



Depositi per la conservazione di archivi digitali: certificazione e autenticità



Maria Guercio
m.guercio@mclink.it

Macerata, 7 maggio 2009

La sfida della conservazione in ambiente digitale richiede il riconoscimento del problema e della sua complessità, nonché la capacità di individuare i requisiti necessari ad affrontarlo

- Il nodo della conservazione a lungo termine delle memorie digitali è ormai riconosciuto in Europa e nel mondo come un fattore rilevante di promozione e sviluppo dei processi di innovazione e miglioramento del sistema Paese (si veda l'*i2010 E-government Action Plan* dell'Unione europea)

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
- PLANETS (pianificazione dei processi di conservazione)
 - <http://www.planets-project.eu>
- DRAMBORA (pianificazione)
 - <http://www.repositoryaudit.eu>
- 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> e moims-rac@mailman.ccsds.org
- REPOSITORY INFRASTRUCTURE
 - <http://repinf.pbworks.com>

Considerazioni preliminari per - 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 quindi migrazioni e copie, normalizzazione dei formati, cattura (il più possibile automatica) dei metadati; implica perdite; si traduce in mantenimento di copie autentiche; implica investimenti organizzativi adeguati
- 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.*

Alcune considerazioni preliminari - 2

- I processi di **documentazione** devono essere pianificati nelle diverse fasi di trattamento
- La conservazione è quindi possibile a costi ragionevoli se è **pianificata** e **gestita precocemente** nell'intero ciclo di gestione degli archivi (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)
- Sono necessari investimenti significativi per la formazione di **profili di competenza** adeguati

I termini generali del processo conservativo - 1

- La soluzione è sempre più largamente ricondotta a un problema di **fiducia** in chi esercita la funzione d'archivio
- La persistenza è ritenuta risolvibile in termini di **accessibilità compatibile** con un livello adeguato di **qualità e controllo** (costi, rischi, fattibilità, grado accettabile di perdita)

I termini generali del processo conservativo - 2

- 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?

Termini e definizioni: le ambiguità di una riflessione complessa

- Deposito digitale
- Deposito digitale istituzionale
- Deposito digitale accreditato o *Trusted Digital Repository (TDR)*: un deposito con il compito di fornire accesso affidabile e a lungo termine agli oggetti digitali in relazione a una specifica comunità di riferimento (OAIS)
- Deposito per la conservazione a lungo termine
- Archivio digitale (OAIS): organizzazione responsabile per la conservazione digitale

Deposito digitale istituzionale: una definizione generale

Clifford A. Lynch:

E' "un luogo" gestito da un soggetto dove memorizzare, consentire l'accesso e conservare oggetti digitali di diversa natura e complessità insieme ai metadati necessari per descrivere e gestire gli oggetti medesimi.

Fornisce **servizi** e gestisce **processi** per i membri di una comunità a fini di gestione e diffusione di materiali/documenti digitali prodotti dalla comunità (ad esempio da una università come soggetto produttore istituzionale e dai suoi ricercatori)
Può includere servi di conservazione a medio e a lungo termine

Deposito digitale e conservazione

Il rapporto del Research Library Group sui *Trusted digital repositories (TDR)* costituisce un punto di partenza per un approfondimento in quanto fornisce una struttura di riferimento per sviluppare concetti di base sui depositi e sulla conservazione digitale (non necessariamente a lungo termine)

I requisiti fondamentali previsti dal rapporto riguardano:

- il deposito e le relative politiche, standard e infrastrutture tecnologiche
- il deposito in quanto sistema **accreditato** (ovvero un sistema di sw e hw fondato su regole certe e verificabile)

Deposito digitale per la conservazione vs *deposito digitale istituzionale*

custodia persistente di documenti per i quali il deposito assume responsabilità in termini di integrità e autenticità e deve rispettare requisiti di accreditamento/certificazione e verifica

versus

gestione corrente finalizzata all'uso e alla distribuzione di risorse digitali

Definizione di responsabilità: affidabilità e monitoraggio

E' implicito il principio di assunzione di responsabilità per la conservazione e tutte le attività correlate inclusa la definizione di metodologie di valutazione e monitoraggio
La prima sfida è quella di trasformare i servizi e i depositi tradizionali in *servizi e depositi digitali affidabili (quindi valutabili) in grado di assicurare tenuta stabile e conservazione nel tempo delle fonti digitali acquisite*

Il nodo cruciale della fiducia

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

Il ruolo positivo dei processi di accreditamento

- La certificazione accresce il grado di fiducia dei portatori di interesse (produttori di documenti e archivi e utenti) sulla possibilità di ottenere qualità nei processi conservativi anche attraverso il ricorso a standard riconosciuti.

Storia e stato dell'arte dei processi di auditing - 1

- L'affidabilità e la conseguente possibile certificazione non possono essere il risultato semplice di una auto-dichiarazione.
- In questi anni, a livello internazionale, a partire dall'AWIICS workshop del 1999, gli investimenti pubblici (nazionali ed europei) si sono concentrati sulla creazione di una vera e propria "Digital Preservation Community".
- L'approvazione ufficiale dell'OAIS reference model nel 2002 e l'avvio del processo di standardizzazione ISO costituiscono uno snodo cruciale.

Storia e stato dell'arte dei processi di auditing - 2

- Il rapporto del Research Library Group sui *Trusted digital repositories (TDR) nel 2002 (Trusted Digital Repositories: Attributes and Responsibilities)* fornisce una prima struttura di riferimento per sviluppare concetti di base sulla conservazione digitale a medio e a lungo termine
 - I requisiti fondamentali previsti dal rapporto riguardano:
 - il deposito e le relative politiche, standard e infrastrutture tecnologiche
 - il deposito in quanto sistema accreditato (ovvero un sistema di sw e hw fondato su regole certe e verificabile)

Storia e stato dell'arte dei processi di auditing - 3

- “Trust” per il deposito vuol dire quindi.
 - garantire la “qualità” delle informazioni;
 - agire in un contesto regolato, conforme a standard fissati dalla comunità internazionale: sulla qualità dell’informazione (ISO 9000 segg.); sulla sicurezza dell’informazione (ISO 17799:2005); sulla gestione della documentazione istituzionale (ISO 15489:2001); sul modello di sistema informativo aperto (ISO 1427 OAIS).

Storia e stato dell'arte dei processi di auditing - 4

- Il documento non è **prescrittivo**:
 - insiste sulla rilevanza degli aspetti organizzativi e gestionali
 - indica l'opportunità di sviluppare progetti di certificazione e
 - Predisporre una prima articolazione di criteri o requisiti per la valutazione e verifica dei depositi digitali.

Storia e stato dell'arte dei processi di auditing - 5

- Dal 2003 una Task Force RLG-NARA, con la partecipazione di esperti di altri paesi tra cui UK, Francia e Olanda e di domini professionali diversi (archivi, biblioteche, centri di ricerca, e-government, comunità professionali) ha affrontato il problema della certificazione dei depositi digitali con l'obiettivo pragmatico di stabilire criteri (o requisiti) per individuare e misurare depositi in grado di conservare archivi digitali.
- Il processo è *bottom-up*: analisi dei depositi esistenti e delle loro funzioni e attività, confronto con i portatori di interesse
- Il primo risultato è stato l'*Audit checklist per la certificazione di Trusted Digital Repository* pubblicato in bozza nel 2005 e approvato nella versione attuale nel 2007 (*Trusted Repository Audit Checklist*)

Il percorso della certificazione secondo TRAC - 1

- Il rapporto finale del 2007 (*Version 1.0, Trustworthy Repositories Audit & Certification: Criteria and Checklist*) ritiene che
 - l'auto-valutazione sia il primo passo essenziale nello sviluppo di un programma di certificazione,
 - gli enti e le istituzioni di conservazione siano in grado di misurare priorità e obiettivi pre-definiti sulla base della metrica individuata dalla *checklist*

Il percorso della certificazione secondo TRAC - 2

Il rapporto è organizzato in quattro sezioni per ciascuna tipologia di requisiti:

- A: Organizzazione (governance, staffing, policies and procedures, financial sustainability and contracts and other obligations)
- B: Funzioni, processi e procedure (the whole range of repository preservation responsibilities including ingest (accessioning), archival storage, description, metadata, access, and preservation strategies)
- C: Comunità di riferimento e uso dell'informazione (the ability of the repository to meet the needs of records creators and users)
- D: Tecnologie e infrastruttura tecnologica = ISO 27001:2005 Certificazione della sicurezza Information Security Management System (security, software and hardware, and similar issues that enable digital preservation)

Il percorso della certificazione secondo TRAC - 3

- I commenti e le critiche hanno sottolineato:
 - la mancanza di indicazioni per la concreta implementazione
 - la genericità dell'approccio
- Dal punto di vista degli archivi si è rilevato che il rapporto costringe la disciplina a una attenta ri-considerazione di alcuni nodi metodologici e di molte pratiche (Bruce Ambacher):
 - la definizione della missione in relazione alla specificità dei fondi conservati
 - l'analisi degli utenti e delle modalità possibili e opportune di consultazione
 - la specifica natura (e le conseguenze operative in termini di diritti e responsabilità) di nuove forme di possesso, tenuta, proprietà degli archivi
 - il grado di impegno e di intervento richiesti per assicurare la persistenza
 - la valutazione dei rischi di perdita e la ridefinizione dei concetti di integrità e autenticità (incluse le modalità concrete per garantirne la verificabilità)
 - la capacità di allargare il quadro e la natura delle competenze tecniche (ad esempio in materia di standard)

Il percorso della certificazione secondo TRAC - 4

- E' risultato evidente che
 - la certificazione è un processo continuo che non si conclude mai,
 - la fase di acquisizione è cruciale e richiede informazioni sufficienti e qualificate (ulteriormente affrontate dalle raccomandazioni ICA-ISO (*Record Exchange Standard BRS - Business Requirements Specification*))
 - il rapporto costituisce solo il primo passo per sviluppare un vero e proprio standard internazionale di certificazione (MOIMS-RAC),
 - è necessario documentare/dare prova della conformità ai requisiti (cfr i criteri di "evidence" sviluppati per ciascun aspetto dal Digital Curation Centre),
 - è opportuno procedere per livelli di qualità: a) auto-valutazione; b) auditing esterno (eventuale),
 - il termine di "comunità di riferimento" è critico: è estraneo alla tradizione delle istituzioni archivistiche nazionali (la cui tipologia di utenti è identificabile con chiunque sia interessato agli archivi conservati); rischia di essere restrittivo in altri contesti.

Iniziative in corso

- 2008: **DRAMBORA**, strumento di auto-valutazione e gestione del rischio per un TDR, a cura di Digital Curation Centre e Digital Preservation Europe, <http://www.digitalpreservationeurope.eu/announcements/drambora/>
- 2008: **NESTOR**, Catalogo dei criteri validi per la certificazione dei depositi digitali, *Kriterienkatalog vertrauenswürdige digitale Langzeitarchive, Version 2*, November 2008, <http://www.langzeitarchivierung.de>
- 2008-2009: **MOIMS-RAC**, definizione delle attività e strumenti di valutazione per la certificazione (gruppo di lavoro costituito in ambito scientifico) <http://wiki.digitalrepositoryauditandcertification.org/pub/Main/moims-rac@mailman.ccsds.org>

Requisiti essenziali per gli archivi digitali - 1

Un punto d'intesa è stato raggiunto a livello europeo nella definizione dei requisiti di base per la costituzione di depositi digitali certificati per la conservazione a lungo termine:

1. Il deposito digitale assume la responsabilità della continuità della tenuta e della conservazione nel tempo degli oggetti digitali per la comunità di riferimento.
2. Dà prove della sua stabilità organizzativa (anche sotto il profilo finanziario, del personale impiegato, delle strutture e dei processi adottati) per far fronte agli impegni assunti.
3. Mantiene e assicura nel tempo la integrità, autenticità e fruibilità degli oggetti digitali.
4. Acquisisce e dispone dei diritti contrattuali per assolvere alle sue responsabilità (in un quadro giuridico definito).

Requisiti essenziali per gli archivi digitali - 2

5. Documenta gli interventi applicati agli oggetti digitali durante il loro ciclo di gestione, in tutti i contesti (acquisizione, gestione e conservazione nel deposito, a fini di fruizione) utilizzando metadati appropriati.
6. Dispone di policy e linee d'azione efficaci ed efficienti.
7. Garantisce la fruizione dei documenti digitali.
8. Acquisisce (in fase di versamento) gli oggetti/documenti digitali sulla base di criteri definiti e conformi ai propri obblighi e alle proprie capacità.
9. Definisce e rispetta strategie di pianificazione e di realizzazione degli interventi per la conservazione a lungo termine degli oggetti digitali.
10. Possiede una infrastruttura tecnica adeguata a garantire la continuità della tenuta e la sicurezza degli oggetti digitali.

L'insieme dei dieci *criteri* rappresenta una base di intesa per nuovi e ulteriori accordi di collaborazione e cooperazione.

I percorsi operativi e i nuovi progetti: il gruppo di lavoro MOIMS-RAC

- Il gruppo di lavoro ha l'obiettivo di definire uno standard ISO di certificazione coerente con le indicazioni previste dal modello concettuale OAIS e finalizzato a fornire elementi concreti di conformità con OAIS.
- Ha elaborato regole e strumenti per lo svolgimento di una corretta attività di auditing nell'ambito di infrastrutture di certificazione: *Metrics for Digital Repository Audit and Certification*, in corso di approvazione come *recommended standard* per le istituzioni che fanno capo al CCSDS

Auditing: il punto di partenza di MOIMS-RAC

- 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) e consentano di misurarla sulla base di parametri approvati con uno standard ISO (RAC)

Auditing: la struttura di MOIMS-RAC

Il documento è destinato agli auditor di depositi digitali e a coloro che operano in questo ambito cui intende fornire elementi per **misurare la credibilità** del deposito oggetto di verifica o per **costruirlo secondo parametri** normalizzati di qualità e documentati

E' organizzato in due sezioni informative, una di tipo normativo e alcuni allegati. I requisiti sono discussi ed esemplificati nel medesimo ordine utilizzato nel rapporto sull'Audit Checklist

- A. Infrastruttura organizzativa
- B. Gestione di oggetto digitale
- C. Gestione del rischio per la infrastruttura tecnica e la sicurezza

Per ogni requisito la metrica di valutazione è accompagnata da esempi concreti dei modi in cui la conformità possa essere documentata

Auditing: RAC-MOIMS, elementi di valutazione - 1

- un **mandato istituzionale** che includa la gestione a lungo termine delle risorse digitali e preveda in forme certe la continuità della funzione anche in caso di trasformazioni istituzionali
- **personale competente** ed esperto per tutti i compiti affidati (sia di natura tecnologica che con riferimento agli aspetti archivistici e gestionali), adeguato in numero
- strumenti e meccanismi per garantire l'**aggiornamento professionale** del personale
- un insieme adeguato e coerente di **procedure e linee d'azione** che includa meccanismi di aggiornamento riferiti sia all'evoluzione tecnologica che ai problemi organizzativi
- meccanismi di **verifica e monitoraggio interni** finalizzati a sostenere la continuità della funzione conservativa in modalità adeguate alle esigenze dei soggetti produttori e degli utenti
- **valutazione esterna periodica** sulla funzionalità del deposito

Auditing: RAC-MOIMS, elementi di valutazione - 2

- **documentazione dei cambiamenti intervenuti nel deposito** (con riferimento alle procedure, al software, all'hardware e in relazione alle strategie di conservazione adottate)
- strumenti per la **verifica dell'integrità** delle risorse digitali conservate
- **pianificazione delle attività** al fine di garantire la sostenibilità del deposito
- **accordi di deposito** con eventuali altri depositi federati
- funzioni, attività e procedure per l'**acquisizione delle risorse digitali** che includano:
 - **gli elementi e gli attributi** per ciascuna tipologia di oggetti (identificati secondo standard internazionali)
 - la definizione mediante **accordi scritti con i soggetti che depositano** delle attività, delle procedure e delle modalità di acquisizione, gestione, accesso
 - **linee guida** per il versamento
 - requisiti per un adeguato **controllo fisico** degli oggetti
 - verifica della **completezza e accuratezza degli elementi informativi** disponibili (secondo i parametri indicati dal modello OAIS)
 - processi per la **continuità del servizio**.

Auditing: RAC-MOIMS, elementi di valutazione - 3

- **documentazione delle responsabilità** assunte per la conservazione
- utilizzo di sistemi in grado di garantire la **identificazione univoca e il legame** con le informazioni di rappresentazione
- utilizzo di meccanismi indipendenti di **verifica dell'integrità** dei contenuti del deposito
- **esistenza e documentazione delle strategie di pianificazione** della funzione conservativa e di migrazione/gestione che includono i riferimenti ai **registri dei formati** secondo gli standard internazionali e la documentazione in grado di sostenere il successo dei processi di conservazione
- **gestione dei metadati descrittivi e gestionali/amministrativi**
- **politiche per l'accesso** in grado di garantire risposte certe e documentate
- sistemi di **tracciamento dei file di log** in grado di testimoniare le criticità degli accessi

Auditing: RAC-MOIMS, elementi di valutazione - 4

- strumenti in grado di garantire l'**esibizione di documenti autentici**
- una **infrastruttura tecnologica** adeguata in termini di sistemi operativi, funzioni di backup, ridondanza e collocazione sicura delle copie, garanzie per il loro allineamento, meccanismi di verifica delle perdite e delle manipolazioni, processi di migrazione, di documentazione dei processi di gestione del ciclo di vita e del loro impatto,
- **sistemi di sicurezza** di alta qualità che includono:
 - l'analisi sistematica dell'ambiente, dei dati, del personale, degli impianti e delle attrezzature, delle esigenze;
 - meccanismi di implementazione;
 - un sistema chiaro di ruoli, responsabilità e autorizzazioni;
 - piani scritti di *disaster recovery* regolarmente verificati;

Indicazioni di approfondimento

- **B. Ambacher**, *Government Archives and the Digital Repository Audit Checklist*, in "JoDI", vol 8, no. 2, 2007, <http://journals.tdl.org/jodi/article/view/190/171>
- **ICA-ISO**, *Record Exchange Standard BRS (Business Requirements Specification)* a cura di un gruppo di esperti del CEN/ISSS/eBES/EG13 sostenuti dall'ICA e dall'ISO, <http://www.ica.org/en/node/38983>
- **Interpares**, www.interpares.org
- **Research Libraries Group**, *Trusted Digital Repositories: Attributes and Responsibilities*, May 2002.
- **RLG-NARA Digital Repository Certification Task Force**, *Trustworthy Repositories Audit & Certification (TRAC): Criteria and Checklist*, 2007
- **M. Day**, *Toward Distributed Infrastructures for Digital Preservation: The Roles of Collaboration and Trust*, Intl. J. Digital Curation 1(3), 2008.
- **R. Moore**, *Towards a Theory of Digital Preservation*, Intl. J. Digital Curation 3(1), 63-75, 2008
- **DCC, DPE, NESTOR, CRL, CORE** Requirements for Digital Archives, Chicago, 2007, <http://www.crl.edu/content/DigArc/metrics2/corerequirements.htm>
- **University of Michigan Tutorial sulla conservazione digitale** <http://www.icpsr.umich.edu/dpm/dpm-ita/contents.html>
- **D. Tarrant, S. Hitchcock**, *Preservation as a Process of a Repository*. In: *The Sun Preservation and Archiving Special Interest Group (PASIG) Fall Meeting*, 18 - 21 November 2008, Baltimore, <http://eprints.ecs.soton.ac.uk/12026>