglia (2004).

<sup>204</sup> MIME: *Multipurpose Internet Mail Extensions*, formato per la codifica, la trasmissione e la gestione di informazione binaria attraverso la posta elettronica o i newsgroup.



#### 2.4.10. Altre caratteristiche del sistema.

Paradigma è attrezzato per riconoscere i duplicati di un documento: di questo ne viene archiviata una singola copia sotto un'unica identificazione. Le nuove referenze si controllano nel DNS, per evitare ripetizioni; se ci sono già, questo documento è considerato una copia e non viene archiviato, le sue referenze sono però aggiunte al documento già archiviato, come varianti. Il riconoscimento di duplicati, dei documenti già archiviati e dei sinonimi delle parole chiave, può contribuire a diminuire il carico di lavoro d'aggiornamento e a prevenire i sovraccarichi di siti Web.

Per quanto riguarda l'intervallo d'aggiornamento, i programmi di raccolta di norma la eseguono in modo completo solo alcune volte d'anno, talvolta senza badare alle presenze di duplicati già archiviati e, così, molti documenti a vita breve possono sfuggire. Nel programma di PARADIGMA viene qui raccomandata un'azione continua, a bassa intensità, che riguardi ogni URL ad intervalli regolari, modificabili. Se non si trova niente di nuovo, l'intervallo fino alla ricerca successiva in quell'indirizzo viene allargato, e viceversa. La variazione nell'intervallo dell'aggiornamento che quel documento presenta a quel certo indirizzo dipenderà dalla natura del documento (es. un periodico verrà riguardato una volta al mese), se l'operatore lo riterrà necessario. Alcuni documenti del Web sono mutevoli, non ha senso aggiornarli: si cercherà, col tempo, di riconoscerli come "banche dati", ed escluderle dalla ricerca.

Programmazione a misura delle richieste o customization: è importante che anche chi non è addetto ai lavori possa ottenere un servizio adatto a lui. È facile modificare alcuni parametri di un programma, più difficile è riscrivere l'intero programma: nel programma PARADIGMA è possibile aggiungere al programma delle funzioni "su misura", composta per particolari scopi, quali percentuali di parole norvegesi, numero di termini, etc. si ha così un "vettore del documento", un insieme di dati che lo riguarda senza aver bisogno di leggerlo tutto. Può servire per decidere se quel documento è importante per l'utente o meno; il vettore non prende decisioni, ma modifica l'importanza d un documento nella lista prodotta dal programma Paradigma, a seconda delle esigenze dell'utente.

I documenti conservati in archivio vi rimangono per molto tempo, quindi, se cambiano gli strumenti di lettura, si dovrà procedere ad una conversione di essi, ma questo può causare perdita di informazioni: si raccomanda allora che il loro contenuto, come era stato preso da Internet, resti intatto. Alcuni documenti sono in formato tale da non essere

praticamente più reperibili in futuro e per questi si producono allora dei formati alternativi in testo standard puro, facilmente leggibili con minimo sforzo anche in futuro: in questo caso si perderanno alcune informazioni, ma almeno si avrà un documento quasi originale. Per estrarre il contenuto di qualche documento servirà, in futuro, qualche strumento che ora non è disponibile. Se però il documento è archiviato si può aspettare finché esso è disponibile o crearlo per conto proprio.

L'accesso per l'utenza potrebbe avvenire con l'uso di 'ticket': per accedere alle collezioni, prodotto da un sistema esterno al DMS; dovrebbe essere un codice criptato e potrebbe garantire l'accesso solo ad alcune funzioni, tipo solo lettura e non stampa ed essere prodotto da soggetti esterni alla biblioteca nazionale, magari anche a pagamento. Per agevolare i gruppi locali di studio nel leggere i documenti della biblioteca senza doverci andare di persona si potrebbe trasferire questi documenti nelle biblioteche locali in forma criptata, leggibili solo mediante un browser fornito in rete dalla biblioteca stessa. Esso non consentirebbe di stampare i testi, ma solo di vederli sullo schermo, e dopo un po' non funzionerebbe più (durata limitata). Anche qui occorrerebbe un ticket d'accesso, fornito da struttura esterna, magari a pagamento.

Per quanto riguarda lo strumento per l'accesso dell'utenza, PARADIGMA prevede che anche con normali Web browser su può accedere all'archivio, ad esempio un programma HTML consente di leggere il documento nella parte bassa dello schermo, mentre sulla parte alta ci sono i comandi e i riferimenti alle precedenti versioni dello stesso documento. Le versioni fornite sono quelle raccolte al momento della catalogazione.

Il progetto disegna una completa architettura per la gestione dei documenti in Internet, dalla loro raccolta all'accesso dell'utente, come base necessaria per la creazione di "una biblioteca di risorse Internet" davvero utile, com'è negli intenti dei responsabili. Se un bilancio sul programma deve essere ancora steso, si possono, però, ribadire i molteplici problemi che il progetto ha trovato davanti a sé e per i quali ha proposto soluzioni che ad oggi sono ancora in via di verifica. Ne vengono denunciati alcuni anche nella presentazione del progetto, pubblicata sul sito Web, e possono rendere l'idea dei problemi iniziali, pratici e teorici, di un lavoro come quello affrontato dai progettisti di PARADIGMA.

Per esempio, lo sbilanciamento tra il dettato della legge e una realtà molto 'aggravata': il lavoro preparatorio della legge del Deposito legale affonda le sue radici negli anni '80, quando i legislatori non potevano forse prevedere l'immensa portata che il patrimonio in formato digitale ha invece in pochi anni già realizzato, basti pensare all'odierna quantità di pubblicazione quotidiana sul WWW. A questa produzione si è associata un'instabilità dei

tradizionali ruoli di autori, produttori, editori e distributori, con una grande difficoltà a stabilire procedure per depositare documenti digitali autentici e validi. Molti dubbi si sono presentati su cosa andrebbe raccolto e come andrebbe registrato: sulla riflessione se il dominio Web nazionale andasse acquisito totalmente, o con qualche limitazione, e si è inserito il fattore della responsabilità di un'operazione di selezione - per quanto scientificamente coerente - rispetto alla preservazione della cultura attuale per le generazioni future.

Poi ci sono i problemi più legati all'aspetto tecnico: come considerare i documenti dinamici, sempre più frequenti, che sono costantemente aggiornati, per esempio i database e i giornali elettronici, ma anche i blog, etc., come devono essere raccolti, quali standard possono essere usati per la loro descrizione e identificazione univoca. Ultimo, ma non ultimo come priorità, la necessità inderogabile di stabilire un sistema e dei servizi per la ricerca, il recupero e l'accesso ai documenti digitali, che siano realmente efficaci e a misura dell'utente.

Nel corso della sua realizzazione il progetto è stato più volte presentato, sotto il profilo delle sue diverse componenti, per una valutazione e per sollecitare un confronto nel dibattito internazionale sul trattamento di preservazione del patrimonio digitale; alcuni di questi contributi, pubblicati tra gli anni 2003-2005 a cura dei responsabili del progetto, sono stati alla base di questa relazione, come spesso segnalato nel corso dell'esposizione del progetto e inseriti nella bibliografia.

### 3. Le risorse digitali e il Diritto: alcuni aspetti.

I problemi giuridici che presentano le nuove tipologie di risorse sono strettamente legati alle loro caratteristiche di non-fisicità e di ‘dinamicità’ nel tempo e nello spazio<sup>205</sup> (come, per esempio, la trasferibilità e la riproducibilità, che sono indubbiamente facilitate), fattori che mettono in crisi o, quanto meno, alla prova, le concezioni dei sistemi tradizionali riguardo alla tutela dei dati personali, alla garanzia di autenticità della risorsa digitale e ai diritti di proprietà, giungendo a richiedere revisioni sistematiche e attualizzazione dei competenti sistemi normativi.

Quando si parla di gestione del patrimonio culturale non si può non ricordare che esistono principi che vanno salvaguardati, come il diritto per tutti alla conoscenza e i diritti morali che l’autore<sup>206</sup> ha sulla sua opera. Tra le tante opportune riflessioni di principio mi sembra particolarmente calzante questa, di Giuseppe Corasaniti:

«La cultura non è oggetto di proprietà, è oggetto di trasmissione della conoscenza umana, e richiede perciò attenzione alle esigenze umane, tanto quelle dei creatori e dei promotori delle opere intellettuali che quelle dei fruitori. Nessuna categoria può immaginarsi autoreferente e le istituzioni pubbliche, dal canto loro, non possono apparire inerti, disinformate, o peggio dominate da interessi diversi da quelli della collettività dei cittadini».<sup>207</sup>

---

<sup>205</sup> Anche Marco Marandola nel suo contributo allo studio del diritto d’autore e la gestione e protezione delle opere in ambito digitale, con una particolare attenzione al problema della digitalizzazione - Marandola (2004) -, afferma che gestire (e proteggere) le opere in ambito digitale comporta una serie di problemi, dovuti alla facilità di diffusione, di duplicazione, di invio e in alcuni casi (come per Internet) alla sua transnazionalità, ma fa notare che questi problemi sono in conseguenza del fatto che la stessa “creazione di opere in ambito digitale” presenta caratteristiche e problematiche differenti da ogni altro supporto.

<sup>206</sup> Tra questi diritti dell’autore ci sono il diritto ad essere riconosciuto come autore e il diritto ad opporsi ad ogni modificazione dell’opera; i diritti morali sono intrasmissibili e non scadono mai.

<sup>207</sup> Corasaniti (2005). Corasaniti, magistrato Presidente del Comitato consultivo per il Diritto d’Autore del Ministero per i beni e le attività culturali, è autore della cosiddetta “bozza Corasaniti”, la bozza di decreto legislativo, elaborata dal Comitato per il riassetto delle disposizioni legislative in materia di proprietà letteraria e diritto d’autore e adeguamento alla normativa comunitaria (Direttiva 2004/48/CE del Parlamento europeo e del Consiglio del 29 aprile 2004 sul rispetto dei diritti di proprietà intellettuale contro gli illeciti, la c.d. *IPR Enforcement*) e agli accordi internazionali (la *Convention on Cybercrime* firmata a Budapest nel 2001). Testo disponibile all’URL: <<http://www.interlex.it/copyright/bozzagcor.htm>>. La bozza, scrive Manlio Cammarata nel suo articolo *La "bozza Corasaniti" è solo un piccolo passo*, “Interlex”, 03.11.05, riporta sul piano della coerenza giuridica alcuni punti oggetto nel tempo di aggiornamenti e tiene conto sia della realtà tecnologica sia del contesto internazionale, ma non risolve tutti i problemi e le contraddizioni della normativa attuale, in particolare per quanto riguarda gli aspetti penali e il netto sbilanciamento del regime dei diritti a sfavore degli utenti (soprattutto per le restrizioni al *fair use*, cioè l’uso lecito dei diritti acquisiti), restando ancora troppo legata a un modello che ha più di sessant’anni.

Nella trentaduesima sessione della Conferenza generale dell'UNESCO, il 17 ottobre 2003, è stata adottata la *Carta per la conservazione del patrimonio digitale*<sup>208</sup> che, all'art. 2, "Accesso al patrimonio digitale", afferma, come scopo della conservazione del patrimonio digitale, quello del mantenimento della sua fruibilità pubblica, per favorirne il più ampio accesso. Questo documento invita alla cooperazione internazionale tra organizzazioni e istituzioni del settore, alla mediazione tra i diversi portatori d'interessi – autori, produttori, fruitori, – per sostenere la creazione di un contesto tecnologico, giuridico e di prassi, di tutela corretta ed equilibrata per tutti gli attori/soggetti, per la realizzazione degli obiettivi che la conservazione si pone.<sup>209</sup> Conservazione, è bene qui ribadire, vista come obbligo etico nei riguardi di un patrimonio che è, come recita l'art. 1, "Portata",

«...un insieme di risorse insostituibili di conoscenza ed espressione umana. Esso comprende risorse culturali, formative, scientifiche e amministrative, come anche informazioni di natura tecnica, giuridica, medica e di altro genere, create in digitale, o convertite in forma digitale da risorse analogiche già esistenti... Questo patrimonio in continua espansione può esistere in qualunque lingua, in qualunque parte del mondo e in qualunque sfera della conoscenza o dell'espressione umana».

Pertanto, essendo così 'distribuita' e anche trasversale a diversi ambiti, la preservazione (come anche la sola gestione) del digitale incontra sulla sua strada problematiche giuridiche proprie dei diversi campi nei quali si trova ad intervenire, ove, di volta in volta, sono considerati preponderanti interessi ed esigenze del Diritto di carattere proprietario o di privacy, etc., riguardanti i contenuti, gli oggetti digitali in sé, le figure che trattano questo patrimonio, da chi ha la responsabilità intellettuale dell'opera, al fruitore e a quant'altri. Qui di seguito vengono presentate alcune questioni relative al Diritto d'autore e alle regolamentazioni sul deposito del patrimonio digitale.

---

<sup>208</sup> UNESCO. *Charter on the Preservation of the Digital Heritage*. 2003. Una traduzione italiana è stata pubblicata in appendice a Futuro memorie (2003). Il testo in lingua originale è disponibile ai seguenti URL: <[http://portal.unesco.org/ci/en/ev.php-URL\\_ID=13366&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/ci/en/ev.php-URL_ID=13366&URL_DO=DO_TOPIC&URL_SECTION=201.html)>; <<http://infolac.ucol.mx/observatorio/memoria/carta%20ingles.pdf>>. V. anche al § 1.2.1.

<sup>209</sup> Per quanto attiene alla volontà di ottemperare alla conservazione a lungo termine degli archivi della Pubblica Amministrazione e della documentazione amministrativa, in Italia un'iniziativa parlamentare è stata presentata il 20 settembre 2006 (di iniziativa del deputato Andrea Colasio, assegnata il 18 gennaio 2007 in sede referente alla I Commissione Affari Costituzionali, progetto di legge n. 1666), dal titolo "Disposizioni per la costituzione della società Banca della memoria digitale S.p.a.", che riguarda una proposta di legge che autorizzi «...il Ministero per i beni e le attività culturali a costituire una apposita società per azioni, la "Banca della memoria digitale Spa" (BMD), struttura tecnologica per la conservazione in modo protetto di tutte le informazioni numeriche memorizzate su supporti digitali e per il costante aggiornamento tecnologico dei supporti di conservazione». All'art. 1., c. 1-2, vengono enunciati scopi e obiettivi: «Per la conservazione, il mantenimento e l'aggiornamento dei supporti relativi alle basi dati informatiche degli enti dello Stato, delle amministrazioni centrali e periferiche dello Stato, delle istituzioni e di tutti i soggetti privati interessati e nel rispetto dei requisiti e delle finalità propri dei beni pubblici... Tutti gli enti pubblici, gli enti di diritto pubblico, nonché tutti i soggetti che detengono per l'esercizio della propria attività istituzionale archivi informatizzati hanno l'obbligo della conservazione e della protezione dei dati secondo i criteri di sicurezza, di accessibilità e di durata nel tempo stabiliti da un apposito protocollo approvato con decreto del Ministro per le riforme e le innovazioni nella pubblica amministrazione».

### 3.1. Il Diritto d'autore.

Si potrebbe commentare che la tipicità del patrimonio digitale, per il fatto stesso di essere non 'materiale', disorienta circa i limiti tra le legittime aspirazioni alla condivisione sociale dei contenuti e alla considerazione del riconoscimento della proprietà intellettuale, dell'equa retribuzione per ogni forma di sfruttamento commerciale, se non espressamente autorizzato da autori ed editori. Il diritto della proprietà intellettuale o IPR (Intellectual Property Right) indica il diritto che tutela la proprietà di un'opera dell'ingegno di carattere creativo, cioè il diritto d'autore. Questo costituisce l'insieme dei diritti morali (a durata illimitata) e patrimoniali (durante la vita dell'autore e 70 anni dopo la sua morte). Per copyright, termine derivato dal Diritto anglosassone,<sup>210</sup> s'intende anche il marchio corredato dell'indicazione del titolare del diritto e dell'anno a cui esso risale.

Fanno ormai pienamente parte dell'attività quotidiana l'interattività, la digitalizzazione dei contenuti, l'interscambio immediato e globale nelle reti di comunicazione, e forme di riproduzione ad uso personale sono pressoché inevitabili, ma i titolari dei diritti debbono essere sempre "conosciuti e riconosciuti". Proprio il riconoscimento delle reciproche esigenze e ragioni è la via dell'accordo per regolare una realtà ove, da un lato, è moltiplicata la possibilità di distribuzione, riproducibilità, manipolazione, in alcuni casi la non semplice tracciabilità dei trasferimenti e, quindi, la non chiara identificazione di possessori di diritti e, dall'altro, esistono meccanismi e sistemi tecnologici in grado di proteggere le opere e anche di bloccare e controllare ogni uso; l'accordo è indispensabile per trovare un giusto equilibrio fra la tutela della proprietà intellettuale e la diffusione di contenuti digitali, rispettando l'interesse pubblico all'accesso alle informazioni e per quanto riguarda l'insegnamento e la ricerca.

Già la Convenzione di Berna (1886, riveduta fino al 1971) - il testo di base su cui si fonda il sistema internazionale di protezione delle opere dell'ingegno - individuava sia diritti esclusivi, da attribuire ai creatori delle opere, sia limiti e eccezioni che possono esserci a questi diritti in nome di interessi superiori, pubblici, che riguardano la cultura, il diritto di accesso all'informazione e alla conoscenza da parte dei cittadini. Veniva intesa come

---

<sup>210</sup> Giuliana De Francesco. *Siti web e diritti di proprietà intellettuale*. In Minerva (2006), p.177-194, in particolare p. 177: «L'espressione "diritti di proprietà intellettuale" (intellectual property rights), coniata in ambito anglosassone, accomuna la industrial property e la copyright law; essa indica quindi, da una parte, i diritti relativi a marchi e segni distintivi in genere e ai brevetti per invenzione e per modelli, da noi disciplinati nell'ambito del diritto industriale, e dall'altra il diritto d'autore, ovvero l'insieme delle tutele cui sono sottoposte le opere dell'ingegno di carattere creativo. Tali tutele si organizzano in differenti sistemi nei diversi paesi; i due sistemi più diffusi sono quello di matrice latino-germanica del diritto d'autore e quello del copyright, proprio dei paesi di common law».

eccezione la riproduzione in casi speciali, possibilità sottostante alle legislazioni nazionali - che si riservano proprie regolamentazioni - e alla mancanza di danno verso uno sfruttamento normale e di pregiudizio ai legittimi interessi dell'autore.

