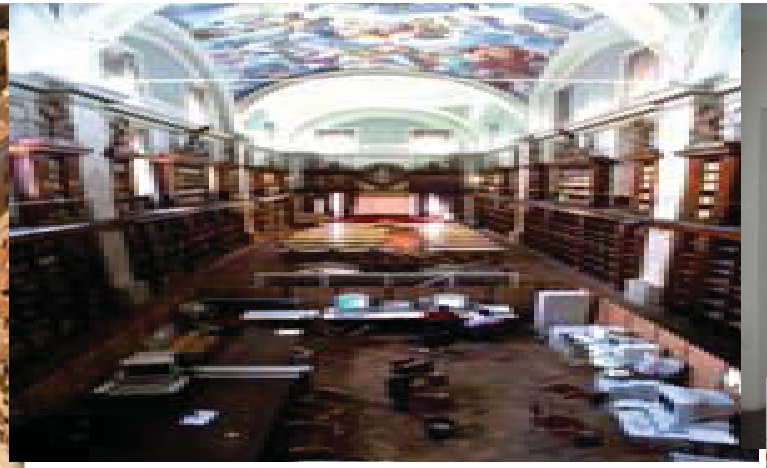
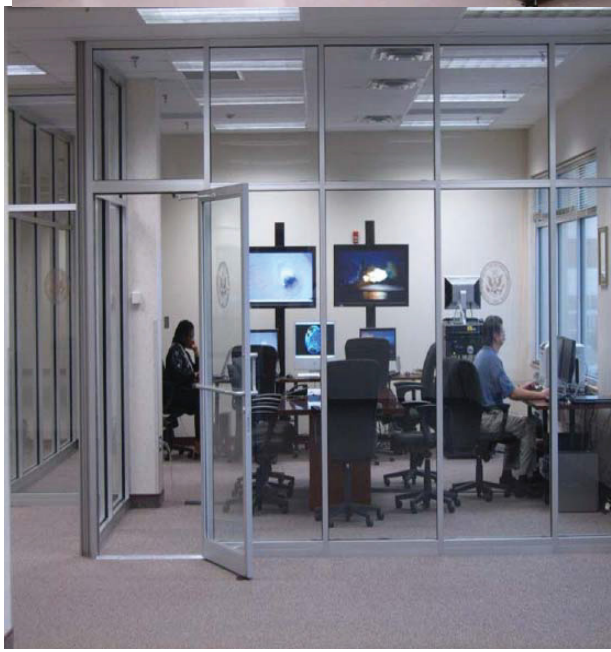


Valutare l'autenticità delle fonti digitali
Le nuove responsabilità degli istituti di conservazione

Mariella Guercio

Cagliari, novembre 2009



Il tema oggetto di riflessione

- Il nodo della conservazione autentica dei contenuti digitali e della possibilità di valutarne il grado di affidabilità è sempre più rilevante in ragione della crescita quantitativa e qualitativa del patrimonio digitale ma anche per la progressiva complessità della sua tenuta
- Gli strumenti per tale valutazione sono ancora in fase di individuazione anche se alcune conclusioni metodologiche sembrano acquisite

Il ruolo dell'archivistica

- La disciplina archivistica offre (come e più di prima) principi e metodi di analisi al servizio degli storici e dei ricercatori non solo in quanto custode imparziale delle memorie digitali, ma anche in quanto supporto (transitorio) per la valutazione critica delle fonti stesse.

I rischi e i costi della conservazione digitale - 1

- dati persi sui **contributi pensionistici per 20 milioni** di giapponesi
- dati persi della **missione su Marte** del 1975 (un miliardo di dollari): impossibilità di decifrare i formati; necessità di inserire nuovamente tutti i dati disponibili su supporto cartaceo nel sistema
- 12 **bad pratice** nel dossier predisposto dall'ICCU nel 2003, *Conservazione delle memorie digitali: rischi ed emergenze* (Alessandra Ruggiero)
- il **10% dei documenti elettronici** prodotti dal governo canadese non sono più leggibili

