

# LA CONSERVAZIONE DIGITALE NELLO SCENARIO EUROPEO E INTERNAZIONALE

## PRINCIPI, METODI, PROGETTI

Roma, novembre 2003

Maria Guercio

Il tema della conservazione digitale è all'ordine del giorno della comunità scientifica da quasi un decennio come ci ricordano gli atti di un convegno internazionale organizzato dalla Fondazione IBM Italia del 1993 significativamente intitolato *L'eclisse delle memorie*. Si è tuttavia fatta ben poca strada da allora in termini operativi e organizzativi, se si escludono le iniziative condotte da parte delle istituzioni di conservazione e degli organismi comunitari che da qualche anno finanziano progetti e iniziative sul tema sia pure limitatamente alla creazione di reti di eccellenza piuttosto che in riferimento a veri e propri progetti di ricerca e sperimentazione. E' il caso ad esempio del progetto ERPANET che ha ottenuto un finanziamento triennale con l'obiettivo di dar vita a una rete di eccellenza per la formazione e informazione degli operatori impegnati nel campo della conservazione ([www.erpanet.org](http://www.erpanet.org)) e del progetto DELOS orientato allo sviluppo di iniziative di *digital library* e recentemente ri-finanziato per gli anni 2004-2007 anche in riferimento a un nucleo di attività dedicate alla conservazione digitale. Assenti sono stati invece i produttori di software, distratti gli amministratori, sostanzialmente disattenti i legislatori nazionali (non solo quello italiano) che hanno concentrato (soprattutto il primo) l'attenzione sui problemi della riproduzione sostitutiva e sulle modalità di autenticazione dei documenti informatici, confondendo spesso le soluzioni orientate alla identificazione della provenienza certa e all'integrità delle risorse digitali nel momento della loro formazione e trasmissione con quello molto più complesso del mantenimento nel lungo periodo dell'autenticità e della definizione degli strumenti che ne consentano a distanza di tempo l'accesso e la necessaria valutazione critica.

Sembra oggi di avvertire un cambiamento significativo nella consapevolezza dell'urgenza e della rilevanza del problema, legato anche alla constatazione che – senza la garanzia della validità giuridica e del valore *permanente* di testimonianza dei documenti informatici – gli stessi investimenti di settore sostenuti dalle pubbliche amministrazioni e dalle imprese, quelli già realizzati e ancor più quelli attesi, rischiano di perdere la loro rilevanza. Numerose sono del resto le iniziative che sembrano preannunciare un cambiamento nel senso ora ricordato e costituire una significativa manifestazione di interesse che non nasce solo dalle preoccupazioni delle comunità degli archivisti e dei bibliotecari, ma trova conforto ormai anche in altri settori professionali e istituzionali.

Il problema è peraltro complesso, richiede investimenti nella ricerca, collaborazione, interdisciplinarietà. Merita perciò qualche considerazione introduttiva e la descrizione sintetica, un primo panorama di ciò che si muove in questo ambito sia a livello internazionale ed europeo che – purtroppo assai limitatamente e senza un consistente e adeguato sostegno istituzionale - sul piano nazionale.

E' evidente che problemi di così ampia e diversificata natura richiederebbero un'analisi approfondita e, soprattutto, avrebbero bisogno di specifici interventi formativi inseriti nei programmi universitari specialistici (master e lauree di II livello)<sup>1</sup>, considerato il fatto che per la loro corretta

---

<sup>1</sup> Un master dedicato alla gestione di sistemi documentari avanzati che include moduli sulla gestione e conservazione di memorie digitali e sull'uso di linguaggi di marcatura è stato avviato per l'anno accademico 2002-2003 e riproposto per il 2004 presso l'Università degli studi di Urbino e prevede lo sviluppo di una serie di corsi on-line ([www.uniurb.it/maslet](http://www.uniurb.it/maslet)). Il

trattazione sono necessari investimenti formativi importanti e in campi differenziati (esegesi delle fonti storiche, metodologie della ricerca documentaria, di informatica giuridica e informatica applicata ed elaborazione di sistemi informativi, ecc.). Nella lunga fase di transizione che attraversiamo in troppi ambiti oltre a quello dell'innovazione tecnologica (dalla riforma del modello amministrativo italiano alle trasformazioni del sistema formativo universitario, ai modi della comunicazione interpersonale e interistituzionale, agli strumenti della didattica, ecc.) sono molteplici e tutti importanti i nodi concettuali e organizzativi che richiedono di essere sciolti di volta in volta: la questione della produzione e formazione di memorie digitali sostitutive dei documenti tradizionali non può essere affrontata concretamente e seriamente senza che molti di essi abbiano effettivamente trovato una dimensione stabile, in grado di orientare gli operatori e i progettisti di sistemi per un uso adeguato delle risorse tecnologiche. E' soprattutto a questi rilevanti interrogativi che è necessario dare risposta al più presto, senza avere la pretesa naturalmente di trovare soluzioni definitive, ma con l'ambizione di fornire qualche ulteriore elemento di riflessione e di conoscenza in un panorama che è sempre più ricco di informazioni, ma sempre più dispersivo e frammentario.

In particolare, mi sembra opportuna in questa sede una riflessione sia pur breve sulla natura del cambiamento che l'innovazione tecnologica introduce nel campo specifico della conservazione delle memorie documentarie che costituiscono da secoli, se non da millenni il luogo privilegiato della memoria delle comunità e degli individui.

E' innanzi tutto indispensabile circoscrivere la questione definendo il significato e i contenuti della conservazione delle risorse digitali a lungo termine come l'insieme delle attività e degli strumenti che assicurano che i documenti informatici siano mantenuti accessibili, utilizzabili (leggibili e intelligibili) e autentici (univocamente identificabili e integri) nel medio e nel lungo periodo, in un ambiente tecnologico *certamente* diverso da quello originario.

In questo contesto la conservazione non può consistere nella semplice conservazione del flusso di bit, poiché implica il mantenimento di informazioni (metadati specifici descrittivi e gestionali) in grado di assicurare la possibilità di interpretazione futura del flusso medesimo (contenuto strutturato, configurazione degli elementi, contesti multipli, comportamenti), né è riducibile a procedure e comportamenti omologati all'ambiente tradizionale o a procedure uniformi. Tanto meno si identifica – come si è già ricordato - con la riproduzione ottica sostitutiva o con le procedure di digitalizzazione che costituiscono solo eventuali utili strumenti al servizio della conservazione.

In sostanza la funzione conservativa si configura sempre più nettamente come un complesso articolato e dinamico di attività, strumenti, procedure che richiedono principi chiari, un quadro normativo di riferimento, luoghi significativi di sperimentazione. Sui principi e sulla struttura concettuale di base, non è semplice – e forse neppure auspicabile – far riferimento a una dimensione unitaria della questione, anche se la convergenza delle memorie digitali è fenomeno destinato a produrre effetti simili anche per il trattamento delle risorse: non nel senso di cancellare le differenze funzionali delle specifiche tipologie di oggetti, ma in relazione ad alcuni aspetti trasversali della gestione – ad esempio nel trattamento complessivo dei metadati per la conservazione (che pur sono alquanto diversi nel merito e nelle modalità di aggregazione e di aggiornamento).

E' la natura stessa della memoria digitale che implica una progressiva – anche se non definitiva – indifferenziazione del trattamento gestionale che si accompagna al progressivo venir meno del mantenimento del contesto originario. Mentre il documento convenzionale ha una storia che lascia segni più o meno significativi in grado comunque di sostenere e accompagnare i futuri processi di

verifica della sua credibilità e di cui ciascun ambito di intervento dovrà tenere conto, dato che l'oggetto si conserva *inalterato* nella sua fisicità, le risorse digitali sono al contrario per natura soggette a un continuo (più o meno frequente, ma comunque inevitabile) processo di trasformazione che ne determina una riproposizione in ambienti e contesti solo in parte fedeli all'originale. Accanto ai rischi di perdite e manipolazioni, è insomma necessario considerare anche l'impatto dei processi conservativi sulla specificità stessa dei documenti e sulla loro configurazione logica e fisica.

La dematerializzazione delle fonti documentarie prodotte e mantenute informaticamente costituisce allo stesso tempo il vantaggio ma anche il limite della dimensione informatica delle nuove memorie: è di grande rilevanza disporre di fonti facilmente e rapidamente migrabili su altri supporti analoghi o migliori, in altri ambienti, su altre piattaforme, ma è inevitabile e rilevante il limite intrinseco di materiali che comunque richiedono – anche se i produttori della fonte hanno ben operato - una molteplicità di mediazioni, strumenti e risorse per poter essere utilizzati (letti e compresi) da utenti anche contemporanei alla loro produzione.

In questo contesto l'affermazione per cui “la conservazione digitale è un processo complesso” acquista una dimensione specifica con rilevanti conseguenze sia organizzative che tecniche e con contraddizioni piuttosto evidenti, a cominciare dalla duplicità concorrente delle esigenze stesse degli utenti, che da un lato necessitano del mantenimento della forma originaria, dall'altro esprimono l'esigenza di un accesso dinamico e interattivo. Non c'è dubbio che i responsabili della conservazione e gli utenti medesimi avvertono la pericolosità e la difficoltà di questa fase, che a sua volta muta in continuazione proponendoci sfide sempre più impegnative. In questo contesto la cooperazione è una indispensabile risorsa, che tuttavia implica costi che non possono essere sottovalutati pur presentando risvolti positivi in termini di formazione permanente e qualificazione delle risorse professionali interne.

Sul piano teorico il primo ambito da curare con attenzione è quello relativo alla identificazione degli elementi e delle procedure che garantiscono l'autenticità, cioè l'identità e l'integrità dei documenti elettronici, per ciascun ambiente di riferimento sull'esempio delle attività di ricerca condotte in materia archivistica dal progetto InterPARES<sup>2</sup>, ma anche dal progetto DOCT realizzato dagli Archivi nazionali degli Stati Uniti e dalla Library of Congress in collaborazione con il S.Diego Supercomputer Center<sup>3</sup> e, in campo biblioteconomico, dal progetto Nedlib che ha coinvolto tra le biblioteche nazionali europee anche la Biblioteca nazionale centrale di Firenze. Da sottolineare che in tutti i casi citati si è fatto riferimento – proprio per definire l'architettura dei metadati e dei flussi informativi – a un modello concettuale approvato come standard ISO – l'Open Archives Information System - elaborato dalle agenzie aerospaziali per la conservazione dei dati e dei documenti di ricerca scientifica prodotti in ambiti differenziati.

Più complesso - almeno per ora – è la dimensione pratica e organizzativa della questione che, per la mole degli investimenti richiesti e per l'insufficiente esperienza maturata, deve essere accompagnata, se non addirittura preceduta

- da iniziative “politiche” di sensibilizzazione indirizzate al legislatore, ai produttori, all'opinione pubblica per poi tradurle solo con il tempo in raccomandazioni, in linee guida, in normative e regolamentazioni opportunamente verificate e mature,
- dal riconoscimento da parte degli istituti di produzione e conservazione delle risorse digitali della centralità dei problemi organizzativi che implicano l'adozione di procedure adeguate e di standard

---

<sup>2</sup> Per i materiali predisposti nell'ambito del progetto e per i rapporti conclusivi si veda il sito del progetto: [www.interpares.org](http://www.interpares.org), nonché la traduzione delle relazioni finali di alcune task force pubblicate in “Archivi & Computer”, 2003, 1-2.

<sup>3</sup> Si veda il sito del progetto <http://www.npaci.edu/DICE/Pubs>.

e formati orientati ai dati e “persistenti” (ad esempio XML), la definizione di regole condivise, l’individuazione di responsabilità certe e riconosciute,

La valutazione, analisi e contenimento dei costi in relazione agli obiettivi e ai metodi possibili o necessari per assicurare la fattibilità della conservazione costituisce una componente esplicita e cruciale dell’intero processo, così come la riqualificazione “di massa” del personale tecnico mediante programmi di formazione permanente che includano l’aggiornamento dei contenuti e l’uso di strumenti didattici avanzati.

Anche se molta strada resta da percorrere in questo ambito, alcuni punti fermi si possono considerare acquisiti, almeno per quel che riguarda gli attuali protagonisti delle iniziative e dei progetti conservativi. In primo luogo è ormai condivisa la consapevolezza che la conservazione digitale implichi responsabilità condivise e precoci e un’adeguata e flessibile produzione sia di regole di natura organizzativa che di standard tecnici di facile utilizzo anche da parte dei soggetti produttori di documenti di piccole dimensioni. Larga parte delle informazioni che garantiscono l’accesso e la verifica dell’autenticità sono disponibili *solo* nella fase attiva di produzione e gestione delle risorse digitali (ad esempio i sistemi di classificazione, i metadati di contesto, lo schema logico dei database, la documentazione delle applicazioni, ecc.). La perdita o il tardivo recupero di tali meta-informazioni possono compromettere irrimediabilmente la possibilità stessa della conservazione o farne crescere notevolmente i costi.

Date queste premesse, è chiaro che sono a grave rischio tutti quei documenti che non si producono in un contesto organizzativo definito e, quindi, controllato e regolamentato, con particolare riferimento alla documentazione di natura scientifica e personale e ai nuovi sistemi elettronici dinamici e interattivi (ad esempio ai sistemi web e ad alcune tipologie di database). Nella fase attuale il nodo critico, a livello europeo e ancor più nazionale, riguarda la mancanza di infrastrutture scientifiche per la ricerca, con particolare riferimento alla necessità di disporre di risorse dedicate, laboratori permanenti e centri di competenza di qualità e livello internazionali, all’esigenza di condividere esperienze di ricerca e soluzioni applicative, di dare continuità degli investimenti, di definire strategie di indagine capaci di adattarsi all’evoluzione tecnologica.

Lo stato dell’arte in questo campo non è particolarmente positivo: da un lato la letteratura di riferimento è sovrabbondante, ma di difficile valutazione e quindi dispersiva, dall’altro le ricerche e le sperimentazioni promettenti fanno fatica a tradursi in indicazioni chiare per chi opera oppure si limitano a qualche elementare indicazione pratica sul monitoraggio dei supporti, dei depositi digitali, del livello di obsolescenza. Sebbene sia ormai riconosciuta (per contenere i costi e mantenere la qualità degli oggetti) la necessità di una approfondita e continua riflessione concettuale dei principi, degli strumenti e dei metodi per il trattamento documentario, ma anche di un’analisi critica degli interventi e dei metodi di conservazione finora proposti (emulazione, migrazione in formati standard, ecc.), le iniziative europee sono ancora scarse, quelle nazionali del tutto inesistenti.

La complessità dei problemi determinati dall’obsolescenza richiede invece un impegnativo e mirato sostegno di lungo periodo che si configuri nella predisposizione di *programmi coordinati* più che di *progetti*, in grado di rafforzare – il che equivale spesso a consentire - la cooperazione internazionale e la partecipazione dei ricercatori europei a indagini promettenti, sostenere la diffusione di informazioni qualificate e favorire l’aggiornamento professionale in un settore in cui è difficile orientarsi tra la grande quantità di fonti e materiali, avviare seri (quindi costosi) progetti di ricerca che coinvolgano ambienti disciplinari diversi finora poco attenti alle esigenze della conservazione permanente, garantiscano una effettiva attività di sperimentazione, includano la partecipazione di società di informatiche realmente impegnate nelle attività di ricerca e predispongano schemi e materiali per lo sviluppo di linee guida e raccomandazioni e favoriscano l’indispensabile sostegno

“politico” a interventi complessi la cui necessità è scarsamente percepita dai governi.

E' inoltre indispensabile riconoscere (in sede di normativa europea e nelle politiche nazionali per la conservazione) che la conservazione è un processo e presuppone responsabilità molteplici e che la responsabilità dei produttori di documenti e dei produttori di software è tanto significativa quanto sconosciuta dagli attori medesimi, i cui sforzi vanno perciò sostenuti con interventi diversificati che tengano conto delle specifiche criticità di settore, a cominciare dalla creazione di un sistema di network e di relazioni stabili e affidabili tra soggetti di ricerca e istituzioni in grado di favorire il coordinamento e l'integrazione di conoscenze eterogenee in contesti caratterizzati da complessità, alta incertezza e numerose fonti di conoscenza, ridurre costi e inefficienze di una formazione auto-referenziale, assicurare l'utilizzo di processi di apprendimento derivanti da varie fonti e favorisce la cooperazione come fonte di “varietà strategica e conoscitiva”, garantire l'accesso a soluzioni e strumenti sviluppati da altri soggetti, facilitare il trasferimento di conoscenza “ricca”, complessa e non codificata, altrimenti non separabile dagli attori che la possiedono.

In questo contesto si è collocato il progetto europeo Erpanet (Electronic Resource Preservation and Access Network, 2002-2004), cui è stato affidato il compito impegnativo di superare l'isolamento degli operatori, soprattutto delle istituzioni di piccole dimensioni e garantire la multidisciplinarietà delle iniziative, coinvolgere come nodi di una rete europea i centri di ricerca e le istituzioni che conservano il patrimonio culturale digitale, ma anche le società di informatica, il mondo economico, la pubblica amministrazione al fine di scambiare e condividere conoscenze e idee rendendole disponibili anche a coloro che non partecipano attivamente ai progetti di ricerca e promuovere lo sviluppo e l'uso di standard e best practice. Senza entrare nei dettagli dell'iniziativa le cui attività sono rese tutte ampiamente disponibili grazie a un uso dinamico della rete<sup>4</sup>, vale la pena ricordare che le linee di intervento e i prodotti di maggior impegno di Erpanet riguardano la messa a disposizione attraverso un lavoro di cernita e valutazione delle fonti dei principali documenti di ricerca, delle linee guida, degli standard più significativi per la conservazione, cui si aggiungono la predisposizione di schede descrittive di sintesi della letteratura di settore, l'elaborazione di strumenti e linee guida per gli aspetti rilevanti della funzione conservativa (gestione del rischio, analisi dei costi, selezione di tecnologie e metodi, definizione di politiche generali e aziendali, identificazione dei metadati necessari alla conservazione, analisi delle funzioni di versamento/acquisizione di fonti digitali, ecc.), l'organizzazione di attività formative di diverso livello specialistico (workshop per esperti) e divulgativo (seminari per un pubblico di operatori e utenti interessati al tema per ragioni professionali)<sup>5</sup>.

Sebbene gli impegni progettuali siano essenziali soprattutto per far crescere la consapevolezza delle comunità interessate e il livello di competenza degli addetti, non c'è dubbio che la dimensione del problema e la sua complessità richiedano iniziative rilevanti, in primo luogo – come si è già peraltro ricordato – di natura politica. L'Unione europea si è mossa in questa direzione promuovendo una serie di incontri tra esperti, avviando una concertazione tra gli Stati che in questi ultimi due anni hanno assunto la presidenza dell'Unione. Anche l'Italia, nonostante ritardi e limiti piuttosto pesanti per quanto riguarda i problemi strutturali e infrastrutturali relativi alla ricerca in tutti i campi – grazie all'iniziativa della Direzione generale per i beni librari e dell'ICCU in particolare ha saputo giocare un ruolo

---

<sup>4</sup> Tutte le iniziative sono ampiamente documentate sul sito del progetto: [www.erpanet.org](http://www.erpanet.org).

<sup>5</sup> Il progetto, nel periodo in cui è finanziato, è chiamato a organizzare 6 workshop per esperti del settore e 6 seminari di aggiornamento (cfr [erpaworkshop](http://erpaworkshop) sul sito [www.erpanet.org](http://www.erpanet.org)). Allo stato attuale si sono tenuti i seguenti eventi: Toledo (22-23 giugno 2002) sulla digitalizzazione, Urbino (9-11 ottobre 2002) su XML per la conservazione digitale, Copenhagen (11-12 dicembre 2002) sullo standard OAIS, Berna (9-11 aprile) sulla conservazione dei database, Kerkira (22-24 maggio 2003) sulla conservazione dei materiali web, Marburg (3-4 settembre 2003) sui metadati per la conservazione, Roma (17-19 novembre 2003) sul *trusted digital repository*.

“politico” di un certo peso. Un primo risultato di quest’azione congiunta è stato il documento approvato nel giugno 2002 dai ministri della cultura che hanno riconosciuto in quell’occasione la complessità della questione e hanno cominciato a porre le basi per un piano d’azione operativo già in parte delineato da un gruppo di lavoro di esperti che si è già riunito a Bruxelles nel marzo 2002 e che ha lavorato via web per la preparazione di un’agenda di lavoro approvata in occasione della Conferenza internazionale sulla conservazione organizzata a Firenze nei giorni 16-17 ottobre 2003.

“Digital resources will not survive or remain accessible by accident. The nature of digital content is such that it requires a proactive approach from the institutions concerned if it is to be usable in the future”. Il riconoscimento della complessità del problema ha implicato – in sede europea, - l’individuazione dei fattori chiave di successo: lo sviluppo di standard, la qualità dei prodotti e l’attenzione ai bisogni degli utenti, la capacità di sviluppare l’analisi dei costi e di valutare la fattibilità dei programmi di conservazione, l’individuazione di responsabilità precise all’interno dei produttori, la definizione di policy nazionali e aziendali, lo sviluppo di programmi di sensibilizzazione e formazione adeguati. In particolare in occasione dell’incontro di Bruxelles sono emersi alcuni preziosi suggerimenti per un’azione comunitaria da condurre anche in parte all’interno dei singoli Stati:

- stabilire un’unica piattaforma di coordinamento per tutti gli Stati membri,
- sviluppare linee strategiche e valutarne l’impatto sulle legislazioni nazionali,
- sviluppare iniziative di aggiornamento e sensibilizzazione,
- identificare i bisogni di nuove competenze professionali,
- promuovere lo sviluppo e la conoscenza di best practice e di standard,
- sostenere la creazione di reti che assicurino la comunicazione e la cooperazione tra le diverse comunità di interesse coinvolte nella conservazione del patrimonio digitale,
- implementare metodologie per la definizione di parametri comuni per definire modelli di analisi costi/benefici.

La conferenza di Firenze è stato il passo successivo nello sforzo di costruire un processo continuativo di iniziativa in questo ambito con i pochi e inadeguati strumenti organizzativi e finanziari disponibili. Innanzi tutto si è predisposto un quadro delle normative e delle regolamentazioni anche tecniche di settore, una sorta di *stato dell’arte* che si è concentrato soprattutto sugli Stati dell’Unione, ma che ha anche coinvolto l’Australia, il Canada, gli Stati Uniti<sup>6</sup>. In quell’occasione è stato presentato anche un dossier sulle emergenze, che raccoglie sei studi di casi di interventi eccezionali e costosi – in alcuni casi impossibili – di recupero delle memorie digitali che le nostre comunità hanno già prodotto e utilizzato da decenni. L’obiettivo immediato di informare da un lato e “spaventare” dall’altro l’opinione pubblica qualificata che ha potenzialmente in mano leve e strumenti di intervento in questo campo è sembrato agli organizzatori un primo passo di un processo di sensibilizzazione e di iniziative che si preannuncia comunque necessariamente lungo e articolato. Il fine più generale è il riconoscimento – in termini anche operativi – del fatto che la rivoluzione tecnologica non consiste semplicemente nel fornire informazione digitale per il presente, ma nell’assicurare al presente una informazione ricca e affidabile, capace perciò di costituirne anche la memoria credibile e duratura.

---

<sup>6</sup> Si veda il dossier predisposto per la conferenza citata, a cura di Erpanet e dell’ICCU, *Normative e linee d’azione per la conservazione delle memorie digitali. Un’indagine conoscitiva*, Firenze 2003.