

Archivi digitali d'autore: ipotesi di lavoro

Pierluigi Feliciati¹, Stefano Allegrezza², Ernesto Belisario³, Salvatore Vassallo⁴, Emmanuela Carbé⁵, Primo Baldini⁶, Paul Gabriele Weston⁷ (coordinatore del panel)

¹ Università di Macerata, Italia - pierluigi.feliciati@unimc.it

² Università di Udine, Italia - stefano.allegrezza@uniud.it

³ e-lex, Italia - ebelisario@e-lex.it

⁴ Archivum Romanum Societas Iesu, Italia - vassallosalvatore@gmail.com

⁵ Università di Pavia, Italia - emmanuela.carbe@gmail.com

⁶ Università di Pavia, Italia - primo.baldini@unipv.it

⁷ Università di Pavia, Italia - paul.weston@unipv.it

ABSTRACT

Il panel mira ad attraversare su più livelli il tema della conservazione del digitale nativo nell'ambito degli archivi di persona, con particolare riguardo alla conservazione e gestione dei materiali di autori contemporanei. L'intento è quello di avviare una riflessione condivisa con la comunità scientifica su metodologie e strumenti da adottare in questo campo di indagine non del tutto esplorato e ancora carico di criticità teoriche e operative. Se è chiaro ormai che la conservazione degli archivi nativi digitali è una sfida sempre più importante nei più diversi settori della nostra società (l'ambito della pubblica amministrazione e quello forense sono tra gli esempi più noti), è altrettanto vero che per gli archivi di persona, nel contesto specifico della produzione letteraria, i casi di studio a nostra disposizione sono ancora sporadici e non sufficienti per individuare del tutto delle buone pratiche. Pare dunque necessario un confronto all'interno della comunità scientifica delle DH, non tanto per condividere risultati quanto per illustrare le problematiche maggiori e individuare delle ipotesi di lavoro. Il panel vuole aprirsi al dibattito con un approccio multidisciplinare, cercando di affrontare le questioni archivistiche più importanti relative all'acquisizione e conservazione dei fondi, alla loro archiviazione, gestione e fruizione. Verrà dato particolare rilievo alla scelta di formati e standard, alla qualità dei dati e dei servizi, ai numerosi aspetti critici sotto il profilo giuridico, alla gestione del versioning e agli strumenti per la ricerca e la didattica. Nel corso della sessione saranno esaminate due esperienze italiane: il lavoro sull'archivio ibrido di Massimo Vannucci e il progetto PAD – Pavia Archivi Digitali, che ha avviato nel 2010 un percorso di sperimentazione su archivi di scrittori contemporanei, perfezionando negli anni alcune strategie per le fasi di preparazione dei conferimenti e acquisizione dei materiali.

PAROLE CHIAVE

Born-digital archiving, Digital Curation, Digital Preservation, Private Papers Archiving, Born-digital papers, Personal Digital Archiving, RiC Model

1. INTRODUZIONE

La sfida rappresentata dalla creazione di depositi per la documentazione nativamente digitale non è data tanto dalla individuazione e messa in atto delle procedure di riversamento dei file, quanto dall'impegno richiesto per consentirne nel tempo la fruizione. Non c'è dubbio che il numero dei file che sono oggetto di archiviazione pone già notevoli difficoltà in quanto solitamente eccede di gran lunga quello di un archivio di natura cartacea, con l'ulteriore difficoltà di gestire una grande varietà di formati e di tenere in vita supporti connessi a generazioni tecnologiche ormai in completo disuso. Sfida particolarmente impegnativa è quella della descrizione degli archivi, in primo luogo per il numero davvero notevole dei file e per la diversità dei formati. È, tuttavia, la fluidità del mezzo digitale che rappresenta per l'archivista-conservatore la vera novità rispetto all'archivio tradizionale. Questa fluidità determina un maggior numero di relazioni e di intrecci tra i contenuti, ma soprattutto la stratificazione dei file. Ogni versamento di documenti, che in un archivio tradizionale è incrementale e non produce sostituzioni, in un archivio digitale può produrre il fenomeno opposto: la sostituzione di contenuti, lo spostamento di un file da una cartella all'altra, i cambiamenti di metadati tecnici derivanti dalla sostituzione di un computer – o di una componente software – non sono che alcune delle manifestazioni più comuni. Inoltre, la gestione dell'archivio digitale presuppone l'aggiornamento costante dei modelli di dati, degli standard e delle procedure per far fronte alla crescente varietà delle fonti documentarie.