Sulla questione degli interessi pubblici (e a più di un secolo di distanza), si riporta quanto evidenziato per giungere ad un'equilibrata contrattazione e alla migliore possibilità di fruizione delle risorse, quindi per potenziare il ruolo e il valore del servizio pubblico, da Anna Maria Mandillo:<sup>211</sup>

«Alcune esigenze appaiono prioritarie:

- seguire con maggiore attenzione l'evoluzione della normativa del diritto d'autore sia a livello nazionale sia internazionale perchè venga delineato un quadro di riferimento certo per la regolamentazione dell'accesso in rete a qualunque tipologia di opere e sia mantenuto e rafforzato il necessario equilibrio tra interessi di mercato e valori culturali;
- riconoscere uno status specifico di pubblica utilità alle istituzioni culturali, sia per i compiti di conservazione della memoria, sia per i servizi educativi e di informazione;
- stabilire tavoli di trattative, a livello nazionale e dell'UE, con le organizzazioni dei detentori dei diritti, per arrivare a stipulare contratti che, mediante uno schema di licenza, assicurino un equo e comune sistema di compensi per tutti i possibili usi delle risorse digitali. Si verifica spesso, infatti, che le istituzioni rappresentino la parte obiettivamente più debole nella fase di contrattazione della licenza d'uso: emerge pertanto la necessità della scelta del consorzio, come aggregazione idonea a rafforzare il potere contrattuale delle istituzioni culturali.
- pervenire a un comune modello di licenza-contratto d'uso a favore delle istituzioni culturali (ad esempio un contratto del Ministero per i beni e le attività culturali) condiviso anche a livello europeo, rispettoso degli interessi, pubblici e privati, delle parti in causa, e garante del diritto all'informazione e alla conoscenza da parte dei cittadini;
- promuovere accordi internazionali tra istituzioni e produttori per individuare e dare certezza agli utenti, ovunque essi siano, sulla qualità dei servizi ottenibili e sui sistemi di tariffazione;
- promuovere investimenti per progetti che sviluppino sistemi di DRM economici, facilmente gestibili, rispettosi dei diritti di accesso e di privacy degli utenti;
- favorire l'avvio di sistemi di Open Access per la diffusione della produzione culturale e scientifica mediante lo sviluppo di collezioni-depositi istituzionali (*institutional repositories*), sostenuti da software di tipo open source, con l'obiettivo di facilitare

---

<sup>211</sup> Mandillo (2005), p. 59-60.

l'accesso e l'uso dei documenti, prevedendo, comunque, meccanismi di tutela e controllo dei diritti di proprietà intellettuale. »

Ogni legislazione nazionale affronta il Diritto d'autore, cioè quelle norme che cercano i punti di equilibrio tra i diversi diritti proprietà intellettuale (che comprendono la proprietà letteraria, scientifica, industriale e tutti i diritti connessi al diritto d'autore compreso il diritto di riproduzione), secondo proprie regolamentazioni e secondo la Convenzione internazionale di Berna, che stabilisce che, per quanto riguarda la competenza territoriale, conta il luogo d'infrazione e non quello di riproduzione dell'opera. Nella tutela del diritto d'autore rientra ogni tipologia di opera, teatrale, cinematografica, musicale e letteraria; per opere letterarie s'intende ogni opera a stampa o in formato leggibile dalla macchina (programmi, banche dati, cd rom, OPAC periodici elettronici, Siti Web, Posta Elettronica, pubblicazioni liberamente disponibili, motori di ricerca e directory e ogni altro documento in Internet).<sup>212</sup>

In Italia il diritto è regolato dalla Legge 22 aprile 1941, n. 633, *Protezione del diritto d'autore e di altri diritti connessi al suo esercizio*, con successive modifiche ed integrazioni ex Direttiva Europea e D. Lgs. 9 aprile 2003 n. 68.

Nel febbraio 2005 è stata pubblicata la relazione informativa sui Contenuti digitali nell'era di Internet, *Rapporto sui contenuti digitali nell'era di internet*, elaborato dal Centro Studi del Ministro per l'innovazione e le tecnologie sulle risultanze della Commissione interministeriale (Ministro Innovazione e tecnologie, Ministero beni culturali, Ministero delle telecomunicazioni).<sup>213</sup>

---

<sup>212</sup> Per un quadro completo e aggiornato sui diritti d'autore nel settore del digitale si rimanda a Marandola (2005) e Marandola (2004). Anche Antonella De Robbio è autrice di diversi contributi sul tema dei diritti di accesso e di IPR e se ne citano qui alcuni: A. De Robbio. *Right to access to contents versus intellectual property rights in the Global Information Infrastructure = Diritto di accesso ai contenuti e diritti di proprietà intellettuale nell'infrastruttura globale dell'informazione*. In *Le risorse elettroniche: definizione, selezione e catalogazione. Atti del Convegno internazionale, Roma, 26-28 novembre 2001*. A cura di Mauro Guerrini, con la collaborazione di Stefano Gambari e Lucia Sardo. Milano: Editrice Bibliografica, 2002. Versione web, 2001: <<http://eprints.rclis.org/archive/00000047>>. Id. *Linee Guida sul copyright per le operazioni in biblioteca basate sulla legge vigente*. Venezia, Università Ca' Foscari, Dipartimento di studi storici. Ultima stesura Ottobre 2005 - Online da Ottobre 2005. WP-LIS-5: Working Papers in Library and Information Science, n. 5 (collana digitale curata da Riccardo Ridi). URL: <<http://lettere2.unive.it/ridi/wplis05.pdf>>. *Diritto d'autore: la proprietà intellettuale tra biblioteche di carta e biblioteche digitali*. A cura di Antonella De Robbio, con la collaborazione di Luisa Marquardt. Roma: AIB, Sezione Lazio, 2001. Sul tema aggiornamenti in linea si possono trovare nei seguenti siti : UNESCO. *Collection of National Copyright Laws, Europe and North America, 1995-2006*. URL: <<http://portal.unesco.org/culture/en/ev.php>>. - EUROPEAN BUREAU OF LIBRARY, INFORMATION AND DOCUMENTATION ASSOCIATIONS (EBLIDA). *European Copyright User Platform, Legislation, 1998-2006*. URL: <<http://www.eblida.org/ecup/lex>>. - ASSOC. PER LA DIFESA DEL DIRITTO D'AUTORE. *Dirittodautore.it, 1999-2005*. URL: <<http://www.dirittodautore.it/default.asp>>. InterLex. Diritto, tecnologia, informazione. Diritto d'autore, 1997-2006. URL: <<http://www.interlex.it/copyright/indice.htm>>.

<sup>213</sup> Consultabile in linea, a partire dal sito <[www.innovazione.gov.it](http://www.innovazione.gov.it)>.



Il documento esordisce con l'affermazione che la grande possibilità di un integrale trasformazione in forma digitale di ogni tipologia di contenuti, con la conquista del protocollo di trasmissione Internet e la rapida diffusione delle reti di comunicazione elettronica e della landa larga, rappresentano la base del cosiddetto "dilemma digitale", propriamente inteso come un rischio, un offrire «...scenari non del tutto rassicuranti» per i produttori di contenuti e per gli autori, nonché per i loro mercati.

Il dilemma digitale consisterebbe nella necessità di ricerca del giusto equilibrio tra diffusione dei contenuti e tutela dell'IPR nell'era della digitalizzazione: in realtà questa visione è presa da *The Digital Dilemma: Intellectual Property and the Information Age*, del National research council, USA, 2003, ma tutta la relazione italiana non delude questo preoccupato atteggiamento iniziale e si occupa per quasi il 90% della difesa dei diritti proprietari, degli aspetti economici, del mercato e del suo sviluppo, mentre la restante parte è dedicata a illustrare "i problemi" della rete e i rischi della Peer-to-Peer (abbrev. P2P). Questo contributo non può non destare alcune critiche: per esempio, dice di non voler criminalizzare le tecnologie P2P,<sup>214</sup> strutture decentralizzate, non gerarchiche, che consentono di trasferire i file direttamente tra utenti, aumentando la capacità di servizio delle reti e la distribuzione dei contenuti digitali, molto sviluppate in ambiente di ricerca scientifica, «... ma con un impatto sulla tutela del diritto d'autore» e con "effetti perversi" e, perciò, se ne auspica fortemente l'opportuna gestione, affinché da semplice strumento non diventi causa di illecito.

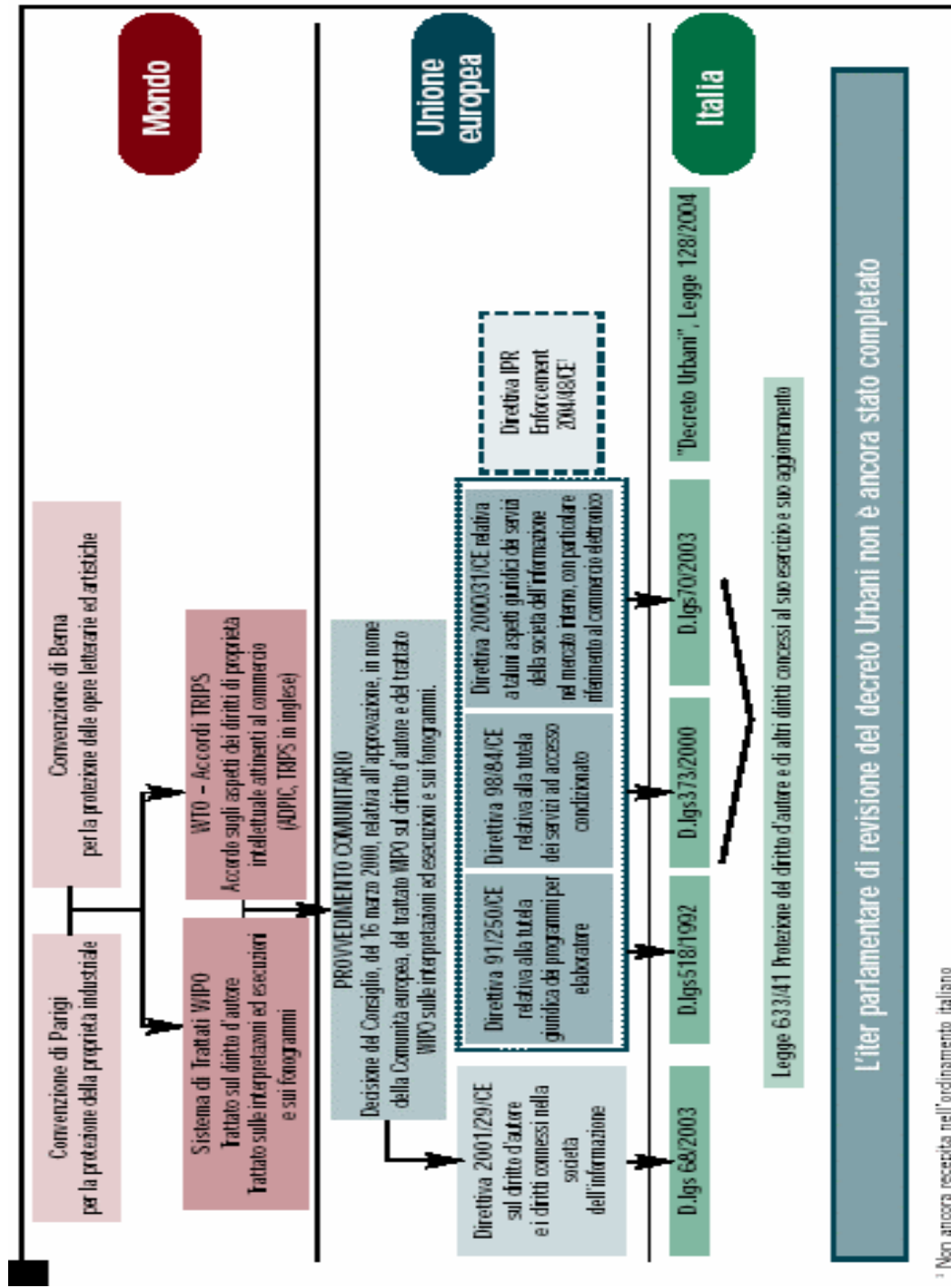
Trattando d'innovazione tecnologica, che può indubbiamente contribuire alla condivisione della conoscenza e a far sì che ogni individuo, anche in situazioni molto svantaggiate, possa prendere parte alla vita culturale partecipando alla comunità e al progresso scientifico, affermare che «la rete non può essere un "territorio senza legge", ma piuttosto tutte le iniziative e anche le misure legislative devono essere "a prova della rete"» denuncia una visione, forse, un po' prevenuta, rigida e poco lungimirante.

Tra gli allegati vi è la pubblicazione del "Patto di Sanremo", sottoscritto il 2 marzo 2005, *Linee guida per l'adozione di codici di condotta ed azioni per la diffusione dei contenuti digitali nell'era di Internet*.

---

<sup>214</sup> Sui produttori di P2P: Centre for Democracy & Tecnology. *CDT Comments: FTC December 2004 Peer-to-Peer File sharing Workshop*, (15-11-2004) <<http://www.cdt.org/copyright/20041115cdt.pdf>>. Si fa qui notare che anche Internet Archive, di cui si è trattato per la conservazione del Web, nato perché la conoscenza che viene oggi trasmessa tramite Internet non scompaia, avviene con un sistema distributivo di contenuti in modalità P2P.

Si riporta qui la tabella illustrativa redatta nel contributo del Centro Studi del Min. Innov. e Tecn.,  
*Rapporto sui contenuti digitali...*, cit., che offre una chiara idea sul quadro normativo del diritto d'autore



Nell'era di Internet la complessità del problema della gestione del diritto d'autore s'è gravata delle nuove modalità tecnologiche di trattamento e del moltiplicarsi (che è anche stato di tipo quantitativo e non solo qualitativo, rispetto alla varietà tipologica) dei soggetti interessati: autori/coautori/editori/produttori - fruitori (gli utenti) - intermediatori (gli operatori delle istituzioni culturali). Ciò ha portato all'esistenza di troppe situazioni di scarsa chiarezza, se non d'infrazione continuata o vero e proprio caos (o, almeno, di "dilemma", come sopra ricordato), nelle applicazioni quotidiane di servizi pubblici che riguardano il digitale, tanto da far dire, e ci si riferisce a due esempi, nel settore della biblioteconomia e nel campo della giurisprudenza:

A. Tammaro: «Non ci sarà biblioteca digitale se non sarà risolto il problema dei diritti d'autore»,

G. Corasaniti: «Senza produttore non ci sono garanzie... ma ogni scelta di consumo o di produzione che si basa solo sul fattore economico e che si lascia alle spalle ogni altra valutazione è destinata ad essere perdente... Il problema vero è quello delle risorse e dei diritti nella società dell'informazione».<sup>215</sup>

Le caratteristiche tipiche del materiale digitale e le modalità tecnologiche del loro trattamento, se da una parte hanno provocato controversie e problematicità di gestione, dall'altro forniscono le stesse soluzioni: per esempio, i contenuti digitali adottano moduli di consumo nuovo, che, attraverso appositi meccanismi, possono assicurare un'ancor più certa difesa dei diritti di autori ed editori.

Per quanto riguarda la riproducibilità e l'uso, già solo la consultazione di una pubblicazione digitale comporta automaticamente la sua libera circolazione all'interno del *server* e la sua riproduzione, in versioni in sostanza identiche all'originale, che non consiste solo, afferma Vitiello,<sup>216</sup> nel «... moltiplicare per via fisica l'opera attraverso meccanismi di *downloading* e *scanning* – operazione estremamente semplificata sulle reti –, ma è anche alla base dei processi di trasmissione elettronica dell'opera attraverso atti incidentali ed effimeri, che vengono cancellati una volta che l'opera è stata trasmessa».

Proporzionalmente alle accresciute facilità e velocità di riproducibilità e trasmissione di documentazione e contenuti, si è levata una richiesta d'estensione *ab libitum* dell'accessibilità alle risorse, evidenziando l'esigenza di poter copiare legalmente e diffondere materiale protetto dal diritto d'autore, quando applicato in ambito educativo, scolastico, esegetico e per altre finalità sociali, pensando di dar vita ad accordi che

---

<sup>215</sup> Tammaro (2006), p. 1683; Corasaniti (2006), p. 57.

<sup>216</sup> Vitiello (2002), p. 163. Sullo stretto rapporto tra riproducibilità e conservazione anche Salarelli, in Salarelli – Tammaro (2006), p. 33, e Giordano (2006), p. 92.

consentano, appunto, una libera fruizione. Come ovvio nella società di mercato, si sono prontamente sollevati tutti i possibili meccanismi di difesa dei diritti di produttori, editori e autori: così, anche all'interno delle reti bibliotecarie, sono state ricercate misure di contrasto alla trasmissione delle copie elettroniche delle pubblicazioni e limitazioni dell'uso, attraverso filtri tecnologici e barriere contrattuali, collocati ad ogni tappa della circolazione del documento, legando così la riscossione dei diritti non tanto al supporto quanto al consumo.

Dice ancora Vitiello:

«La chiave per la risoluzione, almeno parziale, di questi problemi passa per il rovesciamento del paradigma economico di riscossione dei diritti d'autore e per l'approvazione di misure legislative: consiste nel passaggio da una concezione legata all'unità fisica del volume (contenente una o più opere) a una visione orientata verso il consumo dell'opera, e dunque al pagamento di una licenza al detentore del diritto d'autore per il numero di consultazioni effettuate, licenze d'uso che sono quasi la regola in ambiente elettronico, con i cosiddetti contratti legati a una procedura automatizzata (*mouse-click, click-wrap*) in Internet.: il prodotto informativo offerto sulle reti di comunicazione resta inaccessibile e protetto da una barriera elettronica, qualora l'utente si rifiuti di sottostare alle condizioni della licenza d'uso. Se l'utente si rifiuta di sottostare alle condizioni poste dalla licenza d'uso, il prodotto informativo offerto sulle reti di comunicazione rimarrà inaccessibile e protetto da una barriera elettronica».

Le licenze<sup>217</sup> sono accordi, inizialmente, tra autore ed editore e, poi, tra editore e utente, che definiscono i doveri tra i contraenti, gli utenti autorizzati, gli usi consentiti o proibiti, la conservazione, il giudice in caso di discordia.

Sono stati i principi guida elaborati dalla l'Associazione inglese degli editori che hanno aperto la strada in questo campo, suggerendo possibili opzioni per l'utilizzo di opere protette e alcuni modelli di contratto per le licenze e la gestione del diritto d'autore:

≈ la vendita dell'opera nella sua integrità, in modo da liberare l'utente finale da ogni futura obbligazione o, in alternativa,

≈ il pagamento di un abbonamento ad una pubblicazione elettronica, come condizione necessaria per un numero illimitato di consultazioni effettuabili dai singoli utenti,

≈ il pagamento di un corrispettivo ogni qualvolta si acceda al lavoro protetto dal diritto d'autore,

≈ l'imposizione di una tassa periodica grazie alla quale l'utente può consultare un'opera unicamente a partire da una determinata località o da un sito predefinito (costruendo, in un certo modo, una consultazione guidata).<sup>218</sup>

---

<sup>217</sup> Valentina Comba - Vittorio Ponzani. *Licenze e diritto d'autore: la situazione europea. Intervista a Emanuela Giavarra*. "AIB Notizie", 14 (2002), 4, p. 8. URL: <<http://www.aib.it/aib/editoria/n14/02-04giavarra.htm>>.

È in fase di stipulazione di contratto per l'uso della risorsa, per esempio, che deve essere negoziato anche il diritto di archiviazione, o meglio, le clausole che permettano di conservare la copia della risorsa oggetto del contratto, definendo le specifiche e il formato di quanto potrà essere archiviato per un accesso *in perpetuum*.<sup>219</sup>

Quindi le possibilità contemplate dal diritto d'autore per l'uso delle opere sono date dai limiti regolati (come con i contratti, ad esempio licenze elettroniche, o disclaimer, o condizioni d'uso) e dalle eccezioni, che concernono l'uso libero in base ad una serie di condizioni (come l'uso personale e senza scopo di lucro), cioè quando non si crea un danno economico o di altra natura; tra gli scopi per cui è consentito fare una copia, ci sono la didattica, la ricerca scientifica, lo studio, la critica, la rassegna, la cronaca.

La tutela del diritto di proprietà intellettuale avviene attraverso il controllo sulle attività consentite all'utente (copia, diffusione, prestito, adattamento e conversioni in altri formati) e la protezione da usi non consentiti: per le risorse in Internet pubblicamente disponibili, la lettura è consentita sempre, la copia solo per studio e ricerca, la distribuzione o l'uso commerciale devono essere autorizzati; i collegamenti ipertestuali sono consentiti solo se semplici e diretti alla pagina principale di un altro sito WEB.

I Digital Right Management (DRM)<sup>220</sup> sono strumenti per la gestione dei diritti d'autore relativi a contenuti digitali, sono le misure tecniche di controllo degli accessi; per il controllo automatico delle copie digitali non autorizzate esistono sistemi come gli Electronics Copyrights Management System (ECMS), le cui sperimentazioni finora effettuate hanno tentato di delineare dei requisiti di base per garantire l'interoperabilità e la compatibilità delle transazioni effettuate elettronicamente.

La forma più semplice di controllo automatico è quella di apporre la dichiarazione del diritto d'autore ad ogni documento digitale, ma esistono anche procedure di marcature delle fonti e di criptazione dei documenti, inserendo nelle strutture del documento digitale dei metadata (Doi, Onix, Index sono standard sviluppati a questo scopo) che dichiarino i diritti di proprietà degli oggetti digitali, chi li possiede e i termini di accordo delle licenze di accesso.

Oltre a questi aspetti tecnici, vanno definite la natura delle limitazioni a tali diritti di proprietà intellettuale, ma anche le agenzie incaricate della gestione dei diritti elettronici, e, nel caos del momento attuale, alcune condizioni possono valere facilmente per l'utente individuale, ma spesso risultano meno chiare circa le modalità di percezione dei diritti d'autore o gli organismi incaricati di svolgere un ruolo.

---

<sup>218</sup> Vitiello (2002), p. 161.

<sup>219</sup> Giordano (2007).

<sup>220</sup> < <http://drm.info/>>.

Di fronte alla complessità e macchinosità e anche ai lati più difficoltosi della tutela, specie quando si toccano i problemi del controllo in relazione al Diritto alla privacy dei fruitori, si è giunti alla convinzione che un vero controllo dei diritti di proprietà intellettuale è possibile solo attraverso un atteggiamento dell'utenza educata al rispetto del diritto e una vera collaborazione tra editori e tra consorzi di istituzioni come biblioteche e Università:

«Solo l'etica può riempire quel diaframma che c'è tra la norma e la coscienza sociale, tra la consapevolezza diffusa di un'esigenza di tutela e il bisogno di conoscere che può creare, talvolta, nella comunità una domanda deviata e deviante, che non riconosce o ignora i legittimi titolari dei diritti sulle opere, autori e coautori».<sup>221</sup>

Se è vero che sono le stesse nuove forme tecnologiche a suggerire le soluzioni, se le norme hanno la capacità di adattarsi, nei nuovi modelli di diffusione dei contenuti si possono reperire meccanismi di autotutela, e, ancora G. Corasaniti, sottolinea come le tecnologie digitali e le nuove forme di comunicazione multimediale, le opportunità della rete e dell'interattività, possono garantire un sistema di distribuzione dei contenuti e assicurare nuove forme di diffusione che permettano un uso del prodotto originale, applicando un ampliamento dei diritti connessi, delle possibili licenze e delle forme di equo compenso forfetario, che possono essere accettate, se gestite con offerte che possono essere percepite come socialmente ed economicamente convenienti. A ciò si possono affiancare anche forme di autoregolamentazione, cui si può pervenire con una responsabilizzazione degli operatori, produttori di contenuti e responsabili dei servizi di comunicazione e, poiché le tecnologie interattive favoriscono il rapporto diretto autore-pubblico, le categorie interessate, in forma associativa, possono intervenire direttamente e significativamente e moderare gli squilibri nella difesa dei diritti.

Sul copyright le divergenze legislative, filosofiche, economiche e sociali rendono difficile l'elaborazione di accordi mondiali, nonostante la "transnazionalità" delle risorse elettroniche, quindi in ogni nazione vengono elaborati tipi di tutela diversi a livello contrattuale (licenze elettroniche, disclaimer di siti, condizioni di accesso) e tecnologico (sistemi di protezione contro la riproducibilità).

Evoluzioni del sistema del diritto d'autore sono state introdotte negli ultimi anni dal movimento Open Access,<sup>222</sup> con le licenze Open Content<sup>223</sup> e la più attuale Creative Commons, cioè si sono sviluppate nuove prospettive, per un accesso aperto:

---

<sup>221</sup> Corasaniti 2006, p. 54

<sup>222</sup> De Robbio 2006 (b). Tammaro (2006). L'Open Access è il movimento nato e gestito da Carl Lagoze (Cornell) e Herbert Van de Sompel (Los Alamos), finanziato da varie istituzioni, a partire dal 1991, prima ancora della diffusione del World Wide Web, all'interno di comunità di ricercatori e studiosi di fisica, allo

«... le tecnologie “aperte” che danno la possibilità agli autori di rafforzare la tutela della identità creativa e culturale nelle scelte dei metodi comunicativi, nelle strategie di interattività, nelle strategie di interattività, nelle capacità di dialogo con i potenziali consumatori e fruitori delle opere intellettuali».

Nell’ambito della circolazione dei software, per esempio, sta riscuotendo successo il copyleft,<sup>224</sup> termine che designa una prassi pienamente legittima e legale che racchiude in sé l’idea di accesso e uso liberi, a patto che riproduzione e diffusione dei materiali e dei prodotti da essi derivati e/o ispirati, non avvengano a titolo oneroso da parte di un distributore diverso dall’autore o dal detentore dei diritti economici.

Un incontro tra esigenze di accesso e necessità di protezione del diritto degli autori è costituito dai Creative Commons,<sup>225</sup> iniziativa nata nel 2001, che ha rilasciato un’ampia

---

scopo di rendere velocemente visibili i propri lavori e poterli confrontare e commentare, sganciandosi dalla tempistica, decisamente più lunga, della pubblicazione tradizionale; in seguito è stato adottato allo stesso scopo da altre comunità di studiosi. Sviluppa e promuove protocolli e standard per l’interoperabilità e la condivisione degli archivi disseminati nel mondo, e si va diffondendo come modello di diffusione democratica della conoscenza e come fattore di crescita sociale. Legata al movimento Open è l’OAI, Open Archives Initiative, <<http://www.openarchives.org/>>. Gli archivi OAI possono essere centralizzati o distribuiti, possono anche essere basati su un argomento o su una disciplina in particolare (*disciplinary repositories*), oppure possono essere di tipo istituzionale (in questo caso sono detti IR *Institutional Repository*). Sul sito dell’OAI si trova l’OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting). Sull’accesso aperto: IFLA *Statement on Open Access to Scholarly Literature and Research Documentation* (5 dicembre 2003). URL: <<http://www.ifla.org/V/cdoc/open-access04.html>>.

<sup>223</sup> Marandola (2005), p. 19-25.

<sup>224</sup> Marandola (2005), p. 3-17. Si tratta di un movimento sviluppato in ambito informatico, che si batte per la diffusione di un sistema aperto di licenze software con l’uso di contratti del tutto innovativi per regolare l’utilizzo del codice e il diritto del suo d’autore. Si fa iniziare nel 1984, con la fondazione da parte del programmatore Richard Stallman della Free Software Foundation (FSF); uno dei primi prodotti è stato il progetto GNU e la General Public License (GPL). Si veda anche Simone Aliprandi, *Copyleft & Opencontent, l’altra faccia del copyright, con i testi e i commenti delle principali licenze freeware, open source e Creative Commons*, Lodi, PrimaOra, 2005, p. 35-36: Aliprandi riporta che alcuni membri della Free Software Foundation (FSF) riportavano provocatoriamente sui loro contributi la nota «copyleft all rights reversed, ovvero copyleft - tutti i diritti rovesciati» con una C rovesciata invece del canonico copyright (C) - all rights reserved: venivano poi elencate tutte le libertà di cui l’utente era ufficialmente investito e veniva rimarcato l’obbligo di mantenerle intatte in futuro nei confronti degli altri utenti, convertendo le licenze d’uso, da decalogo degli obblighi dell’utente, in una sorta di “statuto” dei suoi diritti, intoccabili nel tempo e invariabili presso terzi.

<sup>225</sup> Marandola (2005), p. 27-39. Si segnala che i Creative Commons sono presenti in Italia già da qualche anno: a Torino, il 16 dicembre 2004, si è tenuta la presentazione dei “Creative Commons per l’Italia”, il lavoro di traduzione/adattamento all’ordinamento giuridico italiano delle licenze di Creative Commons (abbrev. Cc), che si è concretizzato nella messa a disposizione di un servizio che offre all’utente italiano un sistema di scelta di ‘licenza su misura’ tra le varie esistenti. Il progetto garantisce all’utente che naviga in rete di essere informato sul corretto utilizzo di un file musicale, un video o un’immagine o una pagina Web, senza il rischio di incorrere nella violazione dei diritti d’autore. Nel sito Web di Cc, nella sezione “Scegli una licenza”, sono presentate una serie di domande all’autore, rispondendo alle quali si arriva a una licenza “ad hoc” per la propria creazione: si può andare dall’estremo del “Pubblico Dominio” a licenze gradualmente più restrittive, che, per esempio, consentono o non consentono utilizzi commerciali, che incoraggiano le modifiche ma obbligano alla condivisione o che, invece, non consentono modifiche dell’opera. Le licenze previste sono di 6 tipi diversi:

1. obbligatorietà di citare sempre il nome dell’autore;
2. divieto di fare un uso a fini di lucro ;
3. divieto di modificare l’originale;
4. commercializzazione dell’opera presa da Internet e produzione di opere derivate;
5. possibilità di apportare eventuali modifiche ma la circolazione dell’opera deve avvenire con la stessa tipologia di licenza dell’originale (del tipo «fai usare, come tu hai usato»);
6. divieto di utilizzare il prodotto a fini commerciali.

gamma di licenze che spaziano dalla protezione totale all'assenza di diritti riservati: tra queste c'è un insieme di licenze di copyright che sono libere per uso pubblico e rendono possibile la condivisione delle opere, o destinandole interamente al pubblico dominio, o dando una licenza ad uso libero a condizioni e per usi determinati; sono le licenze utilizzate dal Modello OAI, finalizzate a dichiarare gli usi consentiti e studiate per permettere un dialogo tra le esigenze dell'utenza, la tutela dei diritti d'autore e il computer, in uno schema di metadati e con l'apposito linguaggio informatico ODRL<sup>226</sup> (Open Digital Rights Language): è insomma necessario tradurre tutte le documentazioni del diritto con un linguaggio giuridico specifico, non fraintendibile, comprensibile dalla macchina.

---

Si veda all'URL: <<http://www.creativecommons.it/Licenze/Spiegazione>>. L'adattamento italiano è opera dello staff Cc di San Francisco e Berlino che ha lavorato con il professor Marco Ricolfi del Dipartimento di Scienze Giuridiche dell'Università di Torino; dal punto di vista tecnico-organizzativo ha collaborato l'Istituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni (IEIIT-CNR), già dal 2003 istituzione italiana affiliata di Cc, che si è assunta l'incarico di tradurre e adattare i contratti delle licenze dall'originario inglese. Al convegno era seguito, il 18 dicembre 2004 presso l'Università IULM, il convegno 'Condividi la conoscenza: nuovi Commons, nuovi diritti' con Stefano Rodotà e Lawrence Lessig, docente di Diritto presso l'Università di Stanford e ideatore dei Cc, dove si era affrontato il problema della possibilità di condivisione della conoscenza, dell'opportunità di diffusione del Sapere, delle opere dell'ingegno e della creatività umana, alla portata e a vantaggio di tutti, e quindi del valore dei "beni pubblici". Per il suo creatore Cc è un progetto che mira a creare una nuova concezione e nuove regole del 'copyright', più modulabile e adattato alle esigenze del World Wide Web: Commons, i 'beni comuni' da mettere a disposizione attraverso una nuova concezione del brevetto e del diritto d'autore, sono i beni prodotti dall'inventiva umana, il cui valore aumenta a seconda di quanto questi vengano adottati, utilizzati da altri per fare altre cose, "come accade a una lingua che acquista valore quante più persone la parlano". "We use private rights to create public goods", si può leggere nel sito statunitense: poche efficaci parole per definire un intero programma: Cc non è volto a cambiare le leggi sul copyright, dato che nell'ambito delle stesse sono previste per l'autore di un'opera facoltà quali quella di distribuire l'opera nei modi e nelle forme preferite e addirittura nel "pubblico dominio", ma cerca di affiancare al classico modello "All rights reserved" ("Tutti i diritti riservati") la filosofia del "Some rights reserved": così, solo alcuni diritti - per esempio lo sfruttamento commerciale di un brano musicale all'interno di uno spot pubblicitario - restano riservati al titolare dell'opera, mentre è libera ad esempio la distribuzione della stessa senza fini di lucro (ad esempio tramite sistemi P2P o siti Web) e in molti casi la modifica (manipolazioni grafiche, remix di brani musicali, adattamenti di testi letterari o di codici software e così via). Stefano Rodotà ha ribadito nei suoi interventi sul tema (anche in occasione della Conferenza internazionale "*Cultural Heritage on-line. The challenge of accessibility and preservation*", tenutasi a Firenze nel dicembre 2006, URL: <<http://www.rinascimento-digitale.it/indexEN.php?SEZ=396>>), la necessità di rafforzare queste soluzioni alternative e profondamente stimolanti, dal punto di vista culturale, sociale, ma anche secondo una logica di mercato, in quanto, solo capovolgendo logiche che stanno divenendo fortemente limitanti il naturale sviluppo della conoscenza, si possono realizzare i progressi e i vantaggi culturali, in fine anche di tipo economico, per la società.

<sup>226</sup> <<http://odrl.net/>>. L'Open Digital Rights Language (ODRL) Initiative è l'iniziativa internazionale volta allo sviluppo e alla promozione degli standard liberi per l'espressione dei diritti; ODRL intende offrire meccanismi flessibili e interoperabili a supporto di un uso innovativo e trasparente dei contenuti digitali pubblici, in tutti i settori e le comunità che pubblicano, distribuiscono e usano media digitali.



### 3.2. Il deposito legale

Oltre alle licenze d'uso e a riproducibilità, prestito, etc. permessi in base a speciali e determinate deroghe, il Diritto d'autore incontra una terza possibilità d'eccezione:<sup>227</sup> il deposito legale.

Giuseppe Vitiello affronta il capitolo del deposito legale nel suo volume *Alessandrie d'Europa*,<sup>228</sup> in quanto quest'istituto rappresenta lo strumento della raccolta del patrimonio documentario nazionale, che, assieme alla produzione della bibliografia nazionale, rappresenta quella funzione che costituisce «... il requisito minimo per l'esistenza di una biblioteca nazionale». Riconosciuto quindi come lo strumento primario delle politiche d'acquisizione, che permette di raggiungere l'obiettivo di fissare la memoria documentaria di una nazione, quindi con un valore di «... sincronia e diacronia», può essere opportuno riportare queste riflessioni in un periodo, come quello odierno, e particolarmente quello italiano, di faticosa modernizzazione di questo strumento antico, che nasce con scopi strettamente culturali, presto diviene mezzo di censura e controllo, e forse, nell'era di Internet e della ICT,<sup>229</sup> la “Società della conoscenza”, potrebbe finalmente assumere quel ruolo di dispositivo finalizzato esclusivamente e pienamente alla documentazione, svincolato dai più pesanti limiti dei diritti proprietari.