I rischi e i costi della conservazione digitale - 2

- Secondo un recente studio del National Science Foundation (US) il **costo medio per la ricostituzione di 20 MG di dati è di $ 64,000**
- Nel maggio 2006, la società finanziaria Morgan Stanley ha concordato di **pagare $ 15 milioni per non aver presentato** alla Securities and Exchange Commission decine di **migliaia di e-mail precedenti a dicembre 2000**

Alcune indicazioni introduttive - 1

- "per continuare a fare storia, per conoscere in futuro *questo* presente bisogna preoccuparsi di mantenere le possibilità di poterlo fare" (Giuva, Vitali, Zanni Rosiello)
- La fragilità degli archivi digitali deve essere compresa sia da chi li produce che da chi li conserva e da chi li utilizza

Alcune indicazioni introduttive - 2

Ai fini della ricostruzione storiografica nell'era della divulgazione di massa e della produzione di memoria digitale, è rilevante per lo storico (ma anche per l'archivista che opera per la salvaguardia delle fonti) "il **richiamo:**

- **alle regole del mestiere e deontologiche**,
- allo sforzo di una valutazione oggettiva dei fatti,
- alla capacità di distinguere tra giudizio e pregiudizio;
- **all'analisi critica e al rigore filologico**, **all'analisi del contesto di produzione delle fonti**;
- alla conoscenza e all'uso corretto della storiografia e
- alla **consapevolezza degli strumenti culturali e disciplinari indispensabili**, nonché alla **competenza per utilizzarli**;
- al dovere dello storico di non essere parziale e ideologico" (Paola Carucci).

Alcune indicazioni introduttive - 3

Tali impegnativi richiami acquisiscono un'ulteriore pregnanza per la dimensione digitale, che richiede sia allo storico che all'archivista attenzione e cautela di gran lunga superiori a quelle necessarie per utilizzare gli archivi cartacei.

Alcune indicazioni introduttive - 4

- La fragilità dei contenuti digitali dipende infatti da una molteplicità di fattori:
  - dalla **bidimensionalità** più che dalla immaterialità dei documenti che rende impossibile la stratificazione fisica dei segni del tempo sul documento e l'esercizio della critica delle fonti secondo il metodo tradizionale finalizzato a valutare le modifiche e le perdite (accidentali o dolose) su oggetti mantenuti negli archivi nella loro originalità evidente o ricostruibile

Alcune indicazioni introduttive - 5

- dalla certezza, più che dai rischi, **dell'obsolescenza tecnologica** le cui conseguenze (in termini di accessibilità nel tempo e conservazione autentica di originali) implicano la necessità di interventi di migrazione che introducono modifiche nei bit e assicurano la conservazione nella migliore delle ipotesi di *copie autentiche* se non, addirittura, di *componenti digitali autentiche* in grado di essere riprodotte e visualizzate a richiesta dell'utente

La centralità del principio di autenticità e l'esigenza di una definizione condivisa

- Per gli archivisti esiste **una *definizione* ormai acquisita**, che presenta importanti implicazioni in termini di attività da svolgere e di responsabilità da assumere
- E' una definizione utile anche per altri ambiti, anche se le diverse comunità professionali (informatici e giuristi in particolare) non sempre riconoscono la necessità di una ***convergenza*** nell'adozione di definizioni coerenti (ad esempio nel caso del *digital forensic*)

Autenticità: identità e integrità dei documenti - 1

- "The trustworthiness of a record as a record; i.e., the quality of a record that is what it purports to be and that is free from tampering or corruption" (InterPARES)

  L'autenticità di un documento implica la sua identità e la sua integrità (ovvero la gestione e tenuta continuativa e autorevole dei contenuti, delle relazioni, dei contesti)

Autenticità: identità e integrità dei documenti - 2

- L'**identità** di un documento riguarda gli attributi che consentono di caratterizzare **univocamente** un documento:
  - i nomi delle persone che concorrono nella sua formazione,
  - le date di formazione (creazione, ricezione, tenuta) e di trasmissione
  - l'indicazione dell'atto/fatto rappresentato o in cui partecipa
  - l'espressione del vincolo archivistico (indice di classificazione o altro codice univoco di identificazione)

Autenticità: identità e integrità dei documenti - 3

- L'**integrità** di un documento si riferisce alla sua completezza e intangibilità negli *aspetti essenziali*
- L'autenticità è **presunta** sulla base del grado di requisiti cumulativamente rispettati da un sistema di conservazione: più alto è il numero dei requisiti rispettato maggiore è il grado di presunzione dell'autenticità

L'inutilità di strumenti tecnologici complessi nella conservazione permanente

- **Complessità/inutilità/fragilità dei meccanismi di cifratura digitale** basati sulla conservazione del flusso originario di bit (es. firma digitale) a fini conservativi:
  - difficoltà di valutare – a distanza di tempo – la semantica dei meccanismi utilizzati (per funzioni e fini diversi da persone diverse),
  - complessità nella conservazione dell'algoritmo e del software di verifica, rischi nella durata delle Public Key Infrastructure;
  - impossibilità di ricostruire il documento perduto sulla base dell'algoritmo di cifratura
  - assenza di garanzie sul fatto che anche il valore dell'impronta non sia stato manipolato

La conservazione digitale è un processo dinamico - 1

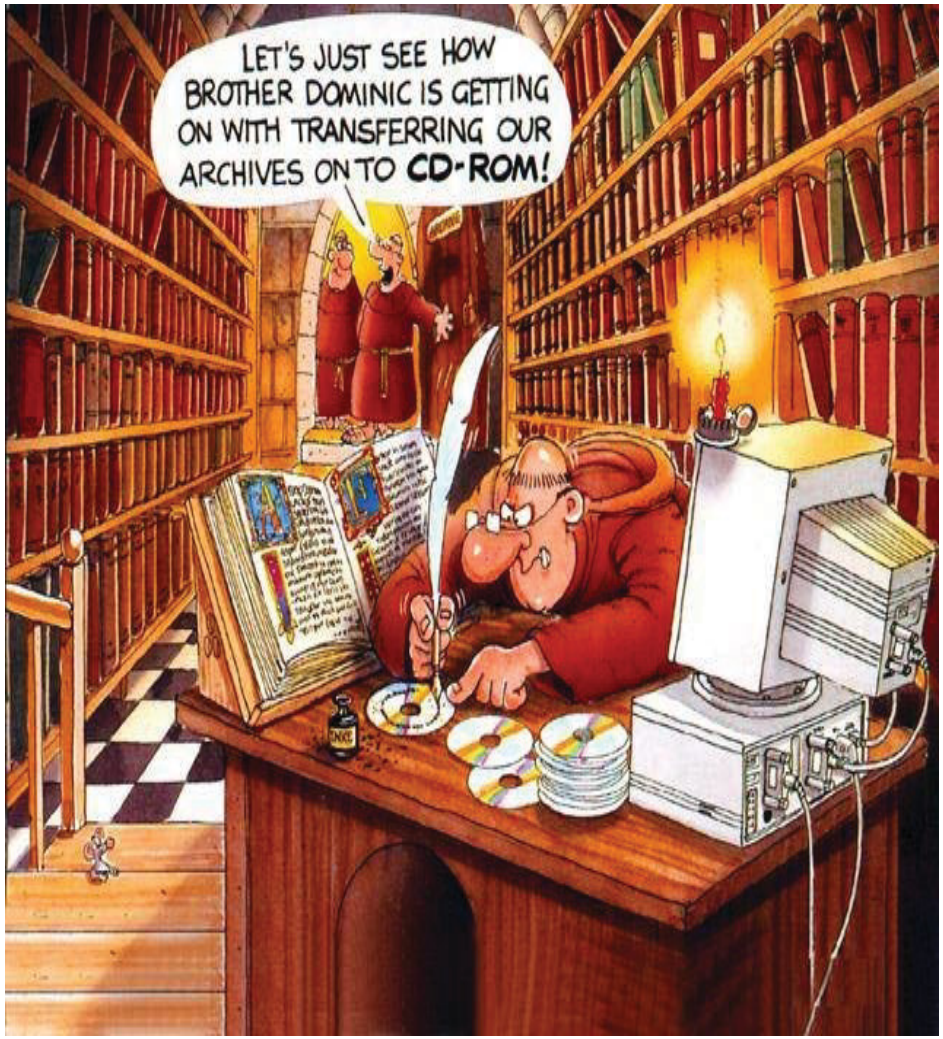
- La conservazione digitale non si riduce al mantenimento statico dei bit originari né alla tenuta dei supporti, ma è un **processo dinamico** finalizzato a mantenere l'accesso alle fonti:
  - richiede migrazioni e copiature, normalizzazione dei formati, cattura (il più possibile automatica) del maggior numero possibile di metadati;
  - implica perdite;
  - si traduce in mantenimento di copie autentiche;
  - implica investimenti organizzativi adeguati