In aggiunta ai noti rischi ai quali sono esposti gli oggetti digitali, nonché alla stratificazione delle versioni prima accennata, intervengono come altri fattori di particolare criticità la tutela dei dati sensibili, la gestione dei diritti, le licenze d'uso degli applicativi software (nell'evenienza di trasferimenti di pacchetti software dal soggetto utilizzatore a quello che si occupa della conservazione), l'organizzazione e le responsabilità della curatela digitale e, infine, il rapporto con l'utenza.

Proprio quest'ultimo fattore rappresenta uno degli elementi di maggiore incertezza rispetto alle soluzioni da adottare per la restituzione dei dati. Idealmente si fronteggiano due approcci: quello archivistico, in cui la descrizione mira a riprodurre il contesto di provenienza, ossia l'organizzazione che l'autore ha dato alle proprie memorie elettroniche, e quello bibliocentrico, in cui sono le singole opere – anche in senso astratto – a fungere da elemento di aggregazione dei contenuti associati: il materiale pretestuale, le diverse stesure del testo, il materiale scambiato con l'editore, le diverse edizioni parziali o integrali, le presentazioni delle opere, le recensioni, la voce dei lettori, i rapporti con le altre opere e così via. Evidentemente questo secondo approccio, che per certe fasce di utenza può essere più funzionale e intuitivo, ci dice assai

poco sulle abitudini di scrittura dell'autore, sul modo in cui lui/lei hanno organizzato i propri spazi mentali, digitali, relazionali.

Il successo di un progetto finalizzato alla conservazione a lungo termine di archivi privati digitali nativi dipende dalla capacità di individuare fin dall'inizio le soluzioni più adeguate (riguardo a linee guida, standard, buone pratiche, ecc.), con l'intesa che abbiano anche la caratteristica della sostenibilità.

Il caso degli archivi nativi digitali letterari presenta, rispetto ad altri casi come quello offerto da PARADIGM (Personal Archives Accessible in Digital Media), una serie di specificità che richiedono soluzioni *ad hoc*. Importanti progetti sono già stati avviati da tempo, come quello dell'Harry Ransom Center e dell'Emory University's Manuscript, Archives and Rare Book Library. Nel corso di questa sessione, che offrirà un attraversamento su più livelli delle tematiche qui menzionate, saranno illustrati anche due casi italiani: il prezioso lavoro sull'archivio ibrido di Massimo Vannucci, che rappresenta nel panorama nazionale forse il primo tentativo di inventariazione e riordinamento di un archivio di questo tipo; e l'esperienza del progetto PAD – Pavia Archivi Digitali, che dal 2010 opera a fianco del Centro Manoscritti di Pavia per acquisire e conservare archivi digitali nativi di autori contemporanei, con lo scopo futuro di metterli a disposizione dell'utenza per attività di studio e ricerca.

Considerando l'insieme delle problematiche che la documentazione, gestione e conservazione di questi archivi sottendono, si rende dunque evidente la necessità di integrare fra loro competenze appartenenti ad ambiti molto diversi: discipline letterarie e filologiche, ma anche giuridiche, tecniche archivistiche, tecnologia dell'informazione, aspetti amministrativi. Si pensi ad esempio alla conservazione a lungo termine e ai risvolti metodologici, tecnologici e organizzativi connessi, conosciuti in gran parte, ma su cui ancora non c'è la garanzia di stabilità e affidabilità che sarebbe auspicabile. O ancora, alla descrizione archivistica: pur basandosi su standard ben conosciuti e in genere periodicamente aggiornati, le caratteristiche di questi archivi hanno un impatto particolare che necessita di specifiche evoluzioni, per la numerosità dei file conservati, il *versioning*, la presenza degli stessi file in più cartelle, l'intreccio di relazioni tra file, eventi, responsabilità e agenti, l'assenza di politiche condivise, di massimari di scarto e di piani di conservazione. Una particolare attenzione al modello FRBRoo può essere opportuna, oltre all'osservazione della diffusione di standard come RDA, modalità di pubblicazione come LOD e formati di scambio come JSON. Altre competenze necessarie riguardano lo studio di ambito filologico, per la corretta valorizzazione delle "carte", attraverso una lettura e interpretazione in cui si tenga conto anche delle caratteristiche tecniche dei file. Poiché è probabilmente irrealistico pensare che in ogni contesto siano presenti da subito queste e altre competenze, è forse opportuno pensare da un lato a progetti di messa a fattor comune della ricerca e dello sviluppo, dall'altro a programmi di formazione mirata.

