

Dig *Italia*

Anno V, Numero 2 - **2010**

Rivista del digitale nei beni culturali

ICCU-ROMA

Il paradosso della conservazione digitale: riflessioni sull'autenticità*

Giovanni Michetti

Università degli studi di Roma "La Sapienza"

L'autenticità è uno dei nodi cruciali della conservazione a lungo termine degli oggetti digitali: nonostante gli sforzi e le iniziative di ricerca promosse a livello internazionale, la comunità scientifica non dispone oggi di una teoria compiuta dell'autenticità in ambiente digitale, pur potendo fare riferimento ad un corpus di analisi, riflessioni e materiali ormai consolidato. Il presente articolo si inserisce nel dibattito in materia e, prendendo le mosse dal carattere paradossale della conservazione digitale, in bilico fra le opposte esigenze di trasformazione e di non alterazione degli oggetti, suggerisce degli elementi di riflessione per una migliore definizione del concetto di autenticità. Sulla base delle considerazioni introdotte, viene poi illustrato sommariamente il modello di autenticità elaborato all'interno del progetto europeo CASPAR e sviluppato coerentemente con l'approccio concettuale proposto nella prima parte del saggio.

L'autenticità è sicuramente un nodo cruciale della conservazione a lungo termine degli oggetti digitali: il ruolo dell'autenticità come requisito preliminare e fondamentale dei processi conservativi è stato analizzato all'interno di vari progetti internazionali, alcuni concentrati sulla conservazione a lungo termine di oggetti digitali nell'ambito scientifico, culturale e della pubblica amministrazione (primo fra tutti InterPARES¹), altri dedicati all'identificazione dei criteri e delle responsabilità per lo sviluppo di depositi digitali affidabili (come ad esempio il progetto tedesco NESTOR² o l'iniziativa congiunta del Research Library Group e del National Archives and Records Administration³ negli Stati Uniti).

Uno dei vincoli fondamentali per lo sviluppo di una teoria dell'*autenticità digitale* risiede nell'impossibilità pratica di conservare le risorse digitali nelle loro condizioni originali⁴, senza alcuna alterazione: nella maggior parte dei casi siamo costretti ad un'incessante attività di migrazione degli oggetti che non può essere ridotta ad un semplice spostamento su nuovi supporti, ma che si configura piuttosto come

* Il testo è una rielaborazione dell'intervento presentato al convegno "I luoghi delle memorie e della conoscenza" (Ravenna, Archivio di Stato, 14 novembre 2008).

¹ <http://www.interpares.org>.

² <http://www.langzeitarchivierung.de/eng/index.htm>.

³ <http://www.oclc.org/research/activities/past/rlg/repositorycert.htm>.

⁴ Si è preferito fare riferimento alle *condizioni* piuttosto che alla forma o allo stato, onde esprimere nella maniera più generale possibile l'insieme delle caratteristiche fisiche e logiche o – se si preferisce – dei caratteri estrinseci ed intrinseci di un oggetto.

una vera e propria trasformazione, in maniera tale che possiamo solo *ri-produrre* gli oggetti originali o – meglio ancora – possiamo solo riprodurre la percezione degli oggetti originali. In breve: non possiamo conservare gli *originali*⁵. Sfortunatamente questa circostanza abbastanza ovvia muove in senso contrario ad una fondamentale assunzione di principio, e cioè che la conservazione dell'autenticità implica la conservazione dell'*identità* e dell'*integrità* dell'oggetto digitale. Non è neppure necessario scomodare la dimensione tecnologica del problema o ricorrere a sofisticate considerazioni teoretiche: il dato di esperienza ci suggerisce di ritenere autentico un oggetto che non sia stato modificato o corrotto, *in primis* nelle sue caratteristiche fisiche (integrità), ma non secondariamente anche in relazione alla sua natura, al contesto di appartenenza o al suo profilo logico (identità). Si genera così una sorta di paradosso – ma neppure tanto, a ben pensarci – ove gli oggetti sono costretti a cambiare per rimanere se stessi: l'obsolescenza tecnologica impone il cambiamento, lo spostamento, il fattore dinamico insomma; l'autenticità invece richiede – o meglio, sembra richiedere – la staticità, la *fissità* degli oggetti. In altre parole, una sorta di *crystallizzazione* della fonte sembrerebbe l'espediente più sicuro per garantire ai posteri la genuinità della fonte stessa, intesa non solo come protezione nei confronti delle eventuali corrotte determinate dalla temperie del tempo, ma anche come consolidamento e stabilizzazione delle caratteristiche complessive dell'oggetto stesso. La conservazione riguarda infatti l'insieme delle caratteristiche di un oggetto, e non *sic et simpliciter* l'oggetto stesso: è un lieve slittamento nella formula linguistica che però sottintende un profondo convincimento culturale e metodologico. Nonostante le apparenze, la conservazione – intesa nella sua accezione più completa e complessa – non è mai rivolta alla sola materialità di un oggetto, sia esso documentale o di altro tipo: la conservazione riguarda l'oggetto *tout court*. Ciò ha significato – fino ad oggi – preservarne la materialità, poiché questa si è posta come ineludibile strumento e veicolo del contenuto informativo dell'oggetto stesso. In altri termini, anche in ambiente tradizionale la conservazione della materialità non è il fine, ma piuttosto il mezzo attraverso il quale noi conserviamo il patrimonio di valori (siano essi informativi, ideali, tecnici o latamente culturali) soggiacenti all'oggetto stesso. Questa prospettiva non esclude ovviamente che vi siano ambiti nei quali la materialità *in sé* comunichi valori degni di conservazione: basti pensare alle opere d'arte o – per restare nel nostro dominio – alle fonti documentarie di epoche lontane, ove la fattura stessa del documento è un valore, se non il valore preminente, e il documento porta inciso sulla sua pelle, metaforicamente e concretamente, un significato che

⁵ Ovviamente, come evidenziato dall'uso del corsivo, si fa qui riferimento alla nozione tradizionale di *originale*, giacché l'ambiente digitale impone a nostro avviso un ripensamento di questa categoria concettuale e l'adozione di un mutato paradigma che consenta di considerare originali anche gli oggetti che abbiano subito delle trasformazioni, a patto che queste non modifichino le proprietà significative degli oggetti stessi.

possiamo per lo meno qualificare come tecnico, ma che qualunque studioso della cultura scritta non esisterà a riconoscere come culturale. Ma non si può non riconoscere che l'ambiente digitale sembra proporci una prospettiva nuova: la conservazione del bene culturale non impone la conservazione della materia originale⁶.

A tutto ciò si aggiunga che, a ben guardare, anche in ambiente tradizionale la conservazione non si è mai limitata alla sola materialità, ma si è estesa all'insieme delle conoscenze che fungono da necessario corredo per la corretta interpretazione di una risorsa: senza una conservazione delle idealità soggiacenti e del contesto culturale, antropologico, tecnico, etc. entro cui l'oggetto è immerso, non riusciremmo oggi a recuperare il *patrimonio culturale*, cioè un insieme di idealità tramutate nella materialità, e avremmo un vuoto oggetto incapace di comunicarci i suoi significati. In altre parole, non occorre dimenticare che – in maniera implicita o esplicita – insieme agli oggetti abbiamo sempre conservato un *corpus* di conoscenze che in una certa misura è parte dell'oggetto stesso, poiché partecipa del suo sistema di significati.

Il paradosso fra staticità e dinamismo si ripercuote sulla *stabilità* dell'autenticità in ambiente digitale: questa non può essere riconosciuta come data una volta per tutte, poiché è inficiata ogni qual volta un oggetto digitale è trasferito nello spazio (ad esempio nello scambio fra utenti, sistemi, applicativi) o nel tempo (per semplice permanenza sul supporto di archiviazione o per migrazione, aggiornamento o in generale trasferimento).

L'estrema labilità e l'evidente incertezza dei contorni dell'autenticità impongono

⁶ Beninteso, la storia delle successive migrazioni da un supporto all'altro deve essere conservata, perché ogni supporto – e ogni sistema logico entro il quale l'oggetto è immerso – ha delle specifiche caratteristiche. Il rigore filologico impone di tenere traccia di questo "percorso", non foss'altro perché il supporto stesso ci comunica informazione: una gestione con cd-rom e juke-box è più adeguata per una conservazione cosiddetta *near-line* o *off-line*; l'uso odierno di un floppy disk denota probabilmente un ritardo tecnologico; una chiave USB è riservata in linea di massima a duplicazioni o spostamenti temporanei di dati, non certo alla conservazione permanente; e così via. Senza calcolare che il passaggio da un supporto all'altro, o fra due supporti dello stesso tipo diversamente formattati, potrebbe causare perdita d'informazione. Insomma, l'elemento materiale non è trascurabile, per lo meno non sempre. Il punto è che in ambiente digitale il suo rilievo è in genere assolutamente minimo, e men che mai paragonabile al suo *status* nell'ambiente tradizionale. Molto concretamente: non è affatto improbabile che un *file*, memorizzato in origine su un floppy disk, giunga fino a noi sotto forma di *file* archiviato su disco fisso. Tale circostanza non sembra problematica: la conoscenza del supporto originale serve certo per immaginare le modalità di fruizione della risorsa (e cioè reperimento del disco, inserimento nel drive, tempi di accesso più lenti, maggiore sensibilità al degrado, ecc.); e la conoscenza di un'avvenuta trasformazione deve renderci consapevoli della possibilità di perdita di dati rilevanti. Ma il cuore della fruizione rimarrebbe inalterato: ci ritroveremmo (sempre che sia stato conservato tutto l'apparato tecnologico necessario e che si possa accedere all'oggetto senza difficoltà) davanti a uno schermo e alla rappresentazione dell'oggetto. Per correttezza filologica potremmo disquisire sul fatto che tale rappresentazione è oggi realizzata grazie a schermi ultratecnologici e iperdimensionati, mentre l'utente del passato disponeva ahilui di un ingombrante terminale VT100 con un piccolo video in bianco e nero; ma ci addentreremmo in un livello di analisi più sofisticato, ove è bene che esercitino le proprie competenze altre discipline.

l'adozione di misure a sostegno di una gestione controllata della risorsa: l'autenticità di una risorsa deve essere comprovata e sostenuta da *prove* associate alla risorsa stessa, attraverso una qualsivoglia forma di *documentazione* che tracci e testimoni la storia dei processi di migrazione e di elaborazione, in una parola dei processi di *trasformazione* dell'oggetto nel corso del tempo. In particolare, occorre documentare l'utilizzo di strumenti, tecniche e strategie mirate a garantire l'*identità* e l'*integrità* delle risorse, o almeno in grado di minimizzare i rischi di cambiamento nel corso del tempo.

La conseguenza di tale approccio è che l'autenticità non è mai limitata alla risorsa oggetto di analisi, non è una qualità della risorsa in sé, ma si estende all'intero sistema informativo/documentario e alla sua affidabilità. Ed è per questo motivo che l'autenticità implica un controllo continuativo tanto del contesto di produzione quanto dell'ambiente di trattamento e conservazione della risorsa.

Non possiamo tuttavia nascondere che tali considerazioni, pur ragionevoli e fondate sulle acquisizioni teoriche di importanti progetti di ricerca internazionali, presentano un *vulnus* che abbiamo finora sottratto volutamente all'attenzione del lettore: in assenza di una chiara e inequivocabile definizione del concetto di autenticità, ogni ragionamento sul tema sarà caratterizzato da un'intrinseca debolezza che inficia la costruzione di una teoria in grado di autosostenersi. Ed è per questo che nelle battute iniziali del nostro saggio abbiamo fatto riferimento al dato di esperienza⁷.

Cerchiamo quindi di approfondire il concetto di autenticità indagandone le fondamenta: secondo la definizione più autorevole e consolidata,

«un documento è autentico se è ciò che dichiara/mostra di essere ed è esente da alterazioni o corruzioni»⁸.

Tale definizione, benché sufficientemente generica da tollerare diverse interpretazioni, ad una lettura più approfondita risulta eccessivamente vaga, e appare insoddisfacente anche in ambito digitale. L'analisi puntuale della formula definitoria solleva delle criticità che vale la pena assumere come materia di ragionamento intorno a questi temi, con l'auspicio di raffinare ulteriormente il concetto di autenticità.

⁷ A onor del vero, anche il profilo esperienziale è tutt'altro che pacifico se solo si consideri che l'autenticità, ben lungi dall'essere una caratteristica immanente o latente degli oggetti, può essere interpretata come fenomeno sociale e, come tale, storicizzato. Tuttavia in questa sede non interessa tanto illustrare le diverse prospettive di analisi del concetto di autenticità, quanto rilevare alcune caratteristiche che ne evidenzino la complessa natura.

⁸ «Authenticity [is] the quality of a record that is what it purports to be and that is free from tampering or corruption». Cfr. InterPARES 2 Project, *International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records*, a cura di Luciana Duranti e Randy Preston, Roma: ANAI, 2008, p. 775, disponibile all'indirizzo http://www.interpares.org/display_file.cfm?doc=ip2_book_complete.pdf.

1. Posta in questi termini, l'introduzione fraudolenta di un documento autentico – di più: giuridicamente autentico – all'interno di un sistema informativo produce un documento autentico? Apparentemente sì, purché il documento sia ciò che dichiara di essere e sia esente da manipolazioni. Ma rispetto alla nostra interpretazione di autenticità⁹, rispetto all'esigenza di riconoscere l'identità di un oggetto ben oltre i confini dell'oggetto stesso, rispetto al ruolo dell'*ambiente* che dà forma e sostanza (cioè, significato) all'oggetto: rispetto a tutto ciò, il documento introdotto fraudolentemente non è un documento autentico, perché basta estendere lo sguardo al di là del semplice oggetto e considerare il *contesto* di produzione per verificare presumibilmente la presenza di una falla nel sistema documentario, nell'insieme dei processi che hanno consentito a quell'oggetto di essere assunto nel sistema senza averne diritto, non rispettando cioè quelle procedure che complessivamente garantiscono l'affidabilità del sistema a tutela dell'autenticità degli oggetti. Il punto è quindi interpretare in maniera assolutamente estensiva la locuzione "essere ciò che dichiara di essere" o – meglio ancora – riformularla, utilizzando ad esempio una formula del tipo: "un documento è autentico se è ciò che *manifesta* di essere", ove tale *presentazione verso l'esterno* sia riferita non solo al contenuto informativo del documento, ma anche al complesso intreccio di relazioni che legano il documento ad un contesto. In questo secondo caso, il documento introdotto fraudolentemente nel sistema non passerebbe il vaglio della definizione, poiché si propone evidentemente non solo come documento caratterizzato da una sua autonomia e da una sua intima natura (documento di debito, contratto, circolare; sottoscritto da Tizio o da Caio in una certa data; ecc.), ma come documento appartenente ad un sistema di relazioni, si *manifesta* cioè come documento del sistema informativo, e ciò sarebbe in contrasto con la realtà dei fatti. Sarebbe – ci si consenta – un falso non tanto dal punto di vista giuridico, quanto dal punto di vista archivistico.
2. Nella definizione, il generico riferimento all'assenza di alterazioni e corruzioni rinvia a quell'ipotesi di cristallizzazione degli oggetti che però abbiamo già escluso dal nostro orizzonte per evidenti motivi tecnici. Di nuovo, evitando di interpretare rigidamente la definizione onde salvarne lo spirito, la formula potrebbe essere riferita agli aspetti rilevanti – o meglio, essenziali – per garantire l'autenticità. Ma così si rischia evidentemente di cadere in una tautologia: un documento è autentico se è esente da alterazioni o corruzioni riferite agli aspetti rilevanti per l'autenticità. La nostra impressione è che il generico riferimento ad alterazioni o corruzioni sia inadeguato e convenga abbandonarlo piuttosto che cercare di modificarlo. Occorre riconoscere che il coraggioso

⁹ *Ibidem*.

tentativo del progetto InterPARES di trasportare in ambiente digitale concetti e metodi delle discipline tradizionali ha avuto valore più per il consistente *corpus* di analisi, riflessioni, materiali prodotti sull'argomento, che non per l'elaborazione di una teoria compiuta dell'autenticità in ambiente digitale. La profondità dei livelli di mediazione imposti dagli oggetti digitali e l'enorme varietà dei formati e delle strutture di dati sono tali che non risulta praticabile l'individuazione a priori di categorie concettuali, di profili, di caratteristiche meritevoli di attenzione specifica. Ad esempio, il degrado di un'immagine attraverso un'operazione di compressione che ne riduca la qualità grafica può essere inaccettabile in alcuni contesti, mentre in generale ciò non vale per un oggetto testuale, la cui qualità grafica è vincolata a soglie di tolleranza ben più ampie (non ci riferiamo ovviamente alla digitalizzazione di una pergamena). Una stessa operazione, quindi, ha ripercussioni diverse su oggetti diversi, perché in alcuni casi l'oggetto potrebbe non essere più riconosciuto come autentico. In altre parole: la qualità grafica è una caratteristica che ha un valore costitutivo e funzionale diverso a seconda dei casi. Ed è pertanto impossibile delinearne a priori il valore in riferimento all'autenticità. Ergo, la costruzione di una teoria dell'autenticità sarà sempre costretta ad assumere connotati di *genericità* e *rarefazione* tali da renderla – ci si passi la provocazione – inutilizzabile se non come quadro generale, come impianto teorico da cui derivare specifiche teorie, metodologie e applicazioni riferibili a specifici contesti e oggetti.

3. A ben pensarci – ulteriore apparente paradosso – ogni oggetto è in sé autentico e integro. Ciò che sembra mancare nella definizione di InterPARES è il legame con la risorsa primigenia: in altri termini, il concetto di autenticità è intimamente legato, anche se non appare all'evidenza, ad un profilo diacronico, ad un movimento nel tempo, ad un'idea quasi deterministica di individuazione del punto iniziale A da cui prendono le mosse i nostri ragionamenti (i fisici direbbero "dal punto t_0 ") e di un punto finale B (t_1). E volutamente usiamo il termine *punto*, intendendo con ciò non solo un istante nel tempo, ma un preciso ente dello spazio-tempo, cioè un evento che deve essere contestualizzato nel suo ambiente di produzione. È vero: le "alterazioni e corruzioni" di cui alla definizione assunta a modello evidentemente rimandano a questo profilo dinamico, ma oggettivamente si tratta di un richiamo piuttosto debole. Le conseguenze di tale approccio *dinamico* possono essere tutt'altro che scontate: in una visione per così dire *statica* noi concentriamo la nostra attenzione su un oggetto e ne seguiamo l'evoluzione verificando – o meglio, cercando di verificare – ad ogni generazione successiva, ad ogni passaggio cruciale nel ciclo di vita del documento, che i caratteri essenziali del documento permangano in esso, così da poter riconoscere nel nuovo oggetto l'antico. Accettiamo cioè una catena ininterrotta di trasformazioni che ci conducono dall'oggetto

A_1 all'oggetto A_n con tutta probabilità perdendo progressivamente traccia dell'oggetto iniziale, ma confidando nelle procedure che ci hanno condotto fino all'oggetto A_n e che garantiscono la sostanziale equivalenza con l'oggetto A_1 dal punto di vista dell'autenticità¹⁰.

In una visione più *dinamica* e fortemente fondata sul concetto di tempo e stati (in relazione al tempo), l'oggetto iniziale è solo l'inesco per una catena di trasformazioni che danno vita a nuovi oggetti, ognuno con una propria dignità: in questo caso, la migliore astrazione per rappresentare queste vicende è un grafo, una sorta di *stemma codicum* che racconti la storia dell'oggetto e che in maniera più neutrale descriva le trasformazioni non fornendo necessariamente garanzie sull'autenticità. La conseguenza immediata di questo approccio è ovviamente la necessità di conservare tutti gli oggetti della catena.

In linea di massima, tanto nell'ambito sociale quanto nei domini di settore finora si è preferito adottare la prima interpretazione, tutta concentrata sull'oggetto e sulla permanenza dei suoi caratteri significativi. In quest'ottica, si potrebbe dunque assumere l'integrità e l'identità di una risorsa come discriminare per l'attribuzione dell'autenticità, il che consentirebbe di proporre una definizione di autenticità più limpida, nella quale sia assorbito anche il profilo dinamico: una risorsa rimane autentica nel corso delle successive trasformazioni se conserva l'integrità e l'identità. Evidentemente, tale definizione non è in sé conclusa poiché rinvia ad altri concetti, ma presenta il duplice vantaggio di assorbire esplicitamente il profilo dinamico (le successive trasformazioni), e di spostare il *focus* del problema sulle due qualità cruciali dell'oggetto (integrità e identità), che pertanto richiedono – se non un'esplicita definizione – un adeguato approfondimento¹¹.

L'*integrità* di una risorsa si riferisce alla sua completezza (*wholeness*): una risorsa è integra quando è completa ed esente da corruzioni nei suoi aspetti fondamentali. Il processo di verifica dell'autenticità dovrebbe analizzare proprio tali aspetti e accertare che resistano agli inevitabili cambiamenti indotti dall'obsolescenza tecnologica. Per essere più concreti, nei processi conservativi il mantenimento del flusso di bit non è quasi mai necessario, mentre è inderogabile la conservazione e la completezza della cosiddetta *forma intellettuale*, con particolare riferimento a quelle caratteristiche che veicolano significati di rilievo. In altre parole, l'integrità

¹⁰ Di fronte ad una copia autenticata da un notaio non riserviamo interesse alcuno alle modalità con cui avviene il processo di autenticazione, confidando nel ruolo di terzietà del pubblico ufficiale e nella sua capacità di produrre documentazione fidefacente. Non importano i meccanismi: ciò che rileva è il risultato finale, un prodotto che – sotto il profilo giuridico – possa ritenersi equivalente all'originale in forza di una valutazione demandata ad un soggetto in grado di cogliere gli aspetti rilevanti del documento e verificarne la permanenza, con tutte le garanzie proprie della sua *auctoritas*.

¹¹ Le nozioni di integrità e identità presentate di seguito affondano le loro radici nel quadro concettuale elaborato all'interno del progetto InterPARES.

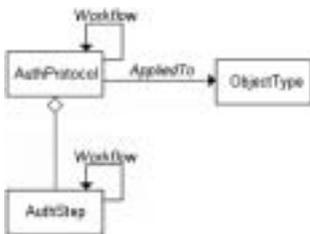
fisica di una risorsa (il *bitstream* originale) può essere corrotta, ma la struttura del contenuto e le componenti essenziali devono rimanere intatte.

L'integrità cioè non è mai totale: c'è sempre un grado di deterioramento. Quindi il problema cruciale è identificare le caratteristiche rilevanti, quelle che definiscono la *forma intellettuale*: questo significa comprendere la natura della risorsa, analizzarne le caratteristiche e valutarne la loro funzione in maniera tale da stabilire quali cambiamenti sono ammessi, senza timore che inficino l'integrità.

L'*identità* di una risorsa deve essere intesa in un'accezione molto ampia: essa si riferisce non solo alla sua univoca denominazione e identificazione. L'identità si riferisce all'insieme delle caratteristiche di una risorsa che univocamente la identificano e la distinguono da tutte le altre: si riferisce cioè non solo alla sua struttura concettuale interna, ma anche al suo contesto generale, cioè al sistema di relazioni cui partecipa (siano esse amministrative, legali, documentali, tecnologiche, perfino sociali). Una risorsa non è una monade isolata, con dei confini ben definiti e una vita propria: una risorsa è un oggetto *nel contesto*, è l'oggetto stesso e l'insieme delle relazioni che danno significato all'oggetto (riferendoci alle considerazioni precedentemente esposte, non è un punto isolato, bensì un punto dello spazio-tempo). Se vogliamo, si tratta di una banale considerazione epistemologica: conoscere gli oggetti significa conoscere un *intorno* di quegli oggetti, sufficientemente ampio da ricavarne un livello di significatività ritenuto discrezionalmente sufficiente. Il problema è che queste relazioni cambiano nel corso del tempo e quindi abbiamo bisogno non solo di comprenderle ed esplicitarle, ma anche di documentarle al fine di avere una *storia* completa della risorsa: non possiamo perdere questa *storia* senza perdere un po' dell'identità della risorsa, con conseguenze (negative) sulla valutazione dell'autenticità della risorsa stessa.

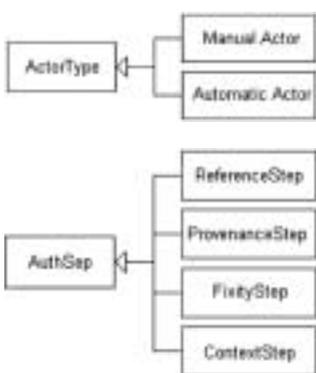
A proposito di valutazione: occorre fare una chiara distinzione fra l'autenticità di una risorsa e la procedura di *validazione* della risorsa, di verifica della sua autenticità. La seconda è parte di un processo più generale mirato ad assicurare che l'oggetto informativo sia conservato *come se, in luogo di*, insomma *funga da* originale. Gli strumenti di gestione dell'autenticità devono quindi monitorare e gestire protocolli e procedure all'interno della catena della custodia, dalla fase di creazione lungo tutto il processo di conservazione. L'autenticità non può essere valutata per mezzo di un indicatore booleano che ci dica se la risorsa è autentica o meno. Nonostante sembri strano, nella valutazione esistono dei *gradi di autenticità*: la certezza dell'autenticità di una risorsa è un obiettivo e i casi certi sono casi limite. La valutazione *tende asintoticamente all'autenticità*, in funzione del grado di controllo di oggetti e processi, e quindi dobbiamo predisporre meccanismi e strumenti conservativi ricordando sempre che alterazioni, corruzioni, perdite di dati significativi e altre modifiche interne o esterne sono dietro l'angolo, anzi sono una certezza; dobbiamo cioè progettare strumenti e *pesi* per comprendere il rilievo di tali cambiamenti e il loro impatto sull'autenticità. Continuando a svolgere il filo del paradosso, dobbiamo fare della probabilità e della statistica il terreno delle certezze sui cui rifondare taluni metodi.

La conseguenza è che gestire l'autenticità implica operare su un insieme appropriato di attributi relativi a contenuto e contesto, e verificare (magari con l'ausilio di un'apposita metrica) la completezza o il grado di alterazione di questo insieme. Nel progetto europeo CASPAR è stato tentato un approccio basato su queste riflessioni¹². L'osservazione cruciale da cui siamo partiti è che la protezione dell'autenticità e la sua valutazione costituiscono un *processo*. Per gestire tale processo abbiamo bisogno di definire le procedure che occorre seguire in relazione a specifiche tipologie di oggetti e di eventi. Abbiamo definito una tale procedura Authenticity Protocol (AP).



Un AP è un insieme di passi fra loro collegati, ognuno dei quali prende il nome di Authenticity Step (AS). Ogni *step* modella una parte del protocollo (AP) e può essere eseguito autonomamente, come un processo autonomo; l'Authenticity Step (AS) costituisce in pratica una fase del processo globale mirato alla valutazione dell'oggetto. Le relazioni fra i vari passi stabiliscono in che ordine questi debbano essere eseguiti nel contesto di uno specifico protocollo. Senza entrare nei dettagli, denotiamo semplicemente con Workflow l'insieme di tali relazioni. A sua volta, ogni Authenticity Protocol può essere usato ricorsivamente per definire altri protocolli, come rappresentato dalla presenza della relazione di Workflow.

Il protocollo di autenticità si applica a un ObjectType, una classe di oggetti con caratteristiche uniformi per l'applicazione del protocollo.



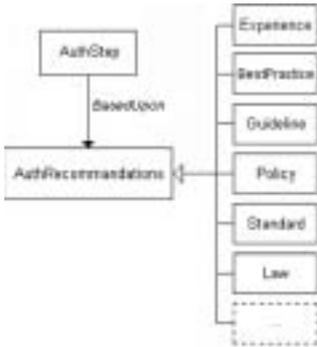
Uno *step* è eseguito da un ActorType, sia esso automatico (hardware o software) o manuale (ente o persona), istanziato da un Actor.

Ci possono essere vari tipi di Authenticity Step: coerentemente con le indicazioni del modello OAIS¹³, distinguiamo gli *step* sulla base delle diverse tipologie di Informazioni sulla conservazione, e quindi avremo *step* riferibili all'identificazione (*reference step*), *step* riferibili alla provenienza (*provenance step*), altri riferibili all'integrità (*fixity step*) e altri ancora riferibili al contesto (*context step*).

Qualunque analisi svolta sull'oggetto può essere riferita a uno di questi *step* o ad una loro combinazione.

¹² Il sito ufficiale del progetto è all'indirizzo <http://www.casparpreserves.eu>. L'Autore ha partecipato al progetto di ricerca come membro del team dell'Università di Urbino che ha collegialmente sviluppato tale modello, con il contributo scientifico dell'ISTI (CNR Pisa) e di Engineering Ingegneria Informatica.

¹³ OAIS: *sistema informativo aperto per l'archiviazione*, a cura di Giovanni Michetti, Roma: ICCU, 2007.



Uno *step* implica un'analisi per una successiva valutazione e pertanto sono necessarie informazioni relative a:

- *best practice*, metodologie e ogni tipo di regolamentazione che deve essere seguita o che possa aiutare nell'analisi e nella valutazione;
- criteri che devono essere eventualmente soddisfatti nella valutazione.



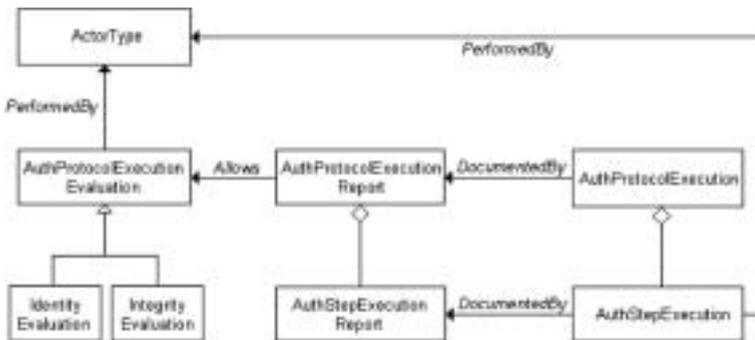
Gli Authenticity Protocol sono eseguiti da un attore su oggetti che appartengono ad una specifica tipologia nel contesto di un'Authenticity Protocol Execution Session. L'esecuzione di un AP è modellata come un Authenticity Execution Protocol (APE). Un APE è correlata ad un AP mediante la relazione ExecutionOf e consiste di vari Authenticity Execution Steps (ASE). Ogni ASE, a sua volta, è correlata ad un AS mediante un'associazione di tipo ExecutionOf, e contiene le informazioni relative all'esecuzione, compresi:

– l'attore che ha fatto l'esecuzione;

- l'informazione utilizzata;
- luogo, tempo e contesto di esecuzione.

Tipi differenti di Authenticity Step Execution hanno strutture differenti e i risultati delle esecuzioni devono essere documentati al fine di raccogliere informazioni su specifici aspetti della risorsa (ad esempio denominazione, consistenza, date e trasformazioni).

Un Authenticity Step Execution Report documenta semplicemente che lo *step* è stato eseguito e raccoglie tutti i valori associati ai metadati analizzati nel corso



Il tema dell'autenticità non può essere certo esaurito in queste poche pagine. Non è stato affrontato ad esempio il profilo del *comportamento*: conservare gli oggetti digitali significa conservare degli oggetti che hanno appunto un comportamento, delle dinamiche. L'autenticità gioca un ruolo cruciale anche in questo ambito, giacché un appiattimento dei fattori dinamici di un oggetto è in contrasto – in generale – con la permanenza dei caratteri di identità dell'oggetto.

Non abbiamo affrontato gli aspetti *organizzativi*: molti ritengono la firma digitale un utile strumento di supporto alle strategie di conservazione e verifica dell'autenticità delle risorse digitali, ma in che misura tale tecnologia può essere d'aiuto? e – soprattutto – qual è il peso della costruzione di una Public Key Infrastructure (PKI) che fornisca affidabilità a tale tecnologia? è possibile prevedere dei modelli organizzativi di rango locale – ad esempio delle infrastrutture di certificazione a livello di ente o di *network* – preservando l'affidabilità di questo strumento¹⁴? e come gestire concettualmente e praticamente una gerarchia dei certificatori? è ipotizzabile una *gerarchia* dell'autenticità? nel caso, come dovremmo differenziare la gestione e la conservazione delle fonti in funzione della loro autenticità?

Non è stato affrontato neppure il profilo *politico*, sovrapposto in parte a quello organizzativo: come ormai consolidato nella letteratura di settore, i depositi digitali certificati necessitano del ruolo del conservatore come di una terza parte fidata, una terza parte cioè che:

- non abbia un interesse soggettivo sui documenti¹⁵;
- non abbia ragioni specifiche per alterare la documentazione oggetto di custodia;
- non consenta ad alcuno di alterare la documentazione, accidentalmente o di proposito.

Come garantire tutto ciò non solo dal punto di vista tecnico, ma soprattutto dal punto di vista *politico*? Come individuare l'architettura complessiva e le figure specifiche che nel nostro ordinamento potrebbero assumere questo ruolo? È vero: esiste una normativa di settore; ed esiste una fitta rete di istituti di conservazione sul nostro territorio. Ma il mutato contesto sociale, le mutate esigenze tecniche, perfino le nuove configurazioni istituzionali e costituzionali richiedono un ripensamento di queste logiche per non fotocopiare semplicemente in ambiente digitale strutture e metodi dell'ambiente tradizionale. Sono cioè necessari interventi di carattere politico-organizzativo, come ad esempio:

¹⁴ La nostra osservazione ha un carattere generale: il legislatore italiano ha già dato una risposta a tale quesito, elaborando un complesso *corpus* di norme – in verità niente affatto scevro da importanti criticità interpretative – sulla firma digitale e sui relativi processi di certificazione.

¹⁵ Efficacemente, in letteratura si parla talvolta di "soggetti che non abbiano una *quota* di partecipazione (*stake*) nei documenti".

- una chiara definizione dei ruoli e delle responsabilità in sede conservativa;
- uno sviluppo coerente di raccomandazioni e politiche per la costruzione di depositi affidabili;
- una precisa identificazione delle singole componenti della funzione conservativa.

Tutto ciò, coerentemente con una chiara definizione degli elementi concettuali sui quali fondare la complessa architettura che sostenga la *mission* conservativa in ambiente digitale.

In breve, i profili d'indagine sono molteplici: in questa sede noi abbiamo inteso solo proporre degli spunti di riflessione, evidenziando la necessità di una rigorosa analisi e di un solido impianto concettuale e metodologico per elaborare un modello grafico-simbolico che sorregga le nostre astrazioni e passi il vaglio dell'implementazione. Viceversa, la rappresentazione grafica del modello, basata sulla notazione formale dei linguaggi di modellazione, non solo conferma l'utilità di metodi e tecniche appartenenti a settori disciplinari diversi dal nostro, ma soprattutto rivela che tali costruzioni simboliche sono veicolo di un'idea, di un preciso approccio e di una determinata interpretazione del mondo, soggiacenti a schemi niente affatto neutri.

L'autenticità richiede una riflessione articolata, uno sforzo in più direzioni, una visione ad ampio raggio¹⁶: non si tratta più di *tradurre* soluzioni consolidate dall'ambiente tradizionale a quello digitale, senza soluzione di continuità. Qui si tratta di *reinterpretare*, rinnovare, cambiare punto d'osservazione. Se occorre, con un pizzico d'incoscienza.

Authenticity is one of the critical aspects in long-term digital preservation. Despite all efforts and research initiatives that have been promoted at the international level, the scientific community cannot yet rely on the existence of a thorough theory on authenticity in a digital environment. On the other hand, an overall framework of reference, made of analytical tools, critical understanding and reference materials, does exist and is by now fairly consolidated. The present paper explores the debate on this issue and, starting from the paradox of digital preservation – constantly thorn by the conflicting need to transform the objects without modifying them – highlights a number of critical aspects that should be considered in order to better define authenticity as a notion. The paper then follows up to its suggestions by briefly illustrating the authenticity model developed by the EU CASPAR project – a model which well adheres to the theoretical approach proposed by the author is the first part of the essay.

¹⁶ «Creating a common understanding about the multiple meanings and significance of authenticity is critical in the digital environment, in which information resources exist in many formats yet are interactive». Cfr. Council on Library and Information Resources, *Authenticity in a digital environment*, Washington D.C., 2000, p. vii.