La conservazione digitale è un processo dinamico - 2

- Le soluzioni sono ancora oggetto di ricerca anche se è ormai assodato che la conservazione richiede la **verifica dell'autenticità** e che l'autenticità è valutabile se il sistema di formazione dei documenti e i modi concreti di gestione sono *documentati* in modo *completo in tutte le fasi del ciclo di vita del documento* e se la **documentazione** è consultabile e comprensibile.

La conservazione digitale implica responsabilità precoci e pianificazione

- I processi di **documentazione** devono essere pianificati nelle diverse fasi di trattamento:
  - la conservazione è possibile a costi ragionevoli se è **pianificata** e **gestita precocemente** nell'intero ciclo di gestione degli archivi



Servizi dedicati e profili di competenza

- Sono necessari investimenti significativi per la creazione di **servizi** e per la gestione di **profili di competenza** adeguati che includano:
  - policy precoci,
  - formati adeguati orientati ai dati e alla persistenza,
  - responsabilità certe e diffuse,
  - regole e procedure stringenti anche in caso di esternalizzazione,
  - analisi dei costi e dei rischi,
  - riqualificazione delle componenti professionali tecniche

Il ruolo crescente e insostituibile del soggetto produttore

- Nel definire profili per la formazione e gestione dei documenti e delle aggregazioni archivistiche
- Nel definire e documentare con continuità e ricchezza di dettagli le policy e le responsabilità degli enti

L'arretratezza delle soluzioni applicative

- Gli strumenti disponibili non sono ***ancora*** in grado di fornire risposte operative ai quesiti sull'autenticità, né le esperienze maturate ci offrono procedure standardizzate soprattutto nel caso delle CMO (Cultural Memory Organizations) di dimensioni piccole o medie che conservano quantità crescenti e scarsamente governati di patrimoni digitali 'pregiati' (progetti innovativi in UK – progetto PARADIGM e a Vienna – progetto PLANETS)

Le scelte e le criticità organizzative

- Infrastrutture di conservazione digitale che prevedono nodi e partner numerosi
- Le strategie possibili:
  1. *gestione decentralizzata* nella continuità del principio del policentrismo della conservazione (scarsi controlli e generale trascuratezza): è una ipotesi impraticabile,
  2. *modelli di concentrazione consortile sul territorio* (organizzativamente complesso): è la soluzione preferibile per gli enti e il deposito volontario,
  3. *sistema decentralizzato ibrido* (flessibile, meno invasivo): è sostenibile – se sostenuto da servizi - nel caso di enti di piccole e medie dimensioni e per gli individui singoli poco propensi ad affidare a terzi il loro patrimonio documentario e di dati (tanto più se digitali).

la custodia affidata agli archivi: un sostegno 'filosofico'

- "è […] in questa *domiciliazione*, in questa assegnazione stabile di una dimora, che hanno luogo gli archivi" […] Niente archivio senza un luogo di consegna, senza una tecnica di ripetizione e senza una certa esteriorità" (J. Derrida, *Mal d'archivio*, p. 13 e p. 22)
- "Il testimone affidabile è quello che può mantenere la sua testimonianza nel tempo […]. L'archivio si presenta come un luogo fisico che protegge il destino della traccia documentaria […]. L'archivio non è soltanto un luogo fisico, spaziale, ma anche un luogo sociale […]" (P. Ricoeur, *La memoria, la storia, l'oblio*, p. 231, 234)

Le soluzioni conservative: le condizioni per la soluzione accentrata

- trattamento *precoce* delle fonti
- *condivisione* delle responsabilità :
  - per la *creazione* di "depositi istituzionali" e archivi digitali correnti e di deposito per gli enti
  - per il *passaggio di custodia* - versamento
  - per la *istituzione* e la *gestione di depositi* (centralizzati) di conservazione a lungo termine
  - per la realizzazione di *servizi di supporto* alle istituzioni medio-piccole e agli individui
- *La soluzione è sempre più largamente ricondotta a un problema di **fiducia** in chi esercita la funzione d'archivio e richiede notevoli investimenti*

Il ruolo e la natura dei depositi d'archivio digitali: il nodo della fiducia

- La conservazione digitale richiede pratiche quotidiane e una **infrastruttura** adeguata sul piano organizzativo che includa un sistema distribuito di **depositi digitali *fidati***

- Gli enti che assumono il compito della conservazione a lungo e medio termine svolgono un **servizio pubblico** (definiscono regole, sostengono processi di informazione e formazione degli utenti, assicurano la funzione conservativa, promuovono la ricerca)

Ma come si definiscono le responsabilità e le funzioni di un sistema di conservazione e come si assicurano processi di certificazione?

Affidabilità dei depositi

- Gli utenti devono poter *misurare/valutare* l'affidabilità di un deposito digitale in termini di:
  - *autenticità dei contenuti* digitali conservati anche in relazione alla qualità dei processi conservativi (documentazione dei processi conservativi, possibilità di ricostruirne la storia, policy, strumenti di monitoraggio interno e di auditing)
  - *capacità di fornire accesso adeguato* ai contenuti medesimi (completezza delle informazioni descrittive, efficienza dei sistemi di reperimento)

e accreditamento

- L'affidabilità e la conseguente possibile certificazione non possono essere il risultato di una auto-dichiarazione
- Dal 2002 la comunità internazionale ha elaborato raccomandazioni e standard che definiscono requisiti e responsabilità necessari alla certificazione dei depositi digitali

- Assicurare la *fiducia* per il deposito vuol dire
  - garantire la "qualità" dei contenuti (documenti, metadati, rapporti);
  - agire in un contesto regolato, conforme a standard fissati dalla comunità internazionale: sulla qualità dell'informazione, sulla sicurezza dell'informazione, sulla gestione della documentazione istituzionale, sul modello del sistema informativo.

Elementi di credibilità e parametri di valutazione

- *E' semplice dichiarare* la propria responsabilità per la cura di archivi digitali; è tuttavia *un impegno tanto costoso e difficile da onorare* quanto è *facile abdicare* successivamente al compito dichiarato (Lynch 2003).
- E' quindi vitale stabilire con chiarezza i criteri che definiscano nel dettaglio gli elementi di credibilità di un deposito (TRAC 2007) e consentano di misurarla sulla base di parametri approvati con uno standard ISO (RAC 2009)

Progetti internazionali di riferimento

- CASPAR (OAIS e architettura per la conservazione)
  - http://www.casparpreserves.eu
- DIGITAL CURATION CENTRE (DCC) (modelli e formazione)
  - http://www.dcc.ac.uk
- DIGITAL PRESERVATION EUROPE (DPE)
  - www.digitalpreservationeurope.eu
- DRAMBORA (pianificazione
  - http://www.repositoryaudit.eu
- ICA-ISO, Record Exchange Standard BRS (Business Requirements Specification)
  - http://www.ica.org/en/node/38983
- INSPECT (Investigating the Significant Properties of Electronic Content Over Time)
  - http://www.jisc.ac.uk/whatwedo/programmes/programme_rep_pres/inspect.aspx
- INTERPARES (autenticità)
  - http://www.interpares.org
- MOIMS-RAC (requisiti per la certificazione)
  - http://wiki.digitalrepositoryauditandcertification.org/pub/Main
- PLANETS (pianificazione dei processi di conservazione)
  - http://www.planets-project.eu
- REPOSITORY INFRASTRUCTURE
  - http://repinf.pbworks.com