

CONSERVAZIONE DEL DIGITALE: DEFINIZIONI, STRATEGIE, APPROCCI

MICHELE SANTORO

1. Alcune definizioni

Per accostarsi nel modo più idoneo al discorso sulla conservazione del digitale, è opportuno partire da alcune definizioni desunte da un recente rapporto sulla gestione dei documenti elettronici¹; vediamo di seguito.

Materiali digitali: termine ampio che include

- a) **surrogati digitali**, creati come risultato della conversione in forma digitale di materiali analogici (**digitalizzazione**), e
- b) materiali “**nati digitali**”, ossia materiali che non hanno un equivalente analogico.

Accesso: utilizzabilità permanente dei materiali digitali, che devono mantenere tutti i requisiti di autenticità, accuratezza e funzionalità essenziali agli scopi per cui sono stati creati e/o acquisiti.

Digitalizzazione: processo di creazione di file digitali attraverso la scansione o altre forme di conversione di materiali analogici; il prodotto della digitalizzazione è un **surrogato digitale**.

Conservazione del digitale (espressione con cui traduciamo il termine inglese *digital preservation*): insieme delle attività realizzate per assicurare l’accesso continuo ai materiali digitali (sia **surrogati** che “**nati digitali**”) per il tempo considerato necessario; essa può essere:

- **conservazione di lungo termine:** accesso continuo ai materiali digitali, o almeno alle informazioni in essi contenute, per un tempo indefinito;

Michele Santoro, Biblioteca del Dipartimento di Scienze Economiche, Università degli Studi di Bologna, e-mail: santoro@spbo.unibo.it. I paragrafi 2 e 5 di questo testo sono ripresi, con leggere modifiche, dal nostro *Dall’analogico al digitale, la conservazione dei supporti non cartacei*, “Biblioteche oggi”, 19 (2001) 2, p. 88-100. Le traduzioni da testi stranieri sono nostre.

¹ MAGGIE JONES, NEIL BEAGRIE, *Preservation management of digital materials*, Workbook, Pre-publication draft, October 2000.

- **conservazione di medio termine:** accesso continuo ai materiali digitali, al di là dei cambiamenti nella tecnologia, per un periodo di tempo definito;
- **conservazione di breve termine:** accesso continuo ai materiali digitali, sia per un periodo di tempo definito ma che non si estende oltre un prevedibile futuro, sia finché tale accesso risulti possibile a causa dei cambiamenti nella tecnologia.

Dunque, se è possibile rendere in modo abbastanza agevole l'espressione **digital preservation**, lo stesso non avviene per un'altra locuzione in cui ci s'imbatte esplorando questi territori, e cioè quella di **preservation reformatting**.

Nel riprendere la definizione del rapporto sopra citato, vediamo che per *reformattare* s'intende "copiare il contenuto informativo da un supporto memorizzato a un diverso supporto memorizzato (**media reformatting**). Se poi si effettua una ricerca in rete, con **preservation reformatting** si ottengono numerose risposte, relative per lo più a dipartimenti o sezioni di grandi biblioteche che si occupano del trasferimento delle informazioni da un supporto ad un altro: dal sito della Columbia University ad esempio apprendiamo che il Dipartimento di Preservation Reformatting "realizza tutte le attività associate con la riformattazione di materiali stampati su carta e che sono diventati deteriorati"²; con maggior chiarezza, l'analoga divisione della Library of Congress dichiara di essere "responsabile del programma della Biblioteca volto a fornire accesso ai materiali a rischio copiandoli su nuovi formati, quali microfilm, copie facsimilari su carta o riproduzioni digitali"³.

Dunque la distinzione fra **digital preservation** e **preservation reformatting** è sottile, i significati non ancora sedimentati e le sovrapposizioni possibili. Così, senza pretendere di arrivare a denotazioni definitive, si può forse dire che quando si parla di **preservation reformatting** si è ancora nella "filiera" del documento analogico, con particolare riferimento a tecniche di conservazione che prevedono un trasferimento di supporto, e che dunque si aggiungono a quelle tradizionali, basate su criteri di conservazione "ex ante" (quali ad esempio il condizionamento ambientale) o "ex post" (come ad esempio il restauro). In ambiente digitale tuttavia l'espressione **preservation reformatting** è a volte usata per indicare l'uso della digitalizzazione come un vero e proprio criterio di conservazione, in particolare se impiegata in alternativa al microfilm⁴.

² Columbia University Libraries, Preservation Reformatting Department, <<http://www.columbia.edu/cu/lweb/services/preservation/PRD.html>>.

³ Library of Congress, Preservation Reformatting Division, <<http://www.loc.gov/preserv/reformat.html>>.

⁴ Library of Congress, Preservation Reformatting Division, *Principles and specifications for Preservation Digital Reformatting*, <<http://www.loc.gov/preserv/prd/presdig/presprinciple.html>>.

Quando invece si parla di **digital preservation** - giusta la definizione precedente, ossia l'insieme delle attività realizzate per assicurare l'accesso continuo ai materiali digitali - il focus è tutto sulla capacità di accedere in maniera permanente ai documenti digitali, siano essi surrogati che nati digitali. Si può allora dire che nel primo caso il fulcro sia il documento, nel secondo l'accesso⁵.

2. Strategie di conservazione

La storia della *digital preservation* è recente e assai intensa. Come punto di partenza si può assumere l'ormai celebre rapporto della Task Force on Archiving of Digital Information⁶ che, a partire dal 1995, ha messo in luce non solo gli aspetti di fragilità e provvisorietà delle informazioni digitali, ma i problemi connessi all'obsolescenza tecnologica dei sistemi informatici - hardware e software - senza dei quali è impossibile l'utilizzo di queste informazioni.

Riportando drammatici esempi di scomparsa di interi archivi digitali, il rapporto ha richiamato l'attenzione della comunità internazionale sui gravi rischi a cui può andare incontro il patrimonio documentario in formato elettronico - sia prodotto in forma digitale, sia convertito dal formato analogico a quello numerico - ed ha indicato al tempo stesso le soluzioni che appaiono più ragionevolmente praticabili.

Il rapporto della Task Force rimane a tutt'oggi lo studio più completo sull'argomento: per quanto alcuni aspetti possano apparire datati, l'analisi complessiva e le scelte di fondo non sono state finora messe seriamente in discussione; il suo impatto sulla comunità internazionale è stato fortissimo, e ad esso sono seguiti una quantità di studi e iniziative volte a individuare le soluzioni più idonee per far fronte a un problema d'importanza cruciale per la società contemporanea.

Le strategie di conservazione individuate dal rapporto e dagli studi che si sono succeduti riguardano sia i surrogati digitali, cioè i prodotti della conversione di

⁵ Difatti, come scrivono Michèle V. Cloonan e Shelby Sanett, "la conservazione delle informazioni digitali è volta non tanto a proteggere gli oggetti fisici quanto a specificare la creazione e la manutenzione di intangibili file elettronici per i quali l'integrità intellettuale è la caratteristica primaria. La conservazione dunque rappresenta qualcosa di più che salvare supporti quali dischi ottici o i nastri magnetici; è il sistema dell'accesso in sé che deve essere conservato" (MICHÈLE V. CLOONAN, SHELBY SANETT, *Comparing preservation strategies and practices for electronic records*, <<http://www.rlg.org/events/pres-2000/cloonan.html>>).

⁶ *Preserving digital information. Report of the Task Force on Archiving of Digital Information*, commissioned by The Commission of Preservation and Access and the Research Library Group, <<http://www.rlg.org/ArchTF/tfadi.index.html>>. Cfr. inoltre NEIL BEAGRIE, DANIEL GREENSTEIN, *A strategic policy framework for creating and preserving digital collections*. Arts and Humanities Data Service, London, King's College, 1998, <<http://ahds.ac.uk/manage/framework.htm>>.

originali analogici, sia i materiali “nati digitali”, cioè quelli per cui non esiste alcun corrispondente analogico; ne riportiamo di seguito le principali⁷.

In primo luogo, al fine di preservare le fragili risorse numeriche, si propone a volte la realizzazione di **hard copies**, cosa che equivale a un ritorno all’analogico, se è vero che tale criterio prevede la stampa su carta o la realizzazione di un microfilm a partire da un originale in formato elettronico. E’ noto infatti che la carta ha tempi di obsolescenza assai lunghi, e si stima che il microfilm possa durare centinaia di anni, ma ciò nonostante la proposta appare paradossale: sia perché siamo in una fase in cui si va verso la digitalizzazione di gran parte del patrimonio analogico esistente su carta o su microfilm e non viceversa; sia perché è un’idea praticamente irrealizzabile, a causa dell’immensa quantità di documenti esistenti in formato digitale; inoltre, essa è tecnicamente svantaggiosa, in quanto sulla *hard copy* non si possono replicare le caratteristiche di flessibilità e maneggevolezza né le capacità ipertestuali e multimediali proprie dei documenti elettronici.⁸

Un’altra proposta che talvolta affiora è quella di mantenere “**come in un museo**” una quantità di hardware e software obsoleti al fine di poterli utilizzare per il corretto utilizzo dei documenti digitali: un’idea, con ogni evidenza, che richiede costi elevati e che non dà garanzie certe di funzionalità dei prodotti conservati. Si parla poi di “**fiducia negli standard**”, soluzione che prevede il trasferimento dei documenti su un nuovo software in grado di condividere con il vecchio gli stessi paradigmi: si tratta di un approccio che da un lato presuppone che tali standard possano mantenersi inalterati nel tempo, e dall’altro richiede costi notevoli per le necessità della loro manutenzione.

A fronte di queste proposte, decisamente più perseguibile appare invece il **refreshing**, ossia la tecnica volta a “mantenere i bit in buona salute”, e che consiste nel copiare, a distanze di tempo piuttosto brevi, le informazioni su nuovi supporti (ad esempio da un floppy disk da 5 pollici e un quarto a uno da 3 pollici e mezzo, oppure da un CD-ROM a un DVD).

Molti consensi riscuote infine la **migrazione**, una strategia che, al pari del *refreshing*, prevede anch’essa il trasferimento sistematico dei documenti su nuovi hardware e

⁷ Cfr. in particolare *Preserving digital information*; JEFF ROTHENBERG, *Avoiding technological quicksand: finding a viable technical foundation for digital preservation*, A report to the Council on Library and Information Resources, Commission on Preservation and Access, Washington, CLIR, 1999, <<http://www.clir.org/pubs/reports/rothenberg/contents.html>>.

⁸ Molto interessante per contro appare il cosiddetto “approccio ibrido”, che propone “di combinare i vantaggi della scansione digitale per l’accesso con quelli della microfilmatura, sia creando un microfilm con i requisiti necessari alla conservazione a lungo termine, e poi ricavando da questo una copia digitale per l’accesso, sia ricavando da una scannerizzazione di alta qualità un COM (Computer Output Microfilm) a scopo di conservazione (GLORIA CIROCCHI, *Conservazione di risorse digitali: quali sfide?*, “Bollettino AIB”, 39 (1999), 3, p. 289-302).

software prima che i precedenti diventino obsoleti ma che, a differenza di quest'ultimo, richiede una ricodifica periodica dei dati affinché possano essere leggibili dalle nuove generazioni di hardware e software.

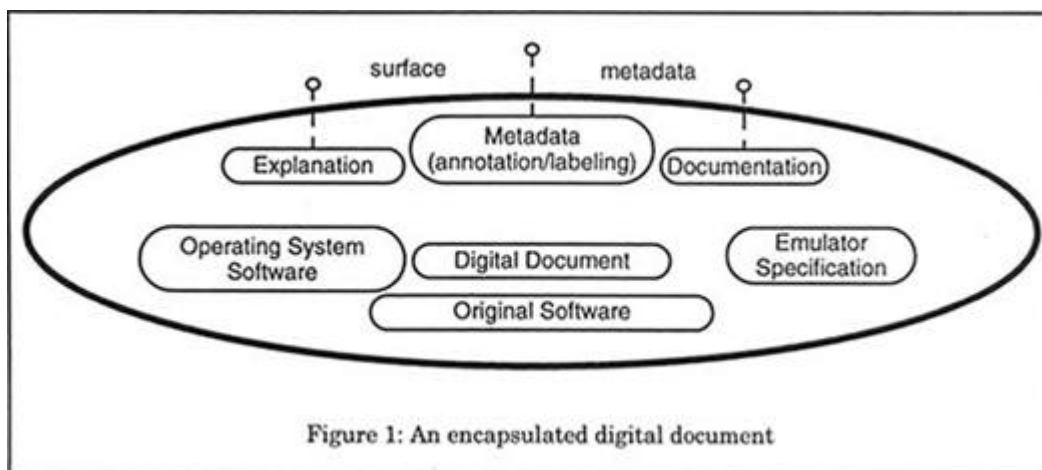
Lo scopo della migrazione infatti è di conservare il contenuto intellettuale degli oggetti digitali e mantenere la capacità di recuperarli, mostrarli o usarli a fronte dei continui mutamenti nella tecnologia. La migrazione, con ogni evidenza, include il *refreshing*, ma differisce da questo in quanto non sempre è possibile ottenere un'esatta copia digitale, o replicare le caratteristiche originarie di database, hardware e software garantendone allo stesso tempo una piena compatibilità con le nuove generazioni di prodotti tecnologici. Difatti, anche per informazioni codificate in forme standard e ormai consolidate (come ad esempio database bibliografici in formato USMARC, o database aziendali di tipo relazionale del tipo SQL), effettuare migrazioni di informazioni verso nuovi standard o programmi applicativi può essere un processo lungo, costoso e molto più complesso di una semplice migrazione.

Una strategia del tutto alternativa alle precedenti è invece l'**emulazione**, che da Jeff Rothenberg, il suo ideatore, è stata definita come “un approccio che permette l'emulazione di sistemi obsoleti su sistemi futuri e non ancora noti, in modo che un software originale di un documento digitale possa essere usato in futuro malgrado sia ormai obsoleto”.⁹

In altre parole, anziché far migrare periodicamente i dati da un supporto a un altro, Rothenberg ritiene di poter utilizzare il software di un documento originale anche se questo, così come l'hardware, sia diventato ormai obsoleto: a tale scopo egli propone di “incapsulare”, insieme al documento digitale, non solo il sistema operativo e il software utilizzati per la sua creazione, ma anche un emulatore della nativa piattaforma hardware, in grado per l'appunto di emulare l'hardware originale su qualsiasi computer futuro; a ciò si aggiungono i metadati, ossia le informazioni sui dati inseriti nella capsula, ed altre informazioni accessorie. Il seguente schema¹⁰ rende esplicita la sua idea:

⁹ JEFF ROTHENBERG, *Avoiding technological quicksand*, cit.

¹⁰ Incluso nel medesimo testo.



La proposta di Rothenberg ha incontrato notevoli consensi ma ha suscitato anche vivaci critiche, in particolare ad opera di David Bearman,¹¹ il quale da un lato ha fatto notare come hardware e software siano di norma “proprietary”, quindi difficilmente “incapsulabili” a causa dei diritti detenuti dai relativi creatori; dall’altro ha messo in luce le difficoltà di lettura dei metadati su piattaforme future e non ancora conosciute, rimarcando l’assenza di istruzioni su come incapsulare tali metadati, che invece avrebbero bisogno di una “architettura” ben definita in vista di una loro corretta interpretazione: lo stesso Bearman infatti riconosce che l’incapsulamento dei metadati può costituire una strategia di successo, e per questo chiede una maggior precisione nei criteri per una loro identificazione, creazione e cattura al momento della costituzione del documento.

3. Migrazione ed emulazione

La validità di queste strategie rappresenta materia di dibattito negli ambienti della *digital preservation*: difatti, come sottolinea il rapporto della Task Force, “nessuna singola strategia si può applicare indifferentemente a tutti supporti dell’informazione digitale, e nessuno degli attuali metodi di conservazione è del tutto soddisfacente”. E’ ormai certo tuttavia che migrazione ed emulazione siano le strategie più idonee alla conservazione dei record digitali¹²; è dunque opportuno esaminarle parallelamente al fine di poterne individuare vantaggi e limiti¹³.

¹¹ DAVID BEARMAN, *Reality and chimeras in the preservation of electronic records*. “D-Lib Magazine”, 5 (1999), 5, <<http://www.dlib.org/dlib/april99/bearman/04/bearman.html>>.

¹² Sulla valutazione delle diverse strategie cfr. fra l’altro TONY HENDLEY, *Comparison of methods and costs of digital preservation*, British Library Research and Innovation Report 106, <<http://www.ukoln.ac.uk/services/elib/papers/tavistock/hendley/hendley.html>>; MICHÈLE V. CLOONAN, SHELBY SANETT, cit.; ANNE R. KENNEY, OYA Y. REIGER, *Moving theory into practice: digital imaging for libraries and archives*. Mountain View, Research Libraries Group, 2000; ANDREW K. PACE, *Digital preservation: everything new is old again*, “Computers in Libraries”, 20 (2000) 2, <<http://www.infotoday.com/cilmag/feb00/pace.htm>>. Per ciò che riguarda in particolare l’emulazione cfr. STEWART GRANGER, *Emulation as a digital preservation strategy*, “D-Lib Magazine”, 6 (2000), 10, <<http://www.dlib.org/dlib/october00/granger/10granger.html>>.

3.1. Migrazione: una volta definita nella sua funzione di trasferimento e ricodifica periodica dei dati, è importante sottolineare che esistono diverse forme di migrazione, che possono variare a seconda se i sistemi di conversione sono semplici e per così dire diretti, oppure se avvengono in seguito a forti cambiamenti nel paradigma tecnologico; in quest'ultimo caso, è ovvio, si può andare incontro a problemi piuttosto rilevanti. E' opportuno dunque esaminare rapidamente, oltre ai vantaggi ed agli svantaggi di questa strategia, anche i requisiti richiesti per la sua realizzazione e le eventuali strategie collegate.

Vantaggi:

- le procedure per migrazioni semplici sono ben definite;
- è la strategia preferita per la conversione della maggior parte degli archivi digitali;
- può diventare più semplice man mano che si sviluppa la tecnologia e si riduce l'ampiezza delle piattaforme;
- sono presenti strumenti per la valutazione dei rischi, in grado di assistere nelle decisioni al riguardo.

Svantaggi:

- costi: è necessario uno specifico programma per migrazioni complesse;
- può richiedere tempi assai lunghi e risultare un procedimento complesso;
- sono probabili perdite di funzionalità, di immagine e di sensibilità rispetto all'originale;
- può compromettere l'integrità dell'originale se avviene senza rigorose procedure di controllo di qualità che siano in grado di assicurarne l'autenticità;
- le risorse digitali più complesse possono essere trasferite con perdite significative di funzionalità;
- è necessario intervenire periodicamente nell'intero ciclo di vita delle risorse.

Requisiti:

- politiche e linee guida scritte, incluse le politiche di selezione dei materiali da sottoporre a migrazione;
- procedure di controllo di qualità;
- rigorosa documentazione delle procedure di migrazione;
- conservazione dei metadati e della documentazione;
- far migrare i dati ogni volta che vi è l'upgrade di un software o l'installazione di un nuovo software;

¹³ Deduciamo le osservazioni che seguono dal già citato studio di Jones e Beagrie.

- assicurarsi che la migrazione provochi perdite minime o nulle nel contenuto o nel contesto;
- impiegare rigorose procedure di controllo di qualità che includano un test del programma di migrazione insieme a un campione di record o di bit/byte, o un confronto fra i dati originari e quelli migrati;
- mantenere copie della risorsa digitale nel suo formato originale ogni volta che una informazione o presentazione della risorsa può andare perduta e/o modificata durante la migrazione.

Strategie collegate:

- archiviazione e mantenimento;
- compatibilità retrospettiva;
- identificazione permanente;
- procedure di validazione;
- conversione a formati standard.

3.2. Emulazione: intesa come un mezzo per superare l'obsolescenza tecnologica di hardware e software attraverso lo sviluppo di tecniche volte a imitare i sistemi obsoleti su future generazioni di computer. Si tratta di una strategia sulla quale si concentra un interesse crescente, se è vero che sono in corso diversi esperimenti¹⁴ ed anche interi progetti rivolti ad esplorarne le possibilità, fra cui il più rilevante è senz'altro il progetto CAMiLEON¹⁵, promosso dalle università del Michigan e di Leeds. Rinviamo a tali lavori per ulteriori approfondimenti, in questa sede è invece opportuno enucleare vantaggi e limiti, nonché requisiti e strategie collegate.

Vantaggi:

- ricrea la funzionalità, l'immagine e la sensibilità dell'originale;
- evita i costi ripetuti associati alla migrazione;
- può offrire la migliore prospettiva per le risorse digitali più complesse.

Svantaggi:

- è ancora in fase sperimentale e richiede ulteriori test pratici;
- può essere capace di emulare soltanto parte della funzionalità, dell'immagine e della sensibilità dell'originale;

¹⁴ JEFF ROTHENBERG, *An experiment in using emulation to preserve digital publications*. Den Haag, Koninklijke Bibliotheek, 2000, <www.konbib.nl/coop/nedlib/results/emulationpreservationreport.pdf>.

¹⁵ CAMiLEON, Creative Archiving at Michigan and Leeds: Emulating the Old on the New, <<http://www.si.umich.edu/CAMiLEON/>>.

- è probabile che sia piuttosto costosa se non si realizzano precise economie di scala: c'è bisogno ad esempio di costruire nuovi emulatori per i principali spostamenti di paradigma dei computer, ed è possibile che tali costi spesso possano essere maggiori dei costi della migrazione;
- c'è bisogno di affrontare gli aspetti relativi al copyright dei software, aspetti che possono risultare estremamente complessi;
- è necessaria una documentazione rigorosa dei requisiti hardware e software; questi aspetti raramente sono stati documentati a un sufficiente livello di dettaglio e richiedono sforzi e risorse comuni.

Requisiti:

- appropriate procedure di memorizzazione e di manutenzione;
- politiche e linee guida scritte;
- metadati per la conservazione;
- documentazione dettagliata su specifiche hardware e software.

Strategie collegate:

- memorizzazione e manutenzione;
- incapsulamento.

4. Archeologia digitale

Sviluppata nella sua forma più ampia e convincente in un recente studio di Seamus Ross ed Ann Gow¹⁶, l'archeologia digitale non è una vera e propria strategia ma un'insieme di criteri e di metodi attraverso cui è possibile recuperare informazioni digitali danneggiate o perdute per le ragioni più diverse, ad esempio a causa dell'obsolescenza tecnologica di hardware e software, oppure per il degrado fisico dei supporti.

Essa riguarda di conseguenza sia gli oggetti digitali propriamente detti (surrogati o "nati digitali"), sia i loro supporti (dischi, nastri, etc.), sia gli strumenti informatici essenziali per il loro utilizzo (hardware e software). L'archeologia digitale dunque comprende tecniche di recupero dei dati che da un lato attengono alla composizione chimica dei supporti, dall'altro alle strutture di funzionamento di hardware e software, incluse naturalmente le strategie di conservazione analizzate in precedenza.

¹⁶ SEAMUS ROSS, ANN GOW, *Digital archaeology: rescuing neglected and damaged data resources*, A JISC/NPO study within the Electronic Library (eLib) Programme on the preservation of electronic materials, February 1999, <<http://www.ukoln.ac.uk/services/elib/papers/supporting/pdf/p2.pdf>>.

Anche questo sistema ovviamente presenta una serie di vantaggi e di svantaggi, che esaminiamo di seguito.

Vantaggi:

- vi è un numero crescente di specialisti esterni che offrono tale servizio;
- è dimostrato che è tecnicamente possibile recuperare una vasta gamma di informazioni da supporti obsoleti o danneggiati (anche se non necessariamente nella medesima forma esistente in precedenza).

Svantaggi:

- sono metodi molto più costosi nel lungo termine rispetto alle strategie “in buona fede” di conservazione del digitale;
- difficilmente riesce ad essere vantaggioso in un rapporto costi-efficacia, se non per le risorse a più alto valore aggiunto;
- possono andare perduti anche materiali potenzialmente utili ma che tuttavia non giustificano i costi;
- c'è il rischio che non tutti i materiali digitali possano essere recuperati con successo;
- la gestione degli investimenti iniziali può risultare difficile.

Tuttavia, per quanto numerose possano apparire le difficoltà, l'archeologia digitale sembra davvero in grado di fornire contributi notevoli nel ripristino di importanti insiemi di record digitali; nel loro studio Ross e Gow riportano illuminanti esempi di recupero a seguito di disastri naturali, o di dati archiviati in formati sconosciuti o ricavati da supporti fortemente danneggiati. Fra questi, molto interessanti sono le esperienze di restauro dei dati contenuti nei nastri magnetici dello Space Shuttle, sprofondato al largo della Florida a seguito dell'esplosione della navetta spaziale nel 1987; o di ricostituzione dei dati digitali danneggiati nel 1995 dal passaggio dell'uragano Marilyn che nel 1995 devastò le Isole Vergini; o ancora il recupero dei record elettronici della ex Germania dell'Est che, in seguito all'unificazione nazionale, risultavano inaccessibili a causa dell'incompatibilità di hardware e software¹⁷.

5. L'impiego dei metadati

Come si è visto, sia la migrazione sia l'emulazione richiedono, per il loro utilizzo efficace, l'impiego dei metadati¹⁸. Difatti, analogamente a quanto avviene nella

¹⁷ Ibid.

conservazione dei supporti tradizionali, anche per i formati digitali è necessaria la presenza di una serie di informazioni relative alla loro struttura, al contesto in cui sono nati e al loro utilizzo nel tempo, informazioni che possono essere fornite solo da specifici metadati: come ha scritto Seamus Ross, “i metadati devono essere capaci di garantire il trasferimento dei record attraverso diverse generazioni di hardware e software; semplificare la ricostruzione dei processi decisionali; creare percorsi di verifica all’interno del ciclo di vita dei record; realizzare record con caratteristiche di auto-selezione ed auto-valutazione; ed essere in grado di catturare la loro documentazione interna”.¹⁹

Pertanto, continua Ross, i metadati da associare a un record digitale in vista della sua conservazione devono includere tutte le informazioni sulla fonte dei dati e su come, quando e perché tale fonte è stata creata; fornire i criteri per aprire e leggere il record in condizioni corrette; riportare infine la storia del record, ossia tutte le modifiche avvenute dopo la sua creazione e le modalità con cui tale record si è incrociato con altri software e con altri record. Tuttavia, commenta l’autore, spesso non si tiene conto della necessità di definire i metadati in termini tali da riflettere l’ambiente in cui hanno avuto origine i dati primari: in altre parole, i metadati non possono essere elaborati a tavolino, ma hanno bisogno di essere ricavati da un’analisi dei requisiti e dei bisogni funzionali e organizzativi dei singoli sistemi, affinché possano assicurare il corretto utilizzo dei dati e garantire un’adeguata conservazione dei documenti a cui sono correlati²⁰.

¹⁸ Si veda tra l’altro *Preservation metadata for digital objects: a review of the state of the art*, A White paper by the OCLC/RLG Working Group on Preservation Metadata, January 31, 2001, <http://www.oclc.org/digitalpreservation/presmeta_wp.pdf>; MICHAEL DAY, *Extending metadata for digital preservation*, “Ariadne” 9, May 1997, <<http://www.ariadne.ac.uk/issue9/metadata/>>; ID. *Metadata for digital preservation: an update*, “Ariadne”, 22 (1999), <<http://www.ariadne.ac.uk/issue22/metadata/>>; GLORIA CIROCCHI, SIMONA GATTA, LUCIA PANCIERA, ENRICO SETA, *Metadati, informazione di qualità e conservazione delle risorse digitali*, “Bollettino AIB”, 40 (2000), 3, p. 309-329.

¹⁹ SEAMUS ROSS, cit.

²⁰ Ibid.; malgrado il crescente consenso, precisa Ross, questo modello di metadati non va esente da una serie di problemi dovuti alla mancanza di adeguati test in rapporto a determinate categorie di dati; all’utilizzo dei metadati stessi per la rappresentazione di una serie di processi (incluso il flusso delle informazioni); alla stratificazione dei documenti e la necessità di una più ampia copertura della documentazione del sistema; alla necessità di incorporare i metadati nel software; e infine alla capacità da parte di sistemi futuri di interpretare e usare questi metadati.