2. P. FELICIATI – LA QUALITÀ DEGLI ARCHIVI LETTERARI DIGITALI NATIVI: DALLA SEDIMENTAZIONE AL SERVIZIO

La complessità richiesta per la corretta progettazione, creazione, gestione e conservazione a lungo termine degli archivi digitali, nella dialettica continua tra sviluppo delle soluzioni tecnologiche, standard tecnici e procedurali e normativa nazionale e internazionale, è nota e molto studiata da molti. Uno degli snodi critici riguarda la garanzia di contestualizzazione del documento digitale, inseguendo da un lato il modello di vincolo proprio della tradizione archivistica e guardando dall'altro alla sfida che pone la multidimensionalità della sedimentazione, della strutturazione degli elementi del documento e le conseguenze che tutto ciò ha sull'accesso finale. Se si considerano gli archivi letterari nativi digitali o ibridi, le questioni si fanno più complesse e intriganti: quale vincoli tra entità informative (che siano *files* o no) sono da mantenere o creare ex-novo per non svilire il senso complessivo della sedimentazione documentaria dell'attività di uno scrittore? Come gestire il corpus documentario, fatto di entità informative e delle loro relazioni qualificate? Quali scelte selettive ci possiamo permettere di compiere? E come rappresentarlo al fine dell'accesso da parte di utenti finali, ammesso che le caratteristiche della *collezione* siano state studiate e costituite stabilmente? Questo contributo si propone di discutere di queste e altre questioni, inserendole in un modello di qualità che garantisca la migliore gestione e il miglior uso possibile degli archivi letterari digitali nativi.

3. S. ALLEGREZZA – PERSONAL DIGITAL ARCHIVING: CRITICITÀ E SOLUZIONI NELLA CONSERVAZIONE A LUNGO TERMINE DEGLI ARCHIVI LETTERARI DIGITALI

L'intervento si propone di affrontare il tema del *Personal Digital Archiving* con particolare riferimento agli archivi di letterati, fornendo da una parte un quadro aggiornato sugli sviluppi più recenti del dibattito in corso e dall'altro descrivendo l'intervento d'inventariazione e riordinamento dell'archivio "ibrido" – contenente cioè documenti sia cartacei che digitali – di Massimo Vannucci (1957-2012), politico e imprenditore. Si tratta probabilmente del primo caso in Italia di intervento su un archivio digitale di persona: la documentazione risale prevalentemente al periodo compreso fra il 2005 e il 2012, quindi all'epoca dell'attività politica di Vannucci, ed è proprio la natura ibrida dei materiali a determinare l'originalità dell'intervento: e-mail archiviate sia nel client di posta che sulla webmail; documenti rinvenuti nel computer e nello smartphone; post sui social media, si aggiungono alla documentazione cartacea, costituendo a tutti gli effetti un archivio "ibrido" di persona. Oltre a illustrare il lavoro svolto, l'intervento si pone l'obiettivo di aprire una riflessione sulle criticità connesse con la formazione degli archivi di persona in ambiente digitale: tematiche di importanza fondamentale se si pensa che gli archivi – non soltanto di persona – diventeranno totalmente digitali nel giro di poco tempo, dato che i trend evolutivi degli ultimi anni hanno registrato una fortissima accelerazione in tale direzione. È urgente, quindi, aprirsi al confronto su

queste tematiche e individuare soluzioni che permettano di conservare per il futuro gli archivi digitali che si stanno oggi formando.

4. S. VASSALLO – GLI STANDARD ARCHIVISTICI ALLA SFIDA DEGLI ARCHIVI LETTERARI

Di fronte alla descrizione di Archivi Letterari e di altri archivi che non ricadono nella stabile casistica degli archivi di enti, gli standard archivistici (come ISAD, International Standard for Archival Description) sembrano fallire nel tentativo di imbrigliare documentazione per sua natura fluida, con versioni differenti dello stesso documento che si sovrappongono come onde, in una sorta di griglia autocompilativa.

Ciò in gran parte è dovuto a un'erronea interpretazione degli standard, degenerati da filosofie descrittive a precetti, tabelle descrittive fino a diventare anche rigidi e unici modelli di visualizzazione e accesso per l'utente.

RiC (Records in Contexts) da questo punto di vista guarda all'utente, lo asseconda rappresentando il mondo non solo e non tanto come una rigida impalcatura gerarchica in cui incasellarlo quanto come una rete di significati e livelli differenti di lettura intrecciati fra di loro.

La sfida raccolta dai nuovi standard archivistici è proprio quella di fornire le chiavi per poter dar conto e poter descrivere un mondo complesso che si snoda a cavallo fra beni culturali differenti e metodologie per descriverli diverse.

Progetti come SNAC (Social Networks and Archival Context), a titolo di esempio, mostrano infatti come i monoliti descrittivi non possano funzionare in un mondo, quello delle digital humanities, fatto di condivisione e comunicazione.

5. E. BELISARIO – CREA, GESTISCI, CONSERVA: LE NORME SUL CICLO DI VITA DEL DOCUMENTO INFORMATICO

Nel corso degli ultimi vent'anni, il legislatore si è occupato ripetutamente delle norme sulla formazione, gestione, archiviazione e conservazione dei documenti informatici.

Obiettivo primario è stato consentire di avviare il difficile percorso di digitalizzazione della pubblica amministrazione e di dematerializzazione dei suoi archivi. Tuttavia, le stesse norme si applicano anche ai privati che vogliono formare documenti e archivi aventi piena validità giuridica.

Con la recente riforma del Codice dell'Amministrazione Digitale (Decreto Legislativo n. 82/2005), operata a seguito della riforma della Pubblica Amministrazione, il quadro regolatorio è ormai completo e costringe tutti i soggetti giuridici ad un complesso adeguamento: tecnologico, organizzativo, giuridico e contrattuale.

Le decisioni da assumere in questa fase sono fondamentali se si vuole conservare non solo l'integrità e la disponibilità, ma anche il valore legale degli archivi.

Per questo motivo, è opportuno definire le caratteristiche delle soluzioni e delle procedure che consentano di gestire e conservare in modo adeguato il patrimonio documentale.

6. E. CARBÉ, P. BALDINI – I PROCESSI DI ACQUISIZIONE DI ARCHIVI LETTERARI DIGITALI NATIVI: IL CASO PAD

I fondi conservati in PAD – Pavia Archivi Digitali, difforni tra loro per mole, tipologia e supporti di provenienza, rappresentano una casistica davvero ampia che dimostra le molteplici criticità nella gestione di un archivio digitale nativo di tipo letterario, a partire dalla fase di ingest. L'intervento mira a illustrare le metodologie e gli strumenti raffinati in questi anni nell'ambito dei processi di acquisizione dei fondi, su cui ci si aspettano ulteriori miglioramenti grazie a un proficuo confronto con la comunità scientifica. La preparazione del conferimento prevede una fase conoscitiva preliminare con l'autore, sia per la definizione delle modalità di acquisizione, regolate per gli aspetti legali da un contratto, sia per una prima raccolta di informazioni utili alla preservazione del contesto di origine del materiale, effettuata anche grazie a un questionario, che viene inviato all'autore insieme con una Guida al conferimento. L'acquisizione degli archivi avviene tramite un hard disk dotato di crittografia hardware, qualora il trasferimento dei file avvenisse da un computer; in altri casi (ad esempio il salvataggio di siti internet, o ancora il recupero di file contenuti in vecchi floppy disk) sono messe in atto soluzioni alternative, sempre documentate per preservare la storia archivistica del fondo. All'arrivo del materiale a Pavia, sono attivate procedure per la messa in sicurezza e per il controllo dell'integrità. Il materiale viene poi verificato insieme con l'autore e infine convalidato. Solo a questo punto il conferimento risulta concluso: si passa dunque al trattamento archivistico dei materiali, ora in fase di sperimentazione, e alle più consolidate procedure di conservazione a lungo termine.

7. BIBLIOGRAFIA

- [1] Alfieri A., Feliciati P. 2017. *Gli archivi online per gli utenti: premesse per un modello di gestione della qualità*. In: J LIS.it, 8, 1: 22-28. DOI: <http://dx.doi.org/10.4403/jlis.it-12269>.
- [2] Allegrezza S., Gorgolini L. 2016. *Gli archivi di persona nell'era digitale. Il caso dell'archivio di Massimo Vannucci*. Bologna, Il Mulino. ISBN: 978-88-152648-00.
- [3] Archivio Massimo Vannucci. URL: <http://www.archiviomassimovannucci.it>.
- [4] Barrera-Gomez J., Erway R. 2013. *Walk This Way: Detailed Steps for Transferring Born-Digital Content from Media You Can Read In-house*. Dublin, OH, OCLC Research. URL: <http://www.oclc.org/content/dam/research/publications/library/2013/2013-02.pdf>.

- [5] Carbé E. 2017. *Walking in My Shoes: A Case Study from a Born-digital Archive*. In: CEUR, 2021: 137-146. URL: <http://ceur-ws.org/Vol-2021/>
- [6] Carroll L., Farr E., Hornsby P., Ranker B. 2011. *A Comprehensive Approach to Born-Digital Archives*. In: *Archivaria*, 72: 61-92. DOI: <https://open.library.emory.edu/publications/emory:cksgv/pdf/>.
- [7] Consultative Committee for Space Data Systems 2012. *Reference Model for an Open Archival Information System (OAIS)*. Washington DC, CCSDS Secretariat. URL: <https://public.ccsds.org/pubs/650x0m2.pdf>.
- [8] Delneri, F. 2017. *Il documento amministrativo informatico: un cammino per approssimazione. Criticità e risposte possibili, tra normativa e prassi, dalla formazione alla conservazione*. In: *JLIS.it*, 8, 3: 26-38. URL: <https://www.jlis.it/article/view/12193/11275>.
- [9] Dooley J. 2015. *The Archival Advantage: Integrating Archival Expertise into Management of Born-digital Library Materials*, Dublin, Ohio, OCLC Research. URL: <http://www.oclc.org/content/dam/research/publications/2015/oclcresearch-archival-advantage-2015.pdf>.
- [10] Erway, R. 2010. *Defining "Born Digital". An Essay*. Dublin, Ohio: OCLC Research. URL: <http://www.oclc.org/research/activities/hiddencollections/borndigital.pdf>.
- [11] Erway, R. 2012. *You've got to Walk Before You Can Run: First Steps for Managing Born-Digital Content Received on Physical Media*. Dublin, Ohio, OCLC Research. URL: <http://www.oclc.org/content/dam/research/publications/library/2012/2012-06.pdf>.
- [12] Erway R., Goldman B., McKinley M. 2014. *Agreement Elements for Outsourcing Transfer of Born Digital Content*. Dublin, Ohio, OCLC Research. URL: <http://www.oclc.org/content/dam/research/publications/library/2014/oclcresearch-born-digital-content-transfer-2014.pdf>.
- [13] Experts Group on Archival Description. *Records in Contexts – Conceptual Model*. Versione per commenti pubblici. URL: <https://www.ica.org/en/egad-ric-conceptual-model>.
- [14] Feliciati P. 2016. *L'usabilità degli ambienti bibliotecari e archivistici digitali come requisite di qualità: contest, modelli e strumenti di valutazione*. In: *JLIS.it*, 7, 1: 113-130. DOI: <http://dx.doi.org/10.4403/jlis.it-11512>.
- [15] Gueguen, G., da Fonseca V.M.M., Pitti D.V., Sibille-de Grimouard, C. 2013. *Toward an International Conceptual Model for Archival Description: A Preliminary Report from the International Council on Archives' Experts Group on Archival Description*. In: *The American Archivist*, 76, 2: 566-583. URL: https://www.ica.org/sites/default/files/EGAD_English.pdf.
- [16] Guercio M. 2013. *Conservare il digitale. Principi, metodi e procedure per la conservazione a lungo termine di documenti digitali*. Roma-Bari, Laterza. ISBN: 978-88-593000-9-0.
- [17] Guercio M. 2016. *Archivi personali: la sfida del digitale. Una riflessione*. In: *Gli archivi di persona nell'era digitale. Il caso dell'archivio di Massimo Vannucci*, a cura di S. Allegrezza e L. Gorgolini. Bologna, Il Mulino: 133-148.
- [18] Kirschenbaum M.G., Farr E.L., Kraus K.M., Nelson N., Peters C.S., Redwine G., Reside D. 2009. *Digital Materiality: preserving access to computer as complete environments*. In: *iPres 2009. The Sixth International Conference on Preservation of Digital Objects*. California Digital Library, UC Office of the President: 105-112. URL: <https://escholarship.org/uc/item/7d3465vg>.
- [19] Kirschenbaum M.G., Ovenden R., Redwine G. 2010. *Digital Forensics and Born-Digital Content in Cultural Heritage Collections*. Washington DC, Council on Library and Information Resources. ISBN: 978-1-932326-37-6. URL: <https://www.clir.org/pubs/reports/pub149/pub149.pdf>.
- [20] Kuny T. 1998. *A Digital Dark Ages? Challenges in the Preservation of Electronic Information*. In: *International Preservation News*, 17. URL: <https://archive.ifla.org/IV/ifla63/63kuny1.pdf>.
- [21] International Council of Archive. *General International Standard for Archival Description*. URL: <https://www.ica.org/en/isadg-general-international-standard-archival-description-second-edition>.
- [22] Lavoie, B. 2014. *The Open Archival Information System (OAIS) Research Model: Introductory Guide (2nd Edition)*. Dublin, Ohio, OCLC Research, Digital Preservation Coalition. DOI=<http://dx.doi.org/10.7207/TWR14-02>.
- [23] Lenzini L. 2016. *Don't Save! Fortini, un Mac e le ultime cose*. In: *L'Ulisse. Rivista di poesia, arti e scritture*, 19, 96-101. URL: <http://www.lietocolle.com/cms/wp-content/uploads/2016/04/ULISSE-19.pdf>.
- [24] Lynch T. J. 2014. *Social Networks and Archival Context Project: A Case Study of Emerging Cyberinfrastructure*. In: *Digital Humanities Quarterly*, 8, 3. URL: <http://www.digitalhumanities.org/dhq/vol/8/3/000184/000184.html>.
- [25] PAD – Pavia Archivi Digitali: pad.unipv.it.
- [26] Stollar Peters, C. 2006. *When Not All Papers are Paper: A Case Study in Digital Archiving*. In: *Journal of the Society of Georgia Archivists*, 24: 22-34. URL: <http://digitalcommons.kennesaw.edu/cgi/viewcontent.cgi?article=1068&context=provenance>.
- [27] Vassallo S. 2013. *L'integrazione tra archivi e biblioteche alla prova del web semantico*. In: *Biblioteche in cerca di alleati. Oltre la cooperazione, verso nuove strategie di condivisione*. Milano: Editrice Bibliografica. ISBN: 978-88-707576-5-1.
- [28] Weston P.G. 2013. *Digital curation: nuove opportunità per bibliotecari e archivisti*. In: *Il libro, gli archivi e la memoria digitale*, a cura di Paul Gabriele Weston, Milano, Unicopli: 9-66. ISBN: 978-88-400168-1-8.
- [29] Weston P.G., Carbé E., Baldini P. 2016. *Hold it All Together: a Case Study in Quality Control for Born-Digital Archiving*. In: *Qualitative and Quantitative Methods in Libraries (QQML)*, 5: 695-710. URL: http://www.qqml.net/papers/September_2016_Issue/5313QQML_Journal_2016_Westonet_695-710.pdf.
- [30] Weston P.G., Carbé E., Baldini P. 2017. *Se i bit non bastano: pratiche di conservazione del contesto di origine per gli archivi letterari nativi digitali*. In: *Bibliothecae*, 6, 1: 154-177. DOI: <https://doi.org/10.6092/issn.2283-9364/7027>.
- [31] *Workbook on Digital Private Papers 2008*. URL: <http://www.paradigm.ac.uk/workbook/index.html>.