Per meglio comprendere le implicazioni di cui può essere portatore, può essere utile riassumere brevemente le diverse evoluzioni che quest'istituto ha seguito nel tempo.

Il primo atto di obbligo di consegna di una copia compare in Europa con l'ordinanza del 1537 di Francesco I, e risponde ad una visione traboccante di spirito umanista che, con lo scopo di arricchire il sapere dei letterati che lo circondavano, permise al re francese di raccogliere prestigiose collezioni librarie; a questo atto, a Oxford nei primi anni del '600, seguì l'accordo per la Bodleian Library con l'associazione dei tipografi per ricevere ogni nuova pubblicazione, e questa volta precisamente a scopo d'incremento del patrimonio della biblioteca. Ma già nel corso del XVII sec., si estenderà in molti paesi europei<sup>230</sup> il decreto di

---

<sup>227</sup> Parlano di “limitazioni”, “eccezioni”, “esenzioni” al diritto d'autore Vitiello (2003), Mandillo (2005), Marandola (2004) e (2005); AIB (2005).

<sup>228</sup> Da Vitiello (2002) sono state attinte molte informazioni qui utilizzate e anche i passi riportati.

<sup>229</sup> Acronimo di Information and Communications Technology, viene utilizzato per indicare le tecnologie dell'Informatica e della Telematica, convergenti per nuovi modi di trasmettere l'informazione.

<sup>230</sup> Esempio diverso sarà quello dell'Olanda, ove dal 1912 il deposito è attuato unicamente in modo volontario. Vitiello (2002), p. 59. G. Vitiello si è già occupato dell'argomento in G. Vitiello. *Il deposito legale nell'Europa comunitaria*. Milano: Bibliografica, 1994.

obbligo di consegna, così facile da applicare in una società che non conosceva libertà di commercio né tanto meno il Diritto d'autore, obbligo che vacillerà tra l'essere strumento di arricchimento delle collezioni e di scambio, o «...di autocontrollo» per i programmi e le dispute universitarie (è questo l'esempio dell'ordinanza dell'Università di Copenaghen, nel regno di Danimarca e Norvegia, del 1623) e l'assunzione – che poi prevarrà definitivamente – di un ruolo di vero e proprio istituto di controllo, e quindi anche di censura, sulla produzione culturale.

Vitiello dice ancora

«Il deposito legale, in particolare quando concerne i documenti elettronici, va considerato oggi come una delle condizioni necessarie per l'esercizio della libertà d'accesso all'informazione; esso garantisce l'esistenza di un archivio di riferimento, dove la memoria editoriale di un paese rimane sempre a disposizione dei cittadini»,

cosicché, da strumento del controllo censorio, il deposito diventa un elemento di supporto della libertà d'informazione.

In più, mentre in un periodo precedente il deposito legale consentiva l'esistenza dello stato di editore e di autore, quindi era la condizione preliminare all'acquisizione del diritto d'autore, nell'epoca delle risorse digitali la risoluzione del Diritto d'autore è il preliminare per l'esistenza del Deposito legale, che viene a coincidere con la preservazione del patrimonio. Pertanto, un'efficace gestione del deposito legale è la scelta politica di investire sulla conoscenza digitale, favorendo una diffusione dell'informazione elettronica attraverso le strutture pubbliche e non unicamente sulla base del consumo privato. Ovviamente, oltre al collegamento tra deposito e preservazione, occorre porre l'accento sul legame con i servizi bibliografici: quindi promuovere una visione di “deposito-bibliografia-catalogo-preservazione-pubblico accesso sul patrimonio”.

In alcuni Stati europei (Germania, Paesi Bassi, Danimarca, Finlandia, Svezia) esistono esempi di cooperazione tra il mondo dell'editoria e le istituzioni pubbliche deputate alle funzioni di conservazione/documentazione nel tempo, per il Deposito legale, che raggiunge alti gradi d'efficacia anche quando si basa sul deposito volontario, e ciò è dovuto al pressoché immediato inserimento nella bibliografia e nel catalogo collettivo nazionali delle risorse acquisite; nel caso danese, per esempio, gli editori mostrano un forte interesse per la bibliografia nazionale, in quanto permette loro di penetrare il mercato bibliotecario, per cui spediscono le pubblicazioni al centro bibliografico prima ancora che esse siano disponibili in commercio. In pratica, una buona progettazione dei passaggi rende così efficiente il servizio pubblico che, proprio per questo, diviene davvero utile, finanche ai “giganti” dell'editoria, che addirittura lo richiedono.

Per rendere l'idea di come la legislazione e gli obiettivi perseguiti attraverso il deposito legale siano anche responsabili dell'atteggiamento dei produttori, editori, tipografi e autori, e, di conseguenza, anche dello stesso risultato finale per i centri deputati alla conservazione del patrimonio e al servizio bibliografico, si riporta qui di seguito la significativa tavola elaborata da Vitiello (op. cit.):

<i>Obiettivi del deposito legale in Europa dal XVI al XX sec.</i>	<i>Periodo pertinente</i>	<i>Atteggiamento delle categorie depositanti</i>
Censura	XVII – XX sec.	Negativo
Conservazione	XVI sec. – oggi	Positivo
Servizi bibliografici nazionali	1950 – oggi	Positivo
Diritto d'autore: deposito come titolo costitutivo (cioè condizione preliminare all'acquisizione del Diritto d'autore)	1709-1886	Negativo
Diritto d'autore: accesso alle pubblicazioni elettroniche	1996 – oggi	Negativo o fortemente restrittivo

### 3.2.1. Deposito legale delle pubblicazioni elettroniche.

L'avvento delle nuove tecnologie dell'informazione e della comunicazione moltiplica vorticosamente il numero dei prodotti culturali su supporto digitale, favorendo in modo eccezionale la diffusione su vasta scala soprattutto dei documenti d'interesse scientifico e di ricerca. Se tali pubblicazioni non entrano a far parte della collezione nazionale, risultano irrecuperabili per la memoria di una nazione e ciò rende oramai fondamentale la loro conservazione e la loro accessibilità nel tempo.

Ma, tenuto conto anche della varietà delle tipologie delle risorse documentali in formato elettronico, disponibili su supporto informatico e direttamente in Internet, off-line e on-line, professionali e occasionali, elettroniche fin dalla loro creazione (born digital) e digitalizzate, statiche e dinamiche, varietà che apporta ulteriore complessità gestionale, nella regolamentazione del deposito delle pubblicazioni elettroniche i problemi giuridici riguardanti l'accesso rappresentano il fattore maggiormente critico, superando i problemi tecnici e quelli organizzativi.

È ovvio che «alimentare, anche con grandi investimenti effettuati per la realizzazione delle infrastrutture tecnologiche, immagazzinamenti di collezioni nazionali potenzialmente “morte” perchè i documenti non sono accessibili»<sup>231</sup> e, di fatto, essere impedita la diffusione elettronica del contenuto, è ben poco comprensibile, ma in realtà il mondo dell’editoria ha apposto una vera resistenza alla libera disponibilità on-line.

Per forza contrattuale e autorità riconosciuta, molti ritengono che una Biblioteca Nazionale, che svolga un ruolo effettivo di coordinamento e di leadership all’interno dei consorzi di biblioteche, cioè un’istituzione depositaria rafforzata nelle sue prerogative e non indebolita, sarebbe in grado di gestire meglio una negoziazione con i produttori e le lobbies editoriali. La trattativa sarebbe equilibrata se condotta in un quadro normativo di legame organico tra regolamentazione sul deposito legale e disposizione sul diritto d’autore, in cui siano previste clausole riguardanti sia l’acquisizione conservativa del documento elettronico, sia le relative politiche d’accesso.

Per una situazione corretta e favorevole della diffusione dell’informazione elettronica attraverso le strutture pubbliche e la conservazione nel tempo del patrimonio digitale, poco prima dell’emanazione della Direttiva dell’Unione Europea sull’armonizzazione del copyright e dei diritti connessi nella società dell’informazione<sup>232</sup> e della nuova legge sul deposito legale in Italia, Vitiello, proponeva le condizioni giuridiche preliminari:

- che l’istituzione depositaria abbia la possibilità di riprodurre il documento elettronico originale online per poterlo poi immagazzinare nel suo server (gli editori, con i costi di aggiornamento costante per i SW di gestione e dei formati di conservazione, non hanno interesse a mantenere un costoso archivio elettronico di contenuti a fini di conservazione); questa minima deroga alle minime condizioni, tuttavia, «... riuscirebbe unicamente a garantire che la circolazione delle pubblicazioni avvenga *intra mœnia* sulle workstations disponibili nella biblioteca nazionale»;
- che la biblioteca depositaria avvii trattative per ottenere una licenza d’uso che non limiti il diritto del pubblico accesso all’informazione, istituendo a tale scopo dei clearing centres che si occupino dell’obbligo di deposito e delle trattative sulle condizioni d’accesso.

---

<sup>231</sup> Vitiello (2002), p. 202.

<sup>232</sup> Decreto legislativo 9 aprile 2003, n. 68, attuazione della direttiva 2001/29/CE, *Sull’armonizzazione di taluni aspetti del diritto d’autore e dei diritti connessi nella società dell’informazione.*

In effetti, come prevedibile date le dimensioni della produzione editoriale in Internet, già dal 2000 in molti paesi europei è stato disciplinato il Deposito legale e dato avvio alla raccolta della documentazione per questa tipologia di materiale. Gli esempi più riusciti sono collocati soprattutto nelle regioni del Nord Europa:

- il Legal deposit Act in Norvegia<sup>233</sup>, del 1989;
- in Finlandia, dal 1999;
- in Danimarca, dove la raccolta è di tipo selettivo e ragionato sulla base del contenuto, in genere secondo criteri già applicati nel campo dei documenti audiovisivi, dal 2001;
- in Svezia, con il decreto del 2002, che autorizza la Biblioteca Nazionale di Stoccolma all'harvesting;
- il Legal deposit Act, in Gran Bretagna, del 2003;
- in Francia, dove si sono sperimentate modalità di raccolta automatica di documenti digitali attraverso harvesting, soprattutto in relazione al deep web, con siti accessibili attraverso moduli di ricerca non disponibili nella loro forma completa ai motori di ricerca tradizionali;
- in Olanda, in cui non esiste una legge sul deposito legale, ma il sistema si basa su accordi di deposito volontario.

L'harvesting viene effettuato sempre dalle rispettive Biblioteche Nazionali e, fatta eccezione per la Danimarca che ha un approccio selettivo, consiste nella raccolta totale, cioè scaricando negli archivi elettronici qualunque pubblicazione sia disponibile sulla rete Internet, sulla base di un criterio di nazionalità negli URL, comprendendo quindi tutte le risorse che ricadono nel dominio nazionale o i siti Web collegati in qualche modo alla cultura nazionale.<sup>234</sup>

Tra la documentazione più qualificata sulla materia, prodotta nelle esperienze internazionali di settore, si segnalano:

- le *Guidelines for legal deposit legislation*, pubblicate nel 1981, promosse dall'UNESCO e dall'IFLA (International Federation Libraries Associations), elaborate da Jean Lunn, della Biblioteca di diritto dell'Università di Ottawa, Canada;

---

<sup>233</sup> V. nel capitolo 2.3.

<sup>234</sup> Fonte: De Robbio(2006 a): Antonella De Robbio, giudicando positivamente quest'immagazzinamento di tipo automatico «..che, non solo non ispira comportamenti volontari di censura sul materiale da conservare, ma non implica neppure dispendiose procedure di selezione», vede come probabile che queste modalità di raccolta possano facilmente diventare la pratica di riferimento per le attività di deposito legale relative alle pubblicazioni elettroniche.

l'edizione più recente si trova anche on-line: *Guidelines for legal deposit legislation*. 2000. URL: <<http://www.ifla.org/VII/s1/gnl/legaldep1.htm> >. (Il capitolo 6 è dedicato alle pubblicazioni elettroniche);

- lo *Statement on codes of practice for the voluntary deposit of electronics publication*, accordo tra la Conferenza delle Biblioteche Nazionali (CENL) e la Federazione degli Editori Europei (FEP), il quale delinea un codice di comportamento per il deposito volontario delle opere elettroniche;
- la *Carta per la conservazione del patrimonio digitale* dell'UNESCO (già cit.), che ha riconosciuto ai documenti in formato digitale lo stesso valore di quelli in formato o supporto tradizionale;
- i documenti dell'International Internet Preservation Consortium (IIPC) al quale partecipano oltre ad importanti biblioteche nazionali di diversi Paesi anche il progetto "Internet Archive".

### 3.2.2. La nuova norma italiana sul deposito legale.

Il deposito legale delle pubblicazioni italiane è regolato dalla Legge 15 aprile 2004, n. 106, *Norme relative al deposito legale dei documenti di interesse culturale destinati all'uso pubblico*, e dal decreto del Presidente della Repubblica 3 maggio 2006, n. 252, *Regolamento recante norme in materia di deposito legale dei documenti di interesse culturale destinati all'uso pubblico*,<sup>235</sup> che attuano una nuova disciplina destinando direttamente alle Biblioteche nazionali centrali di Firenze e di Roma, agli istituti individuati e agli archivi regionali depositari, il deposito delle copie d'obbligo delle pubblicazioni italiane, di qualsiasi natura esse siano, ma rimanda ad una futura disciplina per quanto riguarda i documenti diffusi tramite rete informatica.

La nuova legge del 2004 è stata originata da una proposta della Direzione Generale per i Beni Librari e le Istituzioni culturali, del Ministero per i Beni e le Attività Culturali. Viene a sostituire una legislazione sulla materia durata circa 60 anni,<sup>236</sup> per gestire una produzione

---

<sup>235</sup> L. 106/2004, pubblicata nella Gazzetta Ufficiale 27 aprile 2004, n. 98, e Regolamento attuativo 252/2006, pubblicato nella Gazzetta Ufficiale del 18 agosto 2006 ed entrato in vigore 2 settembre 2006, sostituiscono la precedente L. 2 febbraio 1939, n. 374, come modificata dal decreto legislativo luogotenenziale 31 agosto 1955, n. 660, e del relativo regolamento attuativo (R.D. 12 dicembre 1940, n. 2052).

<sup>236</sup> Vitiello osserva che la raccolta per deposito legale mostra un tasso di evasione elevato fra gli editori e i produttori di materiale non librario, per diversi motivi (ignoranza della legge, errata politica di raccolta inadeguata alla maggiore produzione, scarsa vigilanza). Il tasso di efficacia riguardante le monografie è molto

editoriale molto accresciuta per quantità e prodotta in una notevole varietà di formati (la legge riconosce sedici categorie di documenti destinati al deposito) ed è configurata con un chiaro proposito: l'obbligo del deposito per la creazione di un archivio di livello nazionale e di uno di livello regionale dei documenti destinati all'uso pubblico, in un quadro di conservazione "*della memoria della cultura e della vita sociale*" italiana (art. 1, c. 1), differenziandosi, quindi, dalla precedente norma, risalente al 1939, che, nel pieno di assai diverse condizioni culturali e sociali, rispondeva più ad un'esigenza di controllo preventivo sulla stampa (i prodotti editoriali dovevano essere consegnati alle Prefetture, dipendenti dal Ministero dell'Interno) e non ha favorito la crescita di una consapevolezza del valore del deposito, fattore, invece, determinante per la riuscita dell'istituto del deposito stesso.

La legge, all'art. 1, c. 2, fissa con precisione questi obiettivi:

1. la costituzione dell'archivio nazionale e regionale della produzione editoriale
2. la realizzazione di servizi bibliografici nazionali di informazione
3. l'accesso ai documenti oggetto di deposito legale.

Definisce (art. 2) chiaramente le finalità del deposito legale come:

1. la raccolta e la conservazione dei documenti;
2. la produzione e la diffusione dei servizi bibliografici nazionali;
3. la consultazione e disponibilità dei medesimi documenti, nel rispetto delle norme sul diritto d'autore e sui diritti connessi, nonché sull'abusiva riproduzione di opere librarie;
4. la documentazione della produzione editoriale a livello regionale.

La norma dà molto rilievo al termine "uso pubblico", inteso come distribuzione, immissione in circolazione, in commercio, diffusione al pubblico dei documenti, anche tramite reti informatiche. L'"uso pubblico" è l'elemento che accentra in sé l'obbligo di deposito legale, che è costitutivo dell'obbligo stesso: il documento è oggetto di deposito legale in quanto è destinato all'uso pubblico.

È indiscutibile un certo ritardo delle istituzioni nell'attuazione di nuove norme per assolvere alle esigenze culturali di un Paese che fa parte di una comunità, come quella europea, mediamente già aggiornata su questo fronte, anche solo notando che l'indispensabile decreto attuativo della legge, il D.P.R. 252/2006, è appunto stato emesso due anni dopo la legge. Ma il ritardo appare ancor più evidente per il fatto che il regolamento

---

basso, intorno al 50-70 % in Italia e Grecia, scende ancora per i periodici ed è decisamente inferiore per tutte le tipologie di materiale non librario (contro l'85-90 % di tutti i paesi di Francia, Germania e Gran Bretagna e in tutto il Nord Europa). Vitiello (2003), p. 196.

sulla parte della legge che riguarda l'acquisizione del digitale on-line, ciò che rientra nei formati elettronici e multimediali del patrimonio, è stato ulteriormente rimandato e senza limiti di tempo, indizio della mancanza di una risposta pronta e sicura per la gestione di questo materiale che, comunque, ha invaso tanto il consumo quanto la stessa produzione editoriale e istituzionale italiana ormai da almeno quindici anni.<sup>237</sup>

Vi è, probabilmente, la consapevolezza di disporre sul territorio nazionale di una situazione strutturale disomogenea e non del tutto capace di gestire con certezza la preservazione e l'accesso, così come di non possedere una sufficiente possibilità di dominio sui problemi relativi al copyright, rispetto alle esigenze culturali del Paese di documentare la produzione intellettuale.

Antonella De Robbio, in una relazione<sup>238</sup> sulla legge italiana e sul relativo regolamento di attuazione, sottolinea che tra gli articoli relativi ai documenti distribuiti su supporto informatico e diffusi tramite rete informatica (il termine più corretto 'distribuzione' è della De Robbio, mentre il legislatore utilizza indifferentemente 'diffusione' sia che si tratti di risorse su supporto sia di risorse su rete), il regolamento ha inserito un articolo sull'accessibilità - considerato evidentemente ovvio per le altre tipologie di materiale - «...nel tentativo di conciliare le disposizioni a tutela dei diritti d'autore». Tentativo non pienamente riuscito, dal momento che vengono rilevate queste debolezze nel testo:

- si distingue tra “documenti su supporti informatici già accessibili in rete liberamente e documenti accessibili per via telematica soggetti a sottoscrizioni o modalità di acquisto a pagamento”, come se alle due tipologie di documento corrispondesse sempre, e in questo ordine, la differenza tra “libero” e “soggetto a condizioni”;

- per i documenti in origine accessibili a determinate condizioni, l'archivio di conservazione permetterà l'accessibilità solo ad utenti registrati e che accedono da postazioni situate all'interno degli istituti di deposito, in pratica limitando persone e luoghi.

In questo forzato e non chiaro distinguo tra risorse libere e risorse a determinate condizioni, si nota un'insufficiente determinazione nel chiarire il diritto all'accesso alle risorse: gli articoli dedicati all'accessibilità dei documenti su supporto informatico e distribuiti in rete (artt. 33 e 38), permetterebbero generiche e deboli modalità d'eccezione ai

---

<sup>237</sup> Ponzani (v. succ. nota 216) sottolinea con preoccupazione questa carenza: «...mettendoci in condizione di dover dipendere da archivi internazionali (come ad esempio l'Internet Archive) anche per la documentazione della cultura italiana, che ha importanti espressioni nella rete, in particolare per quanto riguarda la ricerca».

<sup>238</sup> De Robbio(2006 a). A. De Robbio è autrice di molti contributi su questi temi, tra cui si segnalano: Id. *Auto-archiviazione per la ricerca: problemi aperti e sviluppi futuri*. “Bibliotime”, VI, 3, novembre 2003. URL:<<http://www.spbo.unibo.it/bibliotime/num-vi-3/derobbio.htm>>.



diritti proprietari, avendo alle spalle, dice De Robbio, già la sesta direttiva europea sul diritto d'autore in rete<sup>239</sup> assorbita dall'ordinamento italiano con la modifica dell'art. 171-ter della Legge 633/1941 sul Diritto d'autore, quest'eccezione per l'attività di ricerca e di studio personali,

«... nella sua poco chiara formulazione, così recita: E' libera la comunicazione o la messa a disposizione destinata a singoli individui, a scopo di ricerca o di attività privata di studio, su terminali aventi tale unica funzione situati nei locali delle biblioteche accessibili al pubblico, degli istituti di istruzione, nei musei e negli archivi, limitatamente alle opere o ad altri materiali contenuti nelle loro collezioni e non soggetti a vincoli derivanti da atti di cessione o da licenza».

Il rapporto tra obbligo di deposito legale e rispetto delle norme sul diritto d'autore e sui diritti connessi non è del tutto chiarito, né nel Regolamento, per es. all'art. 37, che si focalizza sui documenti in rete e la loro raccolta automatica, per la questione relativa alla riproduzione delle opere, né nella Legge, in merito a quanto prescritto dall'articolo 2, comma 1, lett. c), sulla consultazione e disponibilità dei documenti.

Poiché nell'articolo 37, “*Modalità di deposito e acquisizione dei documenti diffusi tramite rete informatica*”, si demanda ad un altro apposito regolamento, da redigersi dopo una fase sperimentale, condotta su basi volontarie e tramite «...*stipulazione di appositi accordi con i soggetti obbligati al deposito, con i quali verranno definite le modalità tecniche del deposito, prevedendo, ove possibile, anche forme automatiche di raccolta, secondo le migliori pratiche e conoscenze internazionali del settore*», dal cui esito dovrebbe sortire una compiuta disciplina di tale specifico aspetto delle questioni legate al diritto d'autore in ambiente digitale, si spera anche nella riformulazione del testo dell'articolo stesso, e, poiché le questioni correlate al diritto d'autore emergeranno con sempre maggior forza. E, pure,

«si spera che non siano le Università e la ricerca a far le spese di scelte inique o poco logiche dettate da timori, spesso ingiustificati che sempre sono in agguato quando si opera nella sfera digitale, avanzati da categorie che difendono interessi economici a svantaggio del bene pubblico»<sup>240</sup>.

Dopo l'approvazione della L. 106/2004, l'Associazione Italiana Biblioteche (AIB) ha promosso, attraverso la propria Commissione nazionale Biblioteche e servizi nazionali (CNBSN), un confronto di idee tra i diversi soggetti interessati alla riforma,<sup>241</sup> per individuare gli elementi critici in vista di una efficace applicazione della legge. In vista

---

<sup>239</sup> Direttiva 2001/29/CE (Sesta Direttiva) del Parlamento europeo e del Consiglio del 22 maggio 2001, *sull'armonizzazione di taluni aspetti del diritto d'autore e dei diritti connessi nella società dell'informazione*.

<sup>240</sup> De Robbio (2006 a).

<sup>241</sup> *Speciale Deposito legale*, “AIB notizie”, 17 (2004), n. 6.

dell'istituzione dell'“*archivio regionale delle pubblicazioni*” e delle conseguenze che ciò potrebbe comportare a livello organizzativo e per una più razionale applicazione della norma, ha anche avviato un'indagine sulle biblioteche già destinatarie della copia di deposito,<sup>242</sup> i cui risultati, presentati nel maggio 2006, apportano una conoscenza più analitica della realtà bibliotecaria alla quale dovrebbe essere affidata la conservazione dei documenti acquisiti per deposito legale. Più specificatamente in relazione al deposito dei documenti informatici o diffusi tramite rete informatica, l'AIB ha prodotto altri due documenti d'approfondimento:

\* *Nuova legge sul deposito legale e documenti digitali*. Documento di lavoro preparato dal Gruppo AIB Biblioteche digitali. Versione 8 del 1 giugno 2004.

\* *Proposte tecniche per il deposito legale dei "documenti diffusi tramite rete informatica"*. Testo tratto da appunti di lavoro per la preparazione del Regolamento di applicazione della L. 106/2004, aggiornato e adattato alla pubblicazione su AIB-WEB il 2005-12-08. A cura di Giovanni Bergamin.<sup>243</sup>

\* *Nuova legge sul deposito legale e documenti digitali*. (2004)

Come premessa viene riproposta la dichiarazione della *Carta sulla conservazione dell'eredità culturale digitale* (qui già cit. nel cap. 1.2.), che afferma che i documenti digitali hanno la stessa "dignità" dei documenti su carta e l'esigenza di conservarli al fine di mantenerli accessibili nel tempo, individuando, fra gli strumenti chiave necessari al conseguimento di questo fine, il deposito legale o volontario dei documenti digitali in istituzioni pubbliche, come le biblioteche.

Partendo dalla legge 106/2004 sul deposito legale, che rende anche i documenti digitali oggetto di deposito, in quindici punti sintetici vengono affrontate molte problematiche relative al tema della preservazione del digitale e vi si trovano molte considerazioni.

Anzitutto il documento presenta il deposito legale come un servizio che garantisce nel lungo periodo l'accesso alla produzione editoriale, un servizio che, peraltro, le biblioteche nazionali esercitano da molto tempo e che fa parte integrante della loro missione. Le biblioteche nazionali per affrontare l'archiviazione del Web, partono dalle esperienze in corso (tra cui spicca l'esperienza di Internet Archive) e molte (tra cui la Biblioteca Nazionale

---

<sup>242</sup> Vittorio Ponzani. *Deposito legale: il regolamento, all'improvviso*. "AIB notizie", 19 (2006), n. 3-4, p. 3.

<sup>243</sup> Entrambi i documenti sono disponibili on-line. Il primo contributo è stato pubblicato in "AIB notizie", 6, 2004. Copyright AIB 2004-06 ; ultimo aggiorn. 2004-06-15 a cura di Giovanni Bergamin. URL: <<http://www.aib.it/aib/commiss/bdigit/deplegdig.htm>>.

Centrale di Firenze)<sup>244</sup> partecipano al Consorzio internazionale tra biblioteche internazionali e nazionali e Internet Archive, l'International Internet Preservation Consortium (IIPC).

Per la complessità di molte attività e per riuscire a tenere il passo con gli sviluppi delle risorse digitali e delle tecniche di harvesting, si considera ragionevole pensare anche ad un “*Consorzio nazionale tra le biblioteche depositarie*” che definisca priorità, livelli di servizio e di cooperazione: nei confronti dell'utente finale il consorzio si porrebbe come Comunità distributiva di contenuti digitali e dei relativi servizi di accesso. È presumibile che a tali servizi possano accostarsi anche altre biblioteche o altre istituzioni, sulla base di specifici accordi.

Tra le attività che rendono possibile la conservazione (intesa sempre come salvaguardia e possibilità di accesso nel lungo periodo) della ricchezza dei contenuti digitali in rete, c'è l'impegno a garantire l'autenticità dei documenti archiviati e la persistenza nel tempo dei riferimenti «come servizi essenziali per non perdere le radici»; inoltre, il deposito presso un “terzo” - le biblioteche nazionali che istituzionalmente hanno il mandato per il lungo periodo - offre la base di partenza per l'implementazione di un sistema dei servizi di autenticità e persistenza, ad di là di tutte le tecnologie che potranno essere usate.

Viene stigmatizzato l'accesso con la frase “per tutti e con principi di equità”: tutte le risorse digitali archiviate (libere o a pagamento) devono essere accessibili presso le istituzioni centrali depositarie, visto che funzione fondamentale del deposito legale è quella di garantire a tutti condizioni di uguaglianza nell'accesso alla produzione editoriale, funzione che va accuratamente bilanciata con i legittimi interessi del produttore. Nel caso delle risorse digitali libere da diritti si potrebbe pensare ad una fruizione in rete senza il vincolo del luogo e, tuttavia, prendere in conto attentamente implicazioni di tipo giuridico (p. es. siti Web con materiale diffamatorio ecc.). Dovrebbe essere inoltre assicurato l'accesso anche da parte degli utenti delle altre biblioteche individuate dal regolamento come depositarie.

Sono state aggiunte anche alcune considerazioni relativamente ai costi (che in base alle odierne conoscenze possono comunque essere preventivati), al traffico generato dal crawler e per la larghezza di banda richiesta, per i quali è ragionevole ipotizzare che l'harvesting sia gestito solo da una istituzione e, per assicurare livelli elevati di sicurezza e per le problematiche di “disaster recovery”, si valuta ancora ragionevole ipotizzare che vi debba essere un'altra istituzione centrale con una copia continuamente allineata.

Per quanto riguarda il deep web e “i documenti diffusi su supporto informatico”, che richiedono una trattativa diretta con il produttore, è auspicabile una collaborazione in fase di

---

<sup>244</sup> Qui anche Tammaro (2006), p. 289.

acquisizione tra tutte le biblioteche depositarie. Sulla catalogazione, si considera che non tutto il materiale depositato potrà avere lo stesso livello di catalogazione: in alcuni casi (per esempio periodici elettronici scientifici) potrà essere applicata la catalogazione secondo gli standard biblioteconomici correnti, mentre nel caso dei siti Web in generale occorrerà prevedere sistemi di “gestione controllata” del contenuto, con la parametrizzazione dell'harvesting.

*\* Proposte tecniche per il deposito legale dei “documenti diffusi tramite rete informatica”. (2005)*

Il documento è stato elaborato con lo scopo di delineare alcune proposte tecniche relativamente al deposito legale dei “documenti diffusi tramite rete informatica”. Presenta sei note informative sintetiche, relative ad alcune problematiche tecniche e giuridiche collegate alla tecnologia di raccolta automatica, l'harvesting, il lavoro di archiviazione del Web portato avanti da Internet Archive, l'uso dell'harvesting nel contesto del deposito legale, una breve panoramica internazionale sull'applicazione del deposito legale ai documenti in Rete.

### **3.3. Il deposito istituzionale per la preservazione delle risorse culturali tra Diritto d'autore e Deposito legale. Il caso dell'Università italiana dopo la nuova legge sul Deposito legale.**

Già nella Carta dell'UNESCO, sono affermate come prioritarie le esigenze di raggiungere un equilibrio tra interessi di mercato e valori culturali, di favorire lo sviluppo delle collezioni e dei depositi istituzionali, per la diffusione della produzione culturale e scientifica, che permettano accesso e diritto all'informazione e alla conoscenza, uso della documentazione, qualità di servizi, tutela e controllo dei diritti di proprietà intellettuale. I valori di accessibilità alle risorse digitali e di condivisione di esperienze e saperi, ribaditi in quasi tutti i 12 articoli, si riconoscono perfettamente nell'uso delle tecnologie open source e nell'avvio di sistemi di Open Access.<sup>245</sup>

I depositi istituzionali, o institutional repositories,<sup>246</sup> essendo inizialmente legati alle esigenze del mondo della ricerca scientifica,<sup>247</sup> sono ad oggi un fenomeno prevalentemente concentrato nel mondo delle Università, che possono così gestire e mettere a disposizione

---

<sup>245</sup> Sull'Open Access, v. già la nota 222. Si riporta ancora De Robbio (2006 b): «Con il termine Open Access si intende l'accesso aperto alle produzioni intellettuali dei ricercatori e degli studiosi di tutto il mondo. L'Open Access è un movimento che incoraggia scienziati, ricercatori e studiosi a disseminare i propri lavori di ricerca rendendoli liberamente accessibili alle altre comunità di ricerca. Lo scopo dell'Open Access è rimuovere ogni barriera economica, legale o tecnica all'accesso dell'informazione scientifica, ciò al fine di garantire il progresso scientifico e tecnologico a beneficio di tutta la collettività. Anche nel nostro paese, da qualche tempo, le Università e gli enti di ricerca si stanno dotando di archivi aperti allineandosi agli altri paesi del mondo dove l'accesso aperto è già una realtà più manifesta [...] e si sta affermando con sempre maggior forza e interesse da parte di diversi soggetti, sebbene ad oggi siano ancora poche le Università che hanno messo a disposizione degli autori strumenti legati al mondo dell'Open Access, riviste o archivi ad accesso aperto. Inoltre gli archivi ad oggi esistenti non contengono ancora una quantità rappresentativa dell'intera produzione intellettuale dell'istituzione. È dunque fondamentale aumentare la consapevolezza istituzionale, ma soprattutto creare un tessuto fertile a livello nazionale, riguardo alle nuove modalità della comunicazione scientifica e alle nuove possibilità di editoria accademica, attraverso una concreta promozione a livello italiano. L'Open Access è un movimento che incoraggia scienziati, ricercatori e studiosi a disseminare i propri lavori di ricerca rendendoli liberamente accessibili alle altre comunità di ricerca. Possiamo affermare che l'Open Access non è solo un movimento, un insieme di iniziative internazionali con al centro gli scienziati e i bibliotecari coalizzati assieme, ma una strategia».

<sup>246</sup> Risorse Web selezionate sul mondo Open Access e Repository si possono trovare grazie ai portali *Openarchives.eu*, Guida ai repository digitali OAI-PMH, URL:< <http://www.openarchives.eu/>>, e *Pleiadi*, Portale per la Letteratura scientifica Elettronica Italiana su Archivi aperti e Depositi Istituzionali, URL:<<http://www.openarchives.it/pleiadi/>>. PLEIADI è un progetto AEPIC (Academic Electronic Publishing Infrastructure CILEA), dei consorzi interuniversitari italiani CASPUR e CILEA, che ha l'obiettivo di realizzare una piattaforma nazionale per l'accesso centralizzato alla letteratura scientifica depositata negli archivi aperti italiani; si inserisce nel contesto della Budapest Open Access Initiative, che promuove l'accesso libero alla letteratura scientifica, e della Open Archives Initiative, che sviluppa e promuove gli standard per l'interoperabilità degli archivi.

<sup>247</sup> Cfr. anche Giuseppe Vitello. *La comunicazione scientifica e il suo mercato*. "Biblioteche oggi", 21 (2003), 5, p. 37-57, <<http://www.bibliotecheoggi.it/2003/20030503701.pdf>>. Paola Gargiulo. *Gli Atenei italiani per l'Open Access: verso l'accesso aperto alla letteratura di ricerca*. Messina 4-5 novembre 2004, "AIDA Informazioni", 4 (2004), p. 88.

della propria comunità la documentazione digitale prodotta, ma altri modelli di gestione e diffusione di produzione scientifica e culturale dovrebbero essere sperimentati anche con altre istituzioni culturali pubbliche. Tra queste, ad esempio, si evidenziano le strutture centrali tecnico-scientifiche, ma anche gli istituti periferici del Ministero per i beni culturali, quando hanno mission di documentazione sul territorio o siano già impegnate in attività che prevedano anche pubblicazioni elettroniche o produzioni digitali, come accade per molte Biblioteche statali, che possiedono fondi antichi man mano digitalizzati, che dalla fine degli anni '90 possiedono propri siti Web nutriti e complessi, che organizzano e gestiscono nelle proprie strutture Mediateche o sale multimediali.<sup>248</sup>

Anna Maria Tammaro<sup>249</sup> ne dà questa definizione:

«I depositi istituzionali devono considerarsi vere e proprie biblioteche digitali, che si basano su un intero assetto digitale (digital asset): uno spazio virtuale in cui collezioni digitali, comunità di utenti e servizi interagiscono, con la finalità di creare, condividere e usare la conoscenza, rappresentata da risorse digitali».

E ancora: «I depositi istituzionali devono infatti essere considerati vere e proprie biblioteche digitali, di cui hanno tutte le caratteristiche e tutte le opportunità».

Con le nuove tecnologie digitali il campo del commercio di oggetti digitali è diventata un'area molto attiva di ricerca e sviluppo e proprio nell'editoria scientifica, come afferma ancora A. M. Tammaro<sup>250</sup>, sono individuabili i grandi cambiamenti:

«Molti autori hanno affermato che l'editoria digitale potrà salvare la comunicazione scientifica dall'attuale crisi, anche se la prima esperienza di periodici elettronici ha rivelato l'illusorietà di questa speranza, in un mondo dominato da grandi concentrazioni editoriali».

L'opportunità di creazione di contenuti della conoscenza scientifica e i vantaggi per la velocità di fruizione e le dimensioni della disseminazione, sono però stati osteggiati nel suo sviluppo proprio da politiche del *copyright* che restringono e non ampliano le possibilità di diffusione della conoscenza e da sistemi di mercato con sono stati in grado di permettere presto uno sviluppo ottimale dell'editoria elettronica.

L'Open Access è divenuto il paradigma internazionale del principio del diritto di accesso alla letteratura scientifica, divenendo il modello più adeguato, partecipativo, trasparente, nato e sviluppatosi proprio come risposta alla diffusione di accresciuti bisogni e di difficoltà oggettive: ha superato il modello editoriale tradizionale, in crisi per diversi fattori: una crescita delle concentrazioni editoriali internazionali, l'aumento dei prezzi sproporzionato rispetto ai costi di produzione, le politiche del *copyright* sempre più restrittive per quanto

---

<sup>248</sup> Cfr. anche Mandrillo (2005).

<sup>249</sup> Tammaro (2004)

<sup>250</sup> Tammaro (2006), p. 288-291-295.

riguarda i diritti di utilizzazione consentiti, le rigidità dei criteri di selezione di riviste tradizionali, la lentezza dell'iter della pubblicazione che, in alcuni settori, può anche inficiare l'utilità del lavoro stesso.

Ancora Tammaro:

«Un cambiamento davvero rivoluzionario è rappresentato dalla nascita e dalla diffusione di contenuti scientifici di qualità completamente liberi per il destinatario finale, e che comprendono sia la produzione degli archivi aperti e dei depositi istituzionali accademici sia i servizi di informazione (subject gateway, servizi di reference) offerti da alcune istituzioni scientifiche. Alcuni pionieri dei depositi istituzionali [... con nota bibliogr.] sono arrivati ad argomentare che l'informazione scientifica potrà essere disponibile in Internet liberamente o a costi assai più limitati di ora. Gli archivi di e-print, a cominciare dal primo sviluppato al Los Alamos National Laboratory dal fisico Ginsparg, forniscono un accesso efficiente ed economico a milioni di articoli, alternativo alla distribuzione, a volte sotterranea, degli stessi articoli per mezzo di fotocopie, prestito interbibliotecario, document delivery, preprint distribuiti dagli autori ai colleghi».

Il problema dell'accesso e dell'uso delle risorse digitale s'incontra concretamente nei modelli di deposito istituzionali e anche nei nuovi modelli proposti per la comunicazione scientifica, di cui Tammaro definisce caratteristiche e problematiche:<sup>251</sup>

- i modelli editoriali non commerciali (come le Università, le associazioni scientifiche ed altre istituzioni di ricerca che riescono a realizzare sistemi di pubblicazione senza coinvolgere l'editoria commerciale, sono volti alla condivisione dei risultati della ricerca, alla diffusione tramite Internet, a mantenere costi ragionevoli, all'incremento del consumo di informazione scientifica, all'auto tutela governo dei diritti IPR dei prodotti della propria ricerca, alla preservazione a lungo termine delle collezioni elettroniche (ma non solo) prodotte e a garantirne l'accesso),
- i modelli centralizzati di collezioni di risorse digitale (sono i più orientati anche alla conservazione, e a mantenere l'accesso universale alle pubblicazioni scientifiche o per consentire un deposito a editori o fornitori; i problemi di costi e gli aspetti di copyright e di diritto d'accesso necessitano di una organizzazione, pubblica o privata, che se ne faccia carico, ma per garantirne un ruolo di equilibrio tra diritti proprietarie e accesso democratico e conservazione, comunque, sarebbe gestito al meglio dalla collaborazione con consorzi di biblioteche e istituzioni pubbliche,
- modelli di istituzioni no profit (sono i più orientati alla fruibilità immediata, universale, specializzata per target ben individuati, e hanno alle spalle un'organizzazione

---

<sup>251</sup> Tammaro (2006), p. 288-291-295.

come un'agenzia o un'associazione che si occupa dell'infrastruttura, della promozione di standard, dell'efficienza informativa degli archivi eliminando ridondanze e duplicazioni, delle negoziazioni anche sul diritto d'autore, se ottiene il supporto delle Università e può dichiarare lo scopo di "studio e ricerca"),

- o i modelli di pubblicazione a domanda (commerciali, diretti alla fornitura di pubblicazioni, non prevedono intermediazioni, sono le trasformazioni del sistema attuale dell'editoria scientifica, non particolarmente dedicato all'archiviazione e conservazione).

Due ostacoli rendono non pienamente accessibili i contenuti intellettuali: i prezzi alti e i diritti di proprietà intellettuale che non consentono la riproduzione o il riutilizzo dei contenuti. Considerato anche che gli Atenei con la contrazione dei budget delle Università, dei centri di ricerca e delle loro biblioteche, non hanno più le forze per pagare tre volte l'informazione che producono (i.e.: produzione della ricerca, contributo per la pubblicazione, acquisto della pubblicazione).

Polemichamente Rosa Maiello<sup>252</sup> fa notare i paradossi economici che rendono indiscutibilmente vantaggiosa l'offerta del Open Access: per esempio, il fatto che paesi ai primi posti per produzione scientifica (come l'Italia), è all'ultimo posto per investimenti economici su ricerca e innovazione, ma gli investimenti sono prevalentemente pubblici, cioè vi contribuiscono i cittadini, che poi si trovano a dover riacquistare, e a prezzi anche elevati, ciò che hanno già pagato anticipatamente; in più, c'è la considerazione che i centri di ricerca pubblici hanno dimostrato una maggiore produttività, rafforzando

«...l'opinione di quanti sostengono che lo sviluppo delle conoscenze dipende dalla libertà di coltivarle e farle circolare, e deve essere sottratto a logiche puramente di mercato»

ma ciò non basta per riservare, nei contratti con gli editori, la possibilità di avere una copia delle opere in archivi liberamente consultabili.

Osserva Antonella De Robbio:<sup>253</sup>

«...I diritti di proprietà intellettuale "degli autori", finiscono gratuitamente nelle mani degli editori... E, sicuramente non appena gli autori avranno acquisito consapevolezza che il modello tradizionale non può reggere... allora la filosofia del movimento si espanderà. Questo richiederà tempo, ma sono in molti ormai a credere che i risultati delle ricerche, le scoperte scientifiche, le produzioni intellettuali della ricerca spettano a chi li produce e devono poter essere disseminate a beneficio dell'umanità e che tali risultati non possono più

---

<sup>252</sup> Rosa Maiello. *Open Access* o del futuro della comunicazione scientifica. Associazione italiana biblioteche. "Bollettino AIB", 4, 2006. P. 313-316. Da cui è tratto anche l'inciso seguente.

<sup>253</sup> De Robbio (2003), da cui sono presi anche gli incisi seguenti.



continuare ad essere oggetto di lucro da parte di poche grosse multinazionali editoriali».

«Possiamo affermare che l'Open Access non è solo un movimento, un insieme di iniziative internazionali con al centro gli scienziati e i bibliotecari coalizzati assieme, ma una strategia».

«Le potenzialità di un deposito istituzionale possono essere numerose:

- giocano un ruolo significativo nel processo evolutivo della ristrutturazione della comunicazione scientifica;
- offrono una risposta strategica ai problemi esistenti nel sistema che regola i periodici scientifici;
- forniscono un immediato complemento di qualità al modello di pubblicazione scientifica esistente;
- stimolano l'innovazione entro una struttura di editoria disaggregata;
- sono indicatori tangibili della qualità di un'istituzione, incrementandone la visibilità il prestigio e il valore a livello pubblico».

Delle pubblicazioni Open Access, De Robbio e Marchitelli,<sup>254</sup> rimarcano queste caratteristiche:

- sono garantiti a tutti gli utilizzatori il diritto d'accesso gratuito e l'autorizzazione a riprodurle, utilizzarle, distribuirle, trasmetterle, purché ne siano riconosciute la paternità intellettuale;
- il contributo è depositato in un archivio aperto che impiega standard tecnici adeguati e in un formato elettronico che rispetta uno standard internazionalmente riconosciuto.

L'auto-archiviazione o *self-archiving* è un processo essenziale per questo nuovo modello di comunicazione scientifica: è il primo passo all'interno di un circuito di disseminazione informativa caratterizzato dal deposito delle proprie produzioni di ricerca entro spazi "aperti", che diventa funzionale alla ricerca stessa, in quanto non costituisce un'alternativa alla pubblicazione dei lavori in riviste scientifiche, ma piuttosto un'azione parallela, poiché le produzioni di ricerca auto-archivate «...massimizzano e accelerano l'impatto della ricerca, massimizzando l'accesso».

---

<sup>254</sup> De Robbio – Marchitelli (2005)

### 3.3.1. I depositi istituzionali dell'Università.

Nel suo articolo, già ampiamente citato,<sup>255</sup> sui depositi istituzionali intesi come “spazio virtuale al servizio della comunicazione scientifica”, Tammaro indica come finalità di base per lo sviluppo dei depositi istituzionali proprio la condivisione di informazione e nei servizi per gli utenti finali, più che l'archiviazione di una collezione:

«Lo scopo delle istituzioni che realizzano un deposito è quello di utilizzare al meglio le possibilità di condivisione e collaborazione offerte dal web, fornendo un'infrastruttura informativa e comunicativa che serve agli utenti per svolgere la propria attività di studio, di ricerca, di didattica».

In un documento IST<sup>256</sup> della Commissione Europea del 2003 leggiamo che, tra i compiti affidati alle Università nella Società della conoscenza, vi è quello di partecipare alla vita della comunità territoriale ed essere luoghi di consulenza per i cittadini e di confronto tra scienziati e popolazione sullo stato delle conoscenze, pertanto l'avvio di forme di deposito con Open Access da parte delle istituzioni accademiche si può considerare come “un'estensione naturale” della loro responsabilità di generatori di ricerca (*IST Programme User-friendly Information Society Tecnology Programme*, 2003).

Tappe a favore dell'Open Access sono state le Dichiarazioni di Budapest (2002), di Berlino (2003) e di Messina (2004) che hanno costituito un'assunzione d'impegno per i firmatari, oltre a contribuire a diffondere l'informazione:

(2002) *BOAI, Budapest Open Access Initiative*. Iniziativa dell'Open Society Institute, scaturita da un meeting tenutosi nel dicembre 2001 a Budapest per sostenere il movimento per l'accesso aperto: BOAI-I self-archiving e BOAI-II Open Access journals.  
URL: <<http://www.soros.org/openaccess/index.shtml>>.

(2003) *Berlin Declaration ; Dichiarazione di Berlino del 22 Ottobre 2003*.  
URL:<<http://www.zim.mpg.de/openaccessberlin/berlindeclaration.html>> ; trad. it. URL:  
<[http://www.zim.mpg.de/openaccess-berlin/BerlinDeclaration\\_it.pdf](http://www.zim.mpg.de/openaccess-berlin/BerlinDeclaration_it.pdf)>.

---

<sup>255</sup> Tammaro (2004)

<sup>256</sup> Con questa sigla è denominato anche il programma europeo che, per eccellenza, contribuisce alla messa a punto di nuove soluzioni tecnologiche per la cultura; una rassegna esaustiva delle attività culturali in IST si trova all'URL: <<http://cordis.europa.eu.int/ist/digicult/projects.htm>>. Nei suoi sviluppi più recenti è stato articolato in due rami di interesse: Accesso e fruizione-Conservazione di medio e lungo termine degli archivi e delle collezioni digitali. Copre i settori tecnologici d'interesse per il patrimonio culturale così articolati: Digitisation, Digital Libraries, Digital Preservation, Heritage for All and Community Memory, Intelligent Heritage, Preservation of audio-visual objects and films, Supporting activities (attività di consulenza).

Oltre 60 firme di Atenei ed altre istituzioni italiane, alla pagina

<<http://www.zim.mpg.de/openaccess-berlin/signatories.html>>.

*Berlin Declaration on Open Access on Knowledge - Dichiarazione di sostegno all'accesso aperto da parte di Atenei ed enti di ricerca.*

Elenco dei firmatari italiani. URL:

<<http://www.zim.mpg.de/openaccessberlin/signatories.html>>.

(2004) *Dichiarazione di Messina, del 4 novembre 2004,*

<<http://www.aepic.it/conf/viewpaper.php?id=49&cf=1>>.

Alle azioni dichiarative sono seguite anche una crescente diffusione di buone pratiche e l'adozione di regole, come, per esempio, le *Linee guida per lo sviluppo dell'editoria elettronica negli Atenei italiani*,<sup>257</sup> pubblicate nel 2005 dalla Commissione biblioteche della CRUI (Conferenza dei Rettori delle Università Italiane): tra queste regole sono previste la riserva del diritto di deposito in un archivio istituzionale nei contratti di edizione relativi a opere pubblicate su fondi di ricerca, e forme di incentivi e disincentivi in termini di finanziamenti agli autori, secondo che provvedano o meno al deposito.

Uno speciale capitolo è quello che riguarda il deposito legale dei documenti concernenti la produzione scientifica della ricerca italiana, che Antonella De Robbio tratta (nel suo articolo sul Regolamento sul deposito legale, già cit.) in una valutazione critica sulla legge italiana e sul relativo regolamento d'attuazione,<sup>258</sup> nella quale, tra le altre cose, nota che:

«In effetti per quanto attiene la produzione intellettuale accademico-scientifica, la legge sul deposito legale poco dice, se non nell'art. 37, dove accenna all'eventuale stipulazione di accordi che assicurino a livello prioritario la raccolta dei documenti concernenti la produzione scientifica delle Università, dei centri di ricerca e delle istituzioni culturali, laddove poi siano anche previsti sistemi idonei ad assicurare la certezza della data del deposito e l'autenticità del documento depositato, anche al fine di dare certezza sulla data di produzione o di diffusione del documento, nonché sulla provenienza dal suo autore»

---

<sup>257</sup> Sono state pubblicate nel 2006 le *Raccomandazioni per lo sviluppo dell'editoria elettronica negli atenei italiani*. a cura di Patrizia Cotoneschi, Giancarlo Pepeu ; con i contributi di Marco Marandola...[et. al.]. - Firenze : Firenze university press, 2006. Disponibile anche la pubblicazione elettronica URL: <[http://www.cruil.it/data/allegati/links/3290/pubblicazione\\_raccomandazioni\\_editoria.pdf](http://www.cruil.it/data/allegati/links/3290/pubblicazione_raccomandazioni_editoria.pdf)>.

<sup>258</sup> De Robbio (2006) (a). Cit. nota 11.

Altra considerazione è anche che, in base all'art. 37, entro eventuali sperimentazioni opportunamente concordate tramite accordi, saranno determinate le modalità di raccolta automatica o selettiva di documenti digitali prodotti dagli Atenei.

Sempre De Robbio individua tre ambiti di produzione scientifica su cui soffermarsi:

- gli archivi aperti di Ateneo (e-prints): attualmente in Italia sono già stati implementanti una ventina di archivi istituzionali che raccolgono pre-print, articoli, relazioni a convegni e conferenze, tesi di dottorato e altro materiale;<sup>259</sup>
- le tesi di dottorato;
- le tesi.

In merito agli archivi aperti d'Ateneo, sarà necessaria una raccolta (harvesting)<sup>260</sup> di tipo automatico che assicuri, anche a fini conservativi, un deposito finalizzato agli scopi della legge, diventa importante stabilire accordi tra l'ente e la biblioteca nazionale. Per esempio, il CIB di Bologna ha stipulato, già nel marzo 2004, nell'attesa del regolamento di attuazione della nuova legge sul deposito legale, un accordo con la Biblioteca Nazionale Centrale di Firenze (BNFC), per effettuare il deposito volontario delle pubblicazioni scientifiche e didattiche elettroniche multimediali edite dalla biblioteca digitale dell'Università.

Nell'ambito delle tesi di dottorato e tesi in generale: la situazione è più problematica. Entrambe sono oggetto di grande dibattito in termini di diritto d'autore: secondo la giurisprudenza non si può mai prescindere dalla legge sul diritto d'autore, legge sull'esclusiva, e che quindi una tesi, quale opera inedita, non possa essere considerata "documento ad uso pubblico" ma che sia solo il suo autore a decidere se e come pubblicarla o meglio se renderla pubblica, eventualmente a seguito di liberatoria che stabilisca anche su quali diritti si autorizzano determinate azioni: consultazione, prestito, riproduzione, distribuzione, deposito, messa in rete e quant'altro.

---

<sup>259</sup> Anche Bevilacqua (2005).

<sup>260</sup> AIB (2004), § 3, specifica che dal punto di vista del diritto d'autore l'harvesting si configura come una copia: le pagine vengono infatti trasferite - di solito mediante il protocollo "http" - dal server Web alle memorie di massa gestite dal crawler. Nel caso di Google ad esempio la pagina raccolta - oltre ad essere indicizzata - viene anche memorizzata come copia cache (l'archiviazione temporanea tra due sessioni di raccolta). La copia è esplicitamente vietata dalla normativa sul diritto d'autore: è evidente che, nel caso del deposito legale delle risorse elettroniche remote, con l'harvesting - in quanto tecnica di copia - viene data una deroga al "diritto alla copia". Anche T. Giordano, in Giordano (2006), p. 92, afferma: «La complessità che la conservazione digitale presenta dal lato giuridico è grande perché la legislazione sul copyright e le licenze influiscono sulla capacità di provvedere all'accesso a lungo termine. Dall'analisi effettuata su tutti i metodi di conservazione (refreshing, migration, emulation, recreation) e le relative procedure richiedono potenzialmente qualche sorta di copiatura e per tanto implicano autorizzazioni non facili da ottenere da parte dei molteplici detentori dei diritti. Anche le eccezioni in favore delle biblioteche previste in alcuni paesi, sul piano pratico hanno un valore molto limitato, in quanto si riferiscono a pubblicazioni fuori commercio».

Nella normativa che si è susseguita non si è mai sciolto completamente il conflitto tra diritti d'autore e diritto alla consultabilità da parte di terzi, rivelando anche molte contraddizioni:

1980: la norma che istituiva il dottorato di ricerca in Italia (D. P. R. 382/1980 nell'art. 73)<sup>261</sup> stabiliva che il rilascio del titolo di dottore di ricerca è subordinato al deposito di copie, anche non stampate (riferendosi a copie manoscritte, ma la frase così formulata lasciava comunque spazio anche a nuove soluzioni tecnologiche);

1999: il regolamento (D. M. 224)<sup>262</sup> in materia di dottorato di ricerca abroga di fatto l'articolo precedente del D. P. R. 382/80, modificando il disposto e passando all'Università medesima l'onere del deposito di copia della tesi finale presso le biblioteche nazionali di Roma e Firenze, successivamente al rilascio del titolo:

- il rilascio del titolo di dottore di ricerca non è più subordinato al deposito
- l'obbligo del deposito passa all'Università medesima
- non si parla più di rendere pubblica la tesi per almeno trent'anni.

Non sono più citati quei requisiti di “pubblica consultabilità” (che accomunavano la tesi di dottorato alla “pubblicazione”, differenziandola dalla semplice tesi) e si aggiunge il dubbio di legittimità sul fatto che un ente non potrebbe disporre (a seguito di un decreto ministeriale) un deposito di un'opera d'autore, senza il suo consenso, con l'aggravante che il deposito avviene in una biblioteca “pubblica” che per sua missione è tenuta sì alla conservazione, ma anche a permettere l'accessibilità dei documenti che possiede;

2004/2006: la nuova legge sul deposito legale<sup>263</sup> neanche cita le tesi di dottorato tra i documenti soggetti a deposito, come se fosse considerata un'opera come le altre: si confida nella fase di definizione del regolamento, perché disciplini in modo adeguato il deposito in versione digitale con modalità di raccolta automatica dagli archivi di Ateneo o da archivi nazionali opportunamente predisposti, evitando anche la dispendiosa incombenza dell'invio delle copie cartacee, promuovendo una forma di deposito certificato sarebbe a vantaggio della ricerca stessa, in quanto rendere accessibili tali documenti gioverebbe non solo alla conservazione della memoria per le generazioni future, ma al progresso scientifico e tecnologico, visto che costituiscono «la punta di diamante della ricerca italiana e sono l'esito finale di lunghi e complessi lavori di ricerca, svolti anche in equipe, di investimenti anche

---

<sup>261</sup> D. P. Rep. 11 luglio 1980, n. 382 (S. O. G. U. 31 luglio 1980, n. 209), *Riordinamento della docenza universitaria, relativa fascia di formazione nonché sperimentazione organizzativa e didattica*.

<sup>262</sup> D.M. n. 224 del 30/04/99, *Regolamento in materia di dottorato di ricerca*, pubblicato nella G. U. 13/07/99, n. 162.

<sup>263</sup> L. 106/2004, D.P.R. 252/2006.

cospicui, e che in numerosi casi potrebbero anche confluire in applicazioni e sviluppo entro un mercato industriale, a vantaggio delle Università e della collettività intera».

Indicativi dell'utilità riconosciuta al deposito 'accademico' sono queste riflessioni di Peter Suber, che analizza il valore dell'accesso aperto per le tesi e le dissertazioni elettroniche (ETDs)<sup>264</sup> e fornisce ragioni per le quali le Università dovrebbero fornire accesso libero a tali risorse:

1. attualmente la maggior parte delle tesi sono digitali native
2. le ETDs sono la prima fase, lavori di ricerca senza royalties
3. le ETDs non sono formalmente pubblicate
4. le esortazioni non risolvono il problema dell'invisibilità di tesi e dissertazioni
5. le Università sono in una buona posizione per promuovere l'Open Access
6. la promozione dell'Open Access per le tesi educa le future generazioni di ricercatori alla cultura Open Access
7. l'Open Access mostrerebbe fuori al meglio tali risorse
8. un impegno Open Access mostrerebbe che le Università considerano le tesi seriamente.

---

<sup>264</sup> Nel numero di luglio 2006 della SPARC Open Access Newsletter (SOAN), URL: <http://www.openarchives.it/pleiadi/modules/news/article.php?storyid = 122>.

### 3.3.2. Il Deposito delle tesi di dottorato in Archivi aperti delle Università e presso le Biblioteche Nazionali.

Le Università italiane, nell'ambito delle attività della CRUI, sono impegnate nella discussione sulle modalità d'applicazione del deposito legale relativo alle tesi di dottorato, terreno non così semplice, poiché si tratta di opere di un autore, che in quanto tale detiene, oltre alla responsabilità del testo, diritti patrimoniali e di inedito, sulla riproduzione; ma sono anche opere con un alto valore culturale per la collettività, perchè di importanza fondamentale per la ricerca e testimoniano della produzione scientifica condotta all'interno dell'Università, la quale, a sua volta, ha per legge l'obbligo di curarne il deposito presso le Biblioteche nazionali e la conservazione in propri archivi. In più, non va dimenticato che le Università sono soggetti pubblici che godono di una propria autonomia decisionale e che sul territorio nazionale si distinguono in realtà molto variegata sotto molti aspetti.

Dato ciò pare evidente la necessità di giungere, in modo autorevolmente coerente ed omogeneo, a soluzioni semplici e qualitativamente ottimali, che rappresentino un punto di incontro equilibrato, cioè non lesivo degli interessi legittimi, ma mai a svantaggio di un riconosciuto bene pubblico, districandosi nella complessità delle esigenze e nella difficoltà di applicazione di norme, come già notato, in alcuni punti non del tutto precise e complete, ma tenendo fede ad impegni anche internazionali di alto respiro. Anche perché la volontà e il diritto del singolo non può superare né gli obblighi cui il soggetto Università è tenuto per legge, né le sue disposizioni organizzative.

L'Università è obbligata dalla legge italiana sul deposito legale a curare il deposito presso le Biblioteche Nazionali di Roma e Firenze delle tesi di Dottorato di Ricerca, il che comporterà per esse sia il divenire oggetto di "servizi bibliografici nazionali di informazione e di accesso", sia la loro consultabilità per i terzi, considerato questo un diritto (coi limiti circoscritti, temporanei, motivati e nei casi previsti, secondo la disciplina sulla tutela della proprietà intellettuale).

E ha anche l'obbligo della conservazione nei propri archivi, per cui, nel quadro della propria normale attività, ha scelto di aderire alla Dichiarazione di Berlino in materia di Open Access, per quanto si riferisce alla creazione di Archivi istituzionali aperti nei quali inserire, fra l'altro, le tesi di dottorato. Quindi ogni Ateneo, nella propria autonomia, deve provvedere a dare chiaro avviso, nelle modalità di accesso ai corsi di dottorato, nei regolamenti e nei bandi dei dottorati, dell'inserimento delle tesi di dottorato nei propri archivi aperti, che equivale ad un e-print, e del deposito legale presso le Biblioteche nazionali, atto che

comporterà la consultabilità della tesi, intendendo questa come il riconoscimento di un diritto dei terzi e della collettività alla conoscenza e all'informazione (pur prevedendo l'eccezione, per ragioni motivate legate all'esclusiva brevettuale, avallate dal tutor, dell'esclusione dall'accessibilità - c.d. embargo - per un determinato periodo di tempo).

Lo studente che decide di seguire un dottorato di ricerca, firmando una dichiarazione in merito al momento dell'iscrizione, s'impegna ad accettare i fatti cui andrà incontro il suo elaborato: che l'Università è obbligata a provvedere al deposito di legge delle tesi di dottorato al fine di assicurarne la conservazione e la consultabilità da parte di terzi; che, in qualità di autore, continuerà ad essere l'unico responsabile rispetto a contenuto e organizzazione della tesi, che è depositata con procedure di garanzia, soggetta a copyright, ma sono fatti salvi i diritti dell'Università di riproduzione per scopi di ricerca e didattici, con citazione della fonte; che avverrà la pubblicazione in rete nel deposito istituzionale elettronico, se questo è allestito; che l'Università ha aderito alla "Dichiarazione di Berlino per l'accesso aperto alla letteratura scientifica" e all'Open Access e che archiverà e renderà consultabile in rete il testo della tesi attraverso un Archivio istituzionale ad accesso aperto e attraverso il Catalogo ad accesso aperto della Biblioteche Nazionali Centrali di Firenze e Roma.

Poiché la norma del deposito è vigente, ovviamente, va disciplinato anche il periodo di transizione, per giungere all'applicazione della norma a pieno regime, e si prefigura una soluzione di 'declaratoria' per le attività in corso, da far firmare, come condizione, ai dottorandi al momento della richiesta di ammissione all'esame finale del Corso di dottorato.

Insomma, sia sul Deposito legale finalizzato a consultabilità e conservazione a lungo termine, sia sull'Open Access, gli aspiranti dottorandi, devono essere messi al corrente del trattamento che subirà il loro elaborato e, al loro volta, devono dichiarare di essere a conoscenza di questo trattamento e firmare la loro adesione, che non è un consenso, ma esattamente un adeguarsi, o, sarebbe meglio poter dire, un aderire essi stessi ad una "strategia" culturale, ad una concezione sulla preservazione e sulla disseminazione della conoscenza scientifica.



### 3.4. Rassegna di norme sul Diritto d'autore.

L. 22-4-1941 n. 633, *Protezione del diritto d'autore e di altri diritti connessi al suo esercizio*. (Pubblicata nella Gazz. Uff. 16 luglio 1941, n. 166.)

L. 20-6-1978 n. 399, *Ratifica ed esecuzione della convenzione di Berna per la protezione delle opere letterarie ed artistiche, firmata il 9 settembre 1886, completata a Parigi il 4 maggio 1896, riveduta a Berlino il 13 novembre 1908, completata a Berna il 20 marzo 1914, riveduta a Roma il 2 giugno 1928, a Bruxelles il 26 giugno 1948, a Stoccolma il 14 luglio 1967 e a Parigi il 24 luglio 1971*, con allegato. (Pubblicata nel Suppl. ord. alla Gazz. Uff. 2 agosto 1978, n. 214.)

L. 18-8-2000 n. 248, *Nuove norme di tutela del diritto di autore*. (Pubblicata nella Gazz. Uff. 4 settembre 2000, n. 206.)

D. Lgs. 15-11-2000 n. 373, *Attuazione della direttiva 98/84/CE sulla tutela dei servizi ad accesso condizionato e dei servizi di accesso condizionato*. (Pubblicato nella Gazz. Uff. 15 dicembre 2000, n. 292.)

Dir. 22-5-2001 n. 2001/29/CE, *Direttiva del Parlamento europeo e del Consiglio sull'armonizzazione di taluni aspetti del diritto d'autore e dei diritti connessi nella società dell'informazione*. (Pubblicata nella G.U.C.E. 22 giugno 2001, n. L 167. Entrata in vigore il 22 giugno 2001.)

L. 7-2-2003 n. 22, *Modifica al D. Lgs. 15 novembre 2000, n. 373, in tema di tutela del diritto d'autore*. (Pubblicata nella Gazz. Uff. 15 febbraio 2003, n. 38.)

D. Lgs. 9-4-2003 n. 68, *Attuazione della direttiva 2001/29/CE sull'armonizzazione di taluni aspetti del diritto d'autore e dei diritti connessi nella società dell'informazione*. (Pubblicato nella Gazz. Uff. 14 aprile 2003, n. 87, S.O.)

D. M. 14-11-2005 n. 261, *Regolamento recante: Istituzione e funzionamento del fondo previsto dall'articolo 174-quater, comma 1, lettera a, della L. 22 aprile 1941, n. 633, come introdotto dall'articolo 8 della L. 18 agosto 2000, n. 248, e modificato dall'articolo 29 del D. Lgs. 9 aprile 2003, n. 68, destinato al potenziamento delle strutture e degli strumenti impiegati nella prevenzione e nell'accertamento dei reati previsti dalla legge sulla protezione del diritto d'autore e di altri diritti connessi al suo esercizio*. (Pubblicato nella Gazz. Uff. 27 dicembre 2005, n. 300. )

D. Lgs. 13-2-2006 n. 118, *Attuazione della direttiva 2001/84/CE, relativa al diritto dell'autore di un'opera d'arte sulle successive vendite dell'originale*. (Pubblicato nella Gazz. Uff. 25 marzo 2006, n. 71).

Codice Civile: art. 2575 e segg., art. 2598 e segg.

Codice Penale: art. 473 e segg., artt. 517, 518, 528, 529, 725.

## Bibliografia

Tutti i link sono stati controllati in data 13/14-03-2007.

Albertsen (2003) = Ketil Albertsen. *The Paradigma web harvesting environment*. – P. 49-62. In : Julien Masanès - Andreas Rauber - Gregory Cobena (eds). *Proceedings : in conjunction with the 7th European Conference on Research and Advanced Technologies for Digital Libraries, ECDL 2003*. – URL: <<http://bibnum.bnf.fr/ecdl/2003/index.html>>.

Albertsen-Van Nuys (2004) = Ketil Albertsen - Carol Van Nuys. *PARADIGMA : FRBR and digital documents*. In: Patrick Le Boeuf. *Functional requirements for bibliographic records (FRBR) : hype, or cure-all?* “Cataloging & Classification Quarterly”. New York : Haworth Press Inc. 39, n. 3-4, 2004. Pp. 125-149.

*Il documento immateriale: ricerca storica e nuovi linguaggi*. A cura di Guido Abbattista e Andrea Zorzi. In particolare: Riccardo Ridi. *Biblioteche in linea*. “L'indice dei libri del mese”. XVII (2000), 5, suppl.. P. IX. Edizione on-line a cura di Michele Ansani. URL: <<http://lastoria.unipv.it/dossier/ridi.htm>>.

Barocci (2005) = Matteo Barocci. *Prospettive e problemi della catalogazione delle risorse elettroniche. Un dibattito internazionale denso di sviluppi*. “Biblioteche oggi”. 23 (2005), n. 10. P. 7-18.

Bevilacqua (2005) = Fabrizia Bevilacqua. *Aspetti della preservazione digitale. Il caso degli archivi istituzionali*. “Biblioteche oggi”. 23 (2005), 3. P. 9-13.

Bucchioni - Spinelli (2002) = Cinzia Bucchioni - Serafina Spinelli, *A proposito di FRBR*. “Bollettino AIB : rivista italiana di biblioteconomia e scienze dell'informazione”. 42 (2002), 2. P. 201-207.

Buizza (2002) = Pino Buizza. *Dai Principi di Parigi a FRBR*. “Bibliotime”. V (mar. 2002), 1. URL: <<http://didattica.spbo.unibo.it/bibliotime/>>

Conservare il digitale (2006) = *Conservare il digitale. Un Confronto Internazionale. Asolo, 29-30 September 2006*. Relazioni presentate e pre-print (in aggiornamento) a partire dall'URL: <<http://www.ulssasolo.ven.it/webpublic/ddw.aspx?n = 2261&h = -2147477295&Tipod = Conv>>.

Contenuti digitali (2005) = *Rapporto sui contenuti digitali nell'era di internet*. Elaborato dal Centro Studi del Ministro per l'innovazione e le tecnologie; Commissione interministeriale (Ministro Innovazione e tecnologie, Ministero per i Beni e le attività culturali, Ministero delle Telecomunicazioni). 2005. Disponibile sul sito del M. I. T.

Corasaniti (2005) = Giuseppe Corasaniti. *Prospettive di rinnovamento della Legge sul diritto d'autore*. "DigItalia". 2 (2006). P. 52-57.

De Pinedo - De Pinedo (2002) = Isa De Pinedo - Alberto De Pinedo, *Un approccio all'applicazione del modello FRBR alle regole di catalogazione italiane: problemi e possibili soluzioni*. "Bollettino AIB : rivista italiana di biblioteconomia e scienze dell'informazione". 42 (2002), 3. P. 267-280.

Cultural Heritage, Firenze (2006) = Conferenza internazionale *Cultural Heritage on-line. The challenge of accessibility and preservation*", 14-16 dicembre 2006. A cura di: Ministero per i Beni e le Attività culturali, Fondazione Rinascimento Digitale, Associazione CIVITA.  
Preprint on-line, URL: <<http://www.rinascimento-digitale.it/indexEN.php?SEZ = 396>>

De Robbio (2003) = Antonella De Robbio. *Auto-archiviazione per la ricerca: problemi aperti e sviluppi futuri*. "Bibliotime". VI, 3 (novembre 2003). URL: <<http://www.spbo.unibo.it/bibliotime/num-vi-3/derobbio.htm>>.

De Robbio – Marchitelli (2005) = Antonella De Robbio - Andrea Marchitelli. *E-prints for Library and Information Science (E-LIS)*. "Biblioteche oggi". 23 (2005), 10. P. 40-51.

De Robbio(2006 a) = Antonella De Robbio. *Il regolamento di attuazione della legge 15 aprile 2004, n. 106 sul deposito legale dei documenti di interesse culturale destinati all'uso pubblico, per l'università e la ricerca*. *Archivi e Biblioteche d'Italia*. "E-LIS : E-prints in

Library and Information Science (LIS) : The open archive for Library and Information Science”. URL: <<http://eprints.rclis.org/archive/00006656/>>.

De Robbio (2006 b) = Antonella De Robbio. *L'Open Access in Italia*. “DigItalia”. 1 (2006). P. 31-44.

Di Giammarco (2005) = Fabio Di Giammarco. *Conservare il futuro*. “Biblioteche oggi”. 23 (2005), 2. P. 31-33

Di Giammarco (2007) = Fabio Di Giammarco. *Web archiving e ruolo della BNCf. Un'iniziativa di carattere esplorativo*. “Biblioteche oggi”. 25 (2007), 2. P. 80-81.

Duranti (2006 a) = Luciana Duranti. *Un quadro teorico per le politiche, le strategie e gli standards di conservazione digitale: la prospettiva concettuale di InterPARES*. “Bibliotime”. IX, 1 (2006). URL: <<http://didattica.spbo.unibo.it/bibliotime/>>

Duranti (2006 b) = Luciana Duranti. *La questione fondamentale: in quali entità digitali si traduce la memoria del futuro?* “Archivi&Computer”. 16 (2006), 1. P. 24-31.

Faiella (2000) = Feliciano Faiella. *Digital preservation. Le problematiche della conservazione nel contesto dell'informazione digitale*. “Biblioteche oggi”. 18 (2000), 9. P. 64-69.

Futuro memoria, Torino (2004) = *Il futuro della memoria: la trasmissione del patrimonio culturale nell'era digitale. Torino, 10-11 novembre 2004*. CSI Torino. Tutti i materiali del convegno sono disponibili all'URL: <<http://www.memoriadigitale.csi.it/>>.

Futuro memorie, Firenze (2003) = *Futuro delle memorie digitali e patrimonio culturale : atti del Convegno internazionale, Firenze, 16-17 ottobre 2003 : The future of digital memory and cultural heritage*. A cura di Vittoria Tola e Cecilia Castellani. Roma: ICCU, 2004.

FRBR (2000) = *Seminar FRBR : Functional requirements for bibliographic records = Seminario FRBR : Requisiti funzionali per record bibliografici : Firenze, 27-28 gennaio 2000*. Atti a cura di Mauro Guerrini. Roma : Associazione italiana biblioteche, 2000.

FRBR, ICCU (2000) = *FRBR, Requisiti funzionali per record bibliografici : rapporto conclusivo*. IFLA Study Group on the functional requirements for bibliographic records ; approvato dallo Standing Committee della IFLA section on cataloguing. Edizione italiana a cura dell'Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche. Roma: ICCU, 2000.

Giordano (2006) = Tommaso Giordano. *Le collezioni non abitano più qui?* “Biblioteche oggi”. 24 (2006), 3. P. 90-102.

Giordano (2007) = Tommaso Giordano. *Gestione e conservazione delle pubblicazioni elettroniche. Percezione, pratiche, modelli*. “Biblioteche oggi”. 25 (2007), 2. P. 17-27.

Ghilli – Guerrini – Novelli (2003) = Carlo Ghilli – Mauro Guerrini – Antonella Novelli. *FRBR: analisi del record e nuovi codici di catalogazione*. “Bollettino AIB: rivista italiana di biblioteconomia e scienze dell'informazione”. 43 (2003), 2. P. 145-160. Disponibile anche on-line, da: <<http://www.aib.it/aib/boll/2003/>>

Ghilli – Guerrini (2001) = Carlo Ghilli – Mauro Guerrini. *Introduzione a FRBR: Functional requirements for bibliographic record: Requisiti funzionali per record bibliografici*. Milano: Bibliografica, [2001]. (Bibliografia e biblioteconomia; 60)

Gorman (2001) = Michael Gorman. *Metadati o catalogazione? Una falsa alternativa*. Traduzione di Stefano Gambari e Mauro Guerrini. “Biblioteche oggi”. 19 (2001), 5. P. 8-18, oppure <<http://www.bibliotecheoggi.it/2001/20010500801.pdf>>. Ed. originale: *Metadata or cataloging? A false choice*. “Journal of Internet cataloging”. 2 (1999), 1. P. 5-22.

Gorman (2003) = Michael Gorman. *Le risorse elettroniche: quali vale la pena conservare e qual'è il loro ruolo nelle raccolte della biblioteca?* Traduzione di Pino Buizza. In: *Risorse elettroniche (2001) = Le risorse elettroniche: definizione, selezione, catalogazione = Electronic resources. Definition, Selection, Cataloguing: atti del convegno internazionale, Roma, 26-28 novembre 2001*. A cura di Mauro Guerrini, con la collaborazione di Stefano Gambari e Lucia Sardo. Milano: Editrice Bibliografica, 2002. P. 93-97 (testo originale in inglese a p. 467-470). Disponibile anche on-line: “E-LIS : E-prints

in Library and Information Science (LIS) : The open archive for Library and Information Science”. Luglio 2003. URL:<<http://eprints.rclis.org/archive/00000252/>>.

Gorman (2006) = Michael Gorman. *Cataloguing the web: an oxymoron?* Alaska Library Association. Febr. 2006. URL: <<http://mg.csufresno.edu/papers/Metadata.pdf>>.

Gozzetti (2005) = Pietro Gozzetti. *DSpace e i suoi servizi. Un innovativo sistema di biblioteca digitale*. “Biblioteche oggi”. 23 (2005), 6. P. 27-34.

Guercio (2001) = Mariella Guercio. *Rischi e promesse dell'innovazione tecnologica: i conservatori del patrimonio documentario e la cooperazione fra archivisti e bibliotecari*. “Bollettino AIB : rivista italiana di biblioteconomia e scienze dell'informazione”. 41 (2001), 2. P.157-173.

Guercio (2002) = Mariella Guercio. *Archivistica informatica. I documenti in ambiente digitale*. Roma: Carrocci, c2002.

Guercio (2005 a) = Mariella Guercio. *Conservare le risorse digitali: lo stato dell'arte*. (Milano): Fondazione Giangiacomo Feltrinelli Papers, c2005.

Guercio (2005 b) = Mariella Guercio. *Formazione e ricerca per la conservazione digitale: la Scuola estiva di Delos (Antibes, 4-10 giugno 2005)*. “Digitalia”. 0 (2005). P. 105-107.

Guercio (2006) = Maria Guercio. *Certezza documentaria e memoria digitale: una riflessione sul futuro della funzione archivistica*. “Archivi&Computer”. 16 (2006), 1. P. 5-23.

Guerrini (2005) = Mauro Guerrini. *FRBR: analisi del record e nuovi codici di catalogazione*. A cura di Carlo Bianchini, Pino Buizza, Carlo Ghilli, Antonella Novelli, Lucia Sardo. Milano: Silvestre Bonnard, 2005.

Guerrini – Gambari (2002) = Mauro Guerrini - Stefano Gambari, *Definire e catalogare le risorse elettroniche. Una introduzione a ISBD(ER), AACR2 e metadati*. Saggio introduttivo di Paul Gabriele Weston. Milano: Editrice Bibliografica, 2002.

Khale (1997) = Brewster Khale. *Conservare Internet*. “Le scienze”. LVI (1997), 350. P. 89-90.

Internet per il restauro (2003) = *Internet per il restauro, il restauro per Internet*. In *Il Web per la conservazione, la conservazione del Web*. Convegno a Ferrara, 3 aprile 2003. “Kermes. La rivista del restauro”. 4 aprile 2003. P. 5-6. Disponibile all’URL: <[http://www.minervaeurope.org/events/documents/Ferrara2003\\_Finale.pdf](http://www.minervaeurope.org/events/documents/Ferrara2003_Finale.pdf)>.

InterPARES (2006) = *I risultati di InterPARES2*. Seminario Internazionale del 12-13 dicembre 2006 tenuto nel Convegno *ArchiExpo’ Il digitale e la memoria: dal passato al futuro*. Milano, 12-15 dicembre 2006. A cura dell’ANAI, Associazione Nazionale Archivistica Italiana. Interventi di: Philip Eppard, *Making and Maintaining Digital Materials: Guidelines for Individuals*. Kevin L. Glick, *Guidelines for Records Preservation*. Malcolm Todd, *Principles Guiding Policy for Creators and Preservers*. Abstract on-line, URL: <<http://www.anai.org/archiexpo>>.

Mandillo (2005) = Anna Maria Mandillo. *Diritto d’autore e nuovi servizi al pubblico*. “DigItalia”. 0 (2005). P. 47-61.

Marandola (2004) = Marco Marandola. *Il diritto d’autore in ambito digitale secondo la legislazione italiana: un problema aperto*. “Storiadelmondo”. 2 (2004). URL: <[www.storiadelmondo.com/22/marandola.digitalizzazione.pdf](http://www.storiadelmondo.com/22/marandola.digitalizzazione.pdf)>.

Marandola (2005) = Marco Marandola. *Il nuovo diritto d'autore. Introduzione a copyleft, open access e creative commons*. Milano: DEC, 2005.

Miele (2006) = Marzia Miele. *DELOS summer school 2006. Digital preservation in digital libraries: emerging approaches*. “AIB notizie”. 18 (2006), 10. P. 10



Minerva (2005) = National representatives Group coordination mechanisms for digitisation policies and programmes. *Coordinating digitation in Europe : progress report of the National Representatives Group coordination mechanisms for digitisation policies and programmes 2005*. Roma: Michael project, 2006.

Minerva (2006) = Ministero B. A. C., Progetto Minerva . WP5. *Manuale per la qualità dei siti Web pubblici culturali*. A cura di Fedora Filippi, coordinatore del gruppo di lavoro italiano “Identificazione dei bisogni degli utenti e dei criteri di qualità per un accesso comune”. 2. Ed. italiana aggiornata. (Roma: La Tipografia, 2006).

Le Bœuf (2005) = Patrick Le Bœuf. *Indagine su un modello concettuale al di sopra di ogni Codice*. “Bollettino AIB : rivista italiana di biblioteconomia e scienze dell'informazione”. 45 (2005), 2. P. 187-201.

Lunghi (2006) = Maurizio Lunghi. *Presentazione della rete di coordinamento europea sulla conservazione delle memorie digitali (29 giugno 2006)*. Seminario. Fondazione Rinascimento Digitale. Paper disponibile on-line, URL: <<http://www.rinascimento-digitale.it/>> 29 giugno 2006.

Miele (2006) = Marzia Miele. *DELLOS summer school 2006. Digital preservation in digital libraries: emerging approaches*. “AIB notizie”. 18 (2006), 10. P. 10.

Negroponte (2005) = Nicholas Negroponte. *Essere digitali*. Traduzione di Franco e Giuliana Filippazzi. Milano: Sperling & Kupfer, 1995. Ed. originale: *Being digital*. New York: Random house, 1995.

Knusten (2006) = Unni Knutsen. *Controllo bibliografico nei paesi nordici*. “Bollettino AIB : rivista italiana di biblioteconomia e scienze dell'informazione”. 45 (2006), 4. P. 331-341.

Ragazzini (2004) = Dario Ragazzini (et al.) *La storiografia digitale*. A cura di Dario Ragazzini. Torino: UTET libreria, 2004.

Ridi (2000) = Riccardo Ridi. *Il Web bibliotecario come incunabolo digitale*. In: Klaus Kempf - Franz Berger. *Proceedings. Riforma universitaria e rivoluzione dei media: una sfida per le biblioteche universitarie*. P.59-72. Consultata la versione in “E-LIS : E-prints in Library and Information Science (LIS) : The open archive for Library and Information Science”. URL: <<http://eprints.rclis.org/archive/00002118/>>.

Ridi (2004) = Riccardo Ridi. *La biblioteca digitale: definizioni, ingredienti e problematiche*. “Bollettino AIB : rivista italiana di biblioteconomia e scienze dell'informazione”. 44 (2004), 3. P. 273-344.

Ridi (2003) = Riccardo Ridi. *Retaggio digitale: hardware, software e politiche culturali nel mondo delle biblioteche*. Versione italiana aggiorn. al 2003, di: Riccardo Ridi. *Digital heritage: hardware, software and cultural policies in the library world*. In *The future of memory*. Edited by Giulio Blasi, Turnhout (Belgium): Brepols, 2002. P. 35-44. “E-LIS : E-prints in Library and Information Science (LIS) : The open archive for Library and Information Science”. URL: <<http://eprints.rclis.org/archive/00000157/>>.

Risorse elettroniche (2001) = *Le risorse elettroniche: definizione, selezione, catalogazione = Electronic resources. Definition, Selection, Cataloguing: atti del convegno internazionale, Roma, 26-28 novembre 2001*. A cura di Mauro Guerrini, con la collaborazione di Stefano Gambari e Lucia Sardo, Milano: Editrice Bibliografica, 2002.

Rossi (2001) = Federica Rossi. *Un nuovo acronimo per la biblioteca: FRBR*. “Bibliotime”. IV (nov. 2001), 3. URL: <<http://didattica.spbo.unibo.it/bibliotime/>>

Rothenberg (1995) = Jeff Rothenberg. *La conservazione dei documenti digitali*. “Le scienze”. LIV (1995), 319. P. 16-21.

Santangelo (2002) = Sabrina Santangelo. *Long term preservation dei documenti elettronici: il dibattito degli esperti e la risoluzione del Consiglio europeo*. “Archivi&Computer”. 12 (2002), 2. P. 85-89.

Spinazzè (2007) = Linda Spinazzè. *Comunicazione scientifica e tecnologie digitali: a che punto siamo in Italia? A proposito di Open Access e di altro*. "Biblioteche oggi". 25 (2007), 2. P. 63-72.

Spinelli (2002) = Serafina Spinelli. *Al centro di FRBR: l'utente*. "Bibliotime". V (nov. 2002), 2. URL: <<http://didattica.spbo.unibo.it/bibliotime/>>

Serrai (2002) = Alfredo Serrai. *Critica dei Functional requirements for bibliographic records (FRBR)*. "Bibliotheca". 1 (2002), 2. P. 207-215.

Tammaro (2005) = Anna Maria Tammaro. *I metadati amministrativi. Per un'efficace gestione delle risorse digitali*. "Biblioteche oggi", 23 (2005), 6. P. 63-66.

Tammaro - De Gregori (2004) = Anna Maria Tammaro – Teresa De Gregori. *Ruolo e funzionalità dei depositi istituzionali. Uno spazio virtuale al servizio della comunicazione scientifica*. "Biblioteche oggi". 44 (2004), 10. P. 7-19.

Tammaro – Salarelli (2006) = Alberto Salarelli - Anna Maria Tammaro. *La biblioteca digitale*. Nuova ed. interamente riveduta e ampliata. Milano: Editrice Bibliografica, 2006. (Bibliografia e biblioteconomia ; 57)

Van Nuys (2003) = Carol Van Nuys. *The Paradigma project*. "RLG DigiNews". 7 (Apr. 15, 2003), 2. URL: <[http://www.rlg.org/preserv/diginews/v7\\_n2\\_feature2.html](http://www.rlg.org/preserv/diginews/v7_n2_feature2.html)>.

Van Nuys et al. (2004) = Carol Van Nuys et al. *The Paradigma Project and its quest for metadata solutions and user services [online]*. Contributo presentato al World Library and Information Congress: 70th IFLA General Conference and Council, 22-27 August 2004, Buenos Aires, Argentina. URL: <<http://www.ifla.org/IV/ifla70/papers/009e-Nuys.pdf>>. Anche in "International Cataloguing and Bibliographic Control". 34 (Apr./Jun. 2005), 2. P. 29-35.

Van Nuys (2005 a) = Carol Van Nuys. *The Norwegian Digital Library : Easy access to information and knowledge sources*. Contributo presentato al World Library and Information Congress: 71th IFLA General Conference and Council, 14-18 August 2005, Oslo, Norvegia. URL: <[http://www.ifla.org/IV/ifla71/papers/120f\\_trans-vanNuys.pdf](http://www.ifla.org/IV/ifla71/papers/120f_trans-vanNuys.pdf)>

Van Nuys et al. (2005 b) = Carol van Nuys - Ketil Albertsen - Linda Pedersen - Asborg Stendstad. *Paradigma project: il deposito legale delle risorse remote nell'esperienza norvegese*. Trad. italiana di Stefano Gambari e Mauro Guerrini. "Biblioteche oggi". 23 (2005), 1. P. 17-28.

Valacchi (2002) = Federico Valacchi. *Il web per gli archivi e gli archivi nel Web*. "Archivi&Computer". 12 (2002), 3. P. 7-16.

Vitali (2004 a) = Stefano Vitali. *Passato digitale : le fonti dello storico nell'era del computer*. Milano: Bruno Mondadori, c2004. (Le scene del tempo).

Vitali (2004 b) = Stefano Vitali. *Una memoria fragile: il Web e la sua conservazione*. In: Dario Ragazzini (et al.) *La storiografia digitale*. A cura di Dario Ragazzini. Torino: UTET libreria, 2004. P. 101-127.

Vitiello – De Pinedo (1992) = Giuseppe Vitiello - Isa De Pinedo. *Note sul sistema bibliotecario scandinavo*. "Bollettino AIB : rivista italiana di biblioteconomia e scienze dell'informazione". 32 (1992), 1. P. 55-63.

Vitiello (2002) = Giuseppe Vitiello. *Alessandrie d'Europa : storie e visioni di biblioteche nazionali*. Postfazione di Attilio Mauro Caproni. Milano: Sylvestre Bonnard, 2002. (Il sapere del libro)

Weston (2001) = Paul Gabriele Weston. *Catalogazione bibliografica : dal formato MARC a FRBR*. "Bollettino AIB : rivista italiana di biblioteconomia e scienze dell'informazione". 41 (2001), 3. P. 267-286

Weston (2002) = Paul Gabriele Weston, *Tra controllo bibliografico e controllo intellettuale: qualche riflessione sui sistemi bibliografici elettronici*. In: *Risorse elettroniche (2001) = Le risorse elettroniche: definizione, selezione, catalogazione = Electronic resources. Definition, Selection, Cataloguing*: atti del convegno internazionale, Roma, 26-28 novembre 2001. A cura di Mauro Guerrini, con la collaborazione di Stefano Gambari e Lucia Sardo, Milano: Editrice Bibliografica, 2002.