

Dig *Italia*

Anno XIV, Numero 1 - **2019**

Rivista del digitale nei beni culturali

ICCU-ROMA



ICCU

Istituto centrale per il catalogo unico
delle biblioteche italiane e per le informazioni bibliografiche
<https://www.iccu.sbn.it>

Copyright © ICCU - Roma

La riproduzione totale o parziale del contenuto della rivista
è ammessa con obbligo di citazione

Digitalia

Rivista del digitale nei beni culturali
ISSN 1972-6201
Anno XIV, Numero 1 - Giugno 2019

In copertina:

L'immagine è una libera elaborazione grafica della testa della statua di Apollo del I sec. d.c. (Civitavecchia, Museo Nazionale), copia da un originale greco avvicicabile all'Apollo di Leochares (IV sec. a.c.)

Direttore Fondatore

Marco Paoli

Direttore Responsabile

Simonetta Buttò

Comitato di Redazione

Capo Redattore:

Elisabetta Caldelli

Amalia Maria Amendola
Valentina Atturo
Lucia Basile
Laura Borsi
Flavia Bruni
Elisabetta Castro
Massimina Cattari
Silvana de Capua
Carla Di Loreto
Maria Cristina Di Martino
Vilma Gidaro
Egidio Incelli
Maria Cristina Mataloni
Massimo Menna
Lucia Negrini
Paola Puglisi
Alice Semboloni
Vittoria Tola
Maria Lucia Violo

Grafica & Impaginazione

MLA&Partner - Roberta Micchi

Produzione e Stampa

Istituto Poligrafico e Zecca dello Stato S.p.A.
Roma

Editore

ICCU

Istituto centrale per il catalogo unico
delle biblioteche italiane
e per le informazioni bibliografiche
Viale Castro Pretorio, 105
00185 Roma
T +39 06 49.210.425
F +39 06 49.59.302
email: digitalia@iccu.sbn.it
<http://digitalia.sbn.it>

In attesa di registrazione al Tribunale di Roma



Comitato Scientifico

Oswaldo Avallone
Giovanni Bergamin
Dimitri Brunetti
Simonetta Buttò
Rossella Caffo
Rosaria Campioni
Maria Carla Cavagnis Sotgiu
Laura Ciancio
Flavia Cristiano
Gianfranco Crupi
Andrea De Pasquale
Maria Cristina Di Martino
Pierluigi Feliciati
Marina Giannetto
Maria Guercio
Mauro Guerrini
Klaus Kempf
Patrizia Martini

Maurizio Messina
Maria Cristina Misiti
Maria Teresa Natale
Marco Paoli
Don Valerio Pennasso
Alberto Petrucciani
Massimo Pistacchi
Marco Pizzo
Paola Puglisi
Roberto Raieli
Gino Roncaglia
Maria Letizia Sebastiani
Giovanni Solimine
Laura Tallandini
Anna Maria Tamaro
Costantino Thanos
Antonella Trombone
Paul Gabriele Weston

SOMMARIO

giugno 2019

Il Portale delle biblioteche e degli istituti culturali italiani.

Presentazione del progetto

Roma, 11 aprile 2019, Sala Spadolini, MiBACT
di Paola Passarelli, Simonetta Buttò, Giovanni Solimine,
Claudio Leombroni, Alberto Petrucciani,
Gino Roncaglia, Marino Sinibaldi

9

SAGGI

Archivi digitali di persona

PAD - Pavia Archivi Digitali e gli archivi degli scrittori
di Paul Gabriele Weston, Primo Baldini,
Emmanuela Carbé, Laura Pusterla

31

**Through the Looking Glass. Cultural Heritage
Custodians to Populate the Mirrorworld**
di Susan Hazan

55

**I MOOCs, opportunità per la formazione di base
e l'apprendimento continuo: una storia (anche) italiana**
di Matilde Fontanin, Eleonora Pantò

76

PROGETTI

L'attuazione in Italia del Progetto GoogleBooks
di Andrea De Pasquale

103

L'emeroteca digitale dei giornali locali del Piemonte
di Dimitri Brunetti

114

**Urania digitale: il patrimonio storico scientifico
degli osservatori astronomici italiani
in Polvere di stelle e Internet Culturale**
di Antonella Gasperini, Emilia Olostro Cirella

126

**Linked Books: un indice citazionale
per la storia di Venezia** 132
di Giovanni Colavizza, Matteo Romanello, Andrea Giuliano,
Maria Cristina Mataloni, Daniela Grandin

**Phaidra, un archivio digitale FAIR
per la disseminazione e l'accesso
integrato a testi, testimonianze, immagini
e storie del patrimonio culturale** 147
di Laura Tallandini, Lorisa Andreoli, Elena Bianchi,
Linda Cappellato, Yuri Carrer, Gianluca Drago,
Giulio Turetta, Antonella Zane

Da un oggetto racconta la tua scuola 158
di Maria Teresa Natale

SEGNALAZIONI

La Fototeca Tifernate digitale On Line 165
di Alba Ghelli

Patrimonio culturale: reale e virtuale 170
di Maria Teresa Natale

Archivi digitali di persona

PAD – Pavia Archivi Digitali

e gli archivi degli scrittori

Paul Gabriele Weston - Primo Baldini - Università di Pavia

Emmanuela Carbé - Università di Siena

Laura Pusterla - Centro Manoscritti. Pavia Archivi Digitali - Università di Pavia

La diffusione di pratiche di scrittura elettronica e la conseguente istituzione di archivi letterari nativamente digitali ha determinato nel 2009 l'avvio del progetto PAD – Pavia Archivi Digitali, finalizzato a realizzare uno strumento per archiviare e rendere consultabili documenti in formato elettronico di personalità rappresentative della cultura e della società odierna, riconducibili a diverse generazioni e formazioni. Di questo materiale PAD intende garantire la conservazione nel tempo e favorire lo studio, nel rispetto della privacy e del diritto d'autore.

Preservare a lungo termine le memorie collettive e personali degli ultimi decenni è un'impresa resa particolarmente complessa dalla necessità di integrare competenze appartenenti ad ambiti molto diversi fra loro: discipline letterarie, tecniche e archivistiche, tecnologia dell'informazione, questioni giuridiche, aspetti amministrativi. 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.

Fattori di particolare criticità sono, inoltre, la stratificazione delle versioni, la tutela dei dati sensibili, la gestione dei diritti, le licenze d'uso degli applicativi software (particolarmente nel caso 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. 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 e così via), con l'intesa che abbiano anche la caratteristica della sostenibilità.

Prendendo come caso di studio le procedure di archiviazione del fondo del poeta e anglista Franco Buffoni, il cui completamento ha avuto luogo all'inizio dell'anno, l'articolo descrive le soluzioni informatiche ed archivistiche adottate nell'ambito di PAD, individuando al contempo le procedure da mettere a punto nelle ulteriori fasi di sviluppo del sistema.

1. Tempo di bilanci¹

Era il 2009 quando l'Università di Pavia ideava il progetto PAD – Pavia Archivi Digitali: per la prima volta in Italia l'attenzione non si limitava ai soli archivi cartacei degli scrittori, su cui Pavia grazie al suo Centro Manoscritti aveva già una lunga e consolidata tradizione, ma anche al digitale². Sono passati dieci anni e la sensazione di chi scrive è che, sebbene il dibattito sul nativo digitale sia oggi molto più ampio di allora, il tema declinato agli archivi letterari di persona non abbia ancora, almeno nel panorama nazionale, lo spazio che merita³.

Vale dunque la pena ripercorrere la storia di PAD per un primo bilancio del lavoro fin qui svolto, cercando di riflettere sulle opportunità future del progetto⁴. Fu Beppe Severgnini, allora presidente dell'Associazione Alumni dell'Università di Pavia, a lanciare nel 2009 l'idea di un archivio dedicato alle memorie digitali di scrittori, giornalisti e intellettuali contemporanei. PAD avrebbe dovuto raccogliere materiali digitali di diverso tipo, mettere a punto dei sistemi di conservazione a lungo termine ed eventualmente, in base alle disposizioni del soggetto produttore, consentire la consultazione ai fini di studio e ricerca. È su questa missione che venne avviata la prima fase di PAD, grazie a un nucleo di lavoro coordinato da Umberto Anselmi Tamburini insieme con Primo Baldini, tutt'oggi responsabile della parte informatica del progetto, e Annalisa Doneda, che si occupava dei rapporti con gli autori e le case editrici, dei conferimenti, e della comunicazione⁵.

¹ Il presente contributo è il frutto di un'elaborazione collettiva. La responsabilità del primo paragrafo è di Emmanuela Carbé, di Primo Baldini e Paul Gabriele Weston il secondo paragrafo, di Laura Pusterla il terzo, di Paul Gabriele Weston il quarto, nonché il coordinamento del lavoro.

² Per una panoramica degli esempi più rilevanti, anche nel contesto internazionale, si rimanda a Carbé 2018, p. 210-211. Sull'argomento vedi anche Weston — Carbé — Baldini 2017.

³ La sensazione di chi scrive, che andrebbe puntualmente verificata attraverso una sistematica mappatura, è che molti soggetti conservatori, che siano archivi, biblioteche o enti a vario titolo deputati alla conservazione di materiali novecenteschi e degli anni Zero, siano in possesso di materiale nativo digitale, magari custodito ancora su supporti obsoleti come floppy disk, e che non abbiano attivato strategie adeguate per la messa in sicurezza e preservazione a lungo termine. Pensiamo, ad esempio, ai vecchi personal computer rimasti nei depositi di dipartimenti universitari e appartenuti a scrittori/intellettuali che li hanno insegnato.

⁴ Per una storia del progetto si rimanda al contributo su PAD nel terzo volume, di prossima pubblicazione, di *Almum Studium Papiense. Storia dell'Università di Pavia*, a cura di Dario Mantovani, Milano: Cisalpino, 2019.

⁵ Negli anni si ampliò notevolmente il comitato scientifico e PAD riuscì a crescere grazie al lavoro di molti collaboratori. Per la composizione dell'attuale comitato scientifico, coordinato da ottobre 2014 da Paul Gabriele Weston, si rimanda al sito di PAD <www.pad.unipv.it>. Tra i tanti che si sono occupati di PAD si desidera qui almeno menzionare, per l'impegno essenziale nello sviluppo del progetto, Clelia Martignoni, Giorgio Panizza e Simone Albonico. Tra i collaboratori, Grazia Bruttoao è stata per anni responsabile della comunicazione esterna, e Nicoletta Leone dei rapporti tra PAD e Fondazione Corti, nonché responsabile delle attività di segreteria. Chi scrive ha collaborato da maggio 2013 a febbraio 2018; è poi entrata nel gruppo di lavoro Laura Pusterla, a cui si deve, tra le altre cose, la catalogazione del fondo digitale di Franco Buffoni, di cui si dirà oltre.

Si deve a loro, e alla loro perseveranza, la nascita del progetto, che venne avviato con un budget limitato, sufficiente per la realizzazione di un prototipo, inizialmente basato su due NAS indipendenti e due server⁶. I materiali venivano trasferiti nei due server, dopo essere stati oggetto di verifica della loro integrità e identificati sulla base di uno schema essenziale di metadati. Si affiancò a questa soluzione anche l'idea di masterizzare i fondi su DVD *archival grade* con strato metallico in oro, posti in buste sigillate a vuoto e depositati nel caveau della banca che sostiene il progetto. A partire dal quarto conferimento si decise di donare un DVD della stessa tipologia anche all'autore, con una copia dell'intero fondo conferito. Come si vedrà meglio oltre, il sistema di acquisizione e preservazione si è raffinato nel tempo, anche grazie ai problemi che di volta in volta si sono presentati con i nuovi conferimenti. Fu proprio Beppe Severgnini, il 16 ottobre 2010, a donare una copia del proprio archivio digitale proveniente da un Mac⁷. Venne messo a punto un primo contratto standard per il conferimento dei materiali⁸. Severgnini diede a PAD un archivio di 13.227 file, molti dei quali prodotti da Word Processor obsoleti e che all'epoca non potevano essere elaborati facilmente se non con sistemi di virtualizzazione. Ci si era trovati dunque di fronte ad una mole di dati quantitativamente più importante del previsto, e per un considerevole gruppo di file non si sapeva nemmeno il tipo di estensione. I successivi conferimenti, quello di Silvia Avallone e di Gianrico Carofiglio⁹, dimostrarono immediatamente che gli scenari di acquisizione e le tipologie di fondi potevano essere difforni tra loro e che era necessario un protocollo standard di acquisizione, soprattutto per la salvaguardia dei metadati e del contesto di provenienza, e per il controllo di ogni sequenza del processo. Fu quindi creato un software per il controllo della

⁶ In questa fase PAD ottenne, grazie a Beppe Severgnini, una piccola sponsorizzazione da parte di Nokia e Vodafone. UBI Banca aderì al progetto permettendo la custodia di una copia dei fondi presso una filiale di Pavia. Il dipartimento di Chimica-Fisica mise a disposizione dei server per avviare il progetto.

⁷ Il conferimento è avvenuto in presenza di Annalisa Doneda. L'operatore che ha effettuato il trasferimento dei dati era un informatico di fiducia di Beppe Severgnini. Il tecnico generò anche il file "Leggenda Beppe.pdf", contenente l'elenco dei file conferiti. Il documento arrivò a PAD all'interno della cartella principale contenente l'archivio conferito. Ne consegue che il file, pur facendo tecnicamente parte del materiale donato a PAD, è da considerarsi almeno da un punto di vista archivistico escluso dal conferimento, essendo stato generato esclusivamente ai fini di facilitare il processo di acquisizione.

⁸ Il contratto fu elaborato dall'avvocato Luigi Carlo Ubertazzi e subì qualche modifica nel tempo. Nella sostanza l'autore può indicare quali file e cartelle secretare per un certo numero di anni e dare indicazioni sulla fruizione dell'archivio (se ad esempio può essere consultato in remoto o solo nella sede PAD dell'Università di Pavia). L'autore può chiedere modifiche particolari al contratto, che vanno vagliate dal comitato di PAD ed eventualmente dall'ufficio legale dell'ateneo.

⁹ Silvia Avallone ha conferito il suo archivio il 26 luglio 2011 a Bologna, in presenza di Annalisa Doneda, presso le librerie COOP Ambasciatori, dando a PAD una chiavetta USB con materiali già selezionati (69 file). Gianrico Carofiglio fece inviare il materiale via e-mail il 25 maggio 2011 (6 file).

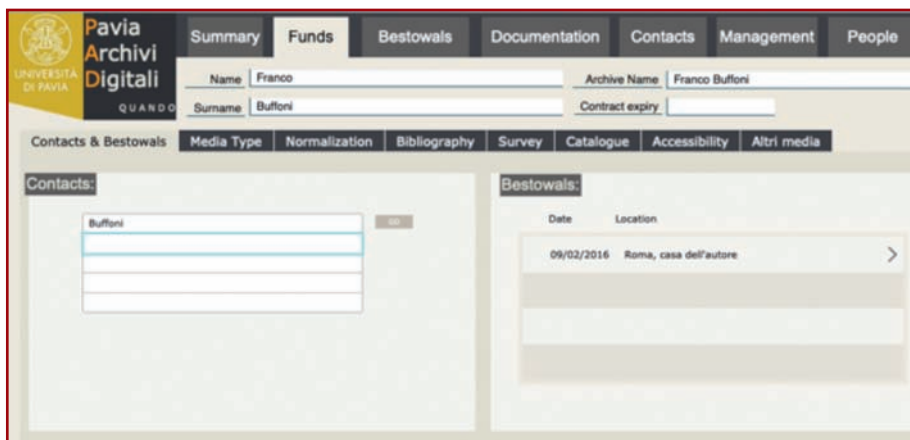


Figura 1. Il gestionale QUANDO

qualità e delle fasi operative (QUANDO)¹⁰, e fu realizzata una guida al conferimento da consegnare all'autore preliminarmente insieme con il contratto che regola l'acquisizione e la gestione del fondo.

Si decise inoltre, a partire dal 2015, di chiedere all'autore di compilare un questionario informativo per raccogliere dati importanti sul contesto di provenienza degli archivi¹¹.

È con questo nuovo assetto che vennero realizzati, a partire dal 2015, i nuovi conferimenti¹². Il primo ad aderire fu Paolo Di Paolo, seguì poco dopo Francesco Pecoraro¹³. Se la mole di file conferito del primo era tutto sommato gestibile (2.982 file), l'arrivo a Pavia del fondo di Pecoraro pose PAD di fronte al caso in assoluto più complesso¹⁴ e tuttora in fase di trattamento: l'archivio consta infatti di 43.718 file, comprende un salvataggio di livello molto profondo del blog non più online di Pecoraro (ma anche pagine collegate al blog, che sono state salvate contestualmente), diversi backup di pc dismessi, e un gruppo di file da secretare su indicazione dell'autore. A questi materiali vanno inoltre aggiunti dei documenti estratti da floppy disk dell'autore, di cui PAD conserva solo il contenuto¹⁵.

¹⁰ Cfr. Weston – Carbé – Baldini 2016.

¹¹ Cfr. Weston – Carbé – Baldini 2017. Si segnala inoltre Feliciati et al. 2018.

¹² Si fece per prima cosa una simulazione completa di acquisizione d'archivio con dei materiali di Emmanuela Carbé, compresi i suoi account sui social network Facebook e Twitter, per capire quali problemi si potevano prevedere di fronte ad acquisizioni di diversa provenienza.

¹³ Il conferimento di Paolo Di Paolo avvenne il 13 gennaio 2015 alla Casa delle Letterature di Roma, con Paul Gabriele Weston e Emmanuela Carbé. Vennero acquisiti materiali provenienti da un MacBook Pro e da un hard disk esterno dell'autore. Il conferimento di Francesco Pecoraro si tenne invece a casa dell'autore il 21 e 22 marzo 2015, con materiali provenienti dal PC dell'autore, da Dropbox e da un hard disk esterno.

¹⁴ Carbé 2017.

¹⁵ Si è trattato del primo recupero di dati da supporti obsoleti. L'autore ha dato in custodia a PAD 35

Venne dunque sviluppato da Primo Baldini, e testato proprio sul fondo Pecoraro, il software PadManager per la catalogazione e la consultazione dei fondi¹⁶.

Risale infine al 2016 il conferimento di Franco Buffoni, che già dal 2011 aveva iniziato a donare a più riprese materiale cartaceo presso il Centro Manoscritti¹⁷, e che concordò con PAD il salvataggio periodico del suo sito web personale. L'archivio Franco Buffoni, come si vedrà oltre, è stata un'acquisizione particolarmente importante anche per ragioni archivistiche, essendo il primo fondo ibrido dell'Università di Pavia¹⁸.

Accanto alla risoluzione di problemi tecnici e archivistici, che sono il cuore della missione di PAD, sono state numerose le iniziative attivate nel corso degli anni: tra queste, ricordiamo i seminari tecnici e letterari che hanno coinvolto esperti di *digital humanities* e gli autori nelle aule universitarie, in libreria e in occasione di manifestazioni letterarie, a dimostrazione che un progetto scientifico, se modulato opportunamente, può diventare un ricco laboratorio interdisciplinare aperto ai giovani studenti e alla cittadinanza¹⁹.

floppy disk (3.5 pollici), che sono stati restituiti dopo aver effettuato l'estrazione dei materiali e una descrizione preliminare dei floppy e delle etichette. Dieci floppy non erano più accessibili, per tutti gli altri sono state attivate procedure di conversione dei file, anche tramite virtualizzazione, e di isolamento dei virus. A marzo 2015 sono stati estratti materiali da tre dischetti (3.5 pollici) appartenuti a Maria Corti e conservati presso la Fondazione Corti di Pavia. Nell'aprile 2019 PAD ha salvato materiali da 58 floppy disk (3.5 pollici) del fondo Jolanda Insana conservato al Centro Manoscritti e sta programmando il salvataggio di sette floppy conservati a Pavia nel fondo cartaceo di Franco Buffoni.

¹⁶ La prima applicazione, realizzata nel 2015 in JavaFx, si chiamava PadExplorer. La sua evoluzione, PadManager, è un software costruito nel 2016 e implementato negli anni successivi.

¹⁷ Il conferimento avvenne il 9 febbraio 2016 presso la casa dell'autore, in presenza di Paul Gabriele Weston e Emmanuel Carbé. I file, tutti provenienti da uno dei Mac dell'autore, vennero trasferiti su una chiavetta a causa di un malfunzionamento dell'hard disk esterno criptato. Per una descrizione dell'archivio cartaceo, a cura di Federico Milone, si rimanda al catalogo online del Centro Manoscritti:

<https://lombardiarchivi.servizirl.it/groups/UniPV_CentroManoscritti/fonds/47480>.

¹⁸ Vale la pena qui sottolineare la grande sensibilità di Buffoni alle tematiche della preservazione e del digitale. Fu anche grazie a lui che nel 2003 venne istituito presso il dipartimento di italianistica dell'Università di Cassino un archivio del giovane autore ispirato al Centro manoscritti di Pavia. L'archivio preservava il materiale, anche nativo digitale, di una ventina di scrittori. Così Giuseppe Antonelli commentava l'iniziativa: «potremmo chiamarla “filologia del presente”. Gli avantesti degli scrittori contemporanei conservati presso il Centro di ricerca su tradizione e traduzione dell'Università di Cassino (quaderni, taccuini, dattiloscritti, stampe da pc, bozze, file su supporto informatico) consentono di affrontare il sedimentarsi di un testo letterario concentrando l'attenzione non sulle opere consacrate dai tempi lunghi del canone, ma sull'altrieri di scrittori che oggi hanno fra i trenta e quarant'anni» (Antonelli 2010).

¹⁹ Tra le iniziative, che si devono soprattutto a Paul Gabriele Weston e Clelia Martignoni, segnaliamo il seminario tecnico “Archiviare il digitale”, con la partecipazione di Fabio Ciotti e Salvatore Vassallo (14 marzo 2016); e due seminari su “Letteratura e web”, con Gilda Policastro, Francesco Pecoraro, Paolo Costa e Alessandro Gazoia (21 marzo e 10 maggio 2016). PAD ha inoltre organizzato un incontro con Franco Buffoni e Francesco Pecoraro a “Tempo di libri” (Milano, edizione 2018) moderati da Gino Roncaglia. I video con le interviste agli autori sono disponibili all'indirizzo

2. L'infrastruttura per l'archiviazione e il trattamento dei file²⁰

L'esigenza di assicurare l'archiviazione a medio-lungo termine di archivi letterari digitali nativi ha richiesto che, nella fase iniziale tra il 2010 e il 2011, un impegno particolare venisse dedicato allo studio di fattibilità dell'infrastruttura informatica destinata ad accogliere i dati e rendere possibili tutte le procedure occorrenti al trattamento dei file. Lo studio ha condotto a privilegiare le seguenti necessità: una architettura ad un tempo semplice e scalabile per consentire ampliamenti successivi conseguenti alla crescita dei dati, più che ad esigenze strutturali, e l'implementazione di nuovi servizi; l'archiviazione e il trattamento veloce di grandi quantità di dati, considerando che anche archivi di modeste dimensioni possono essere costituiti da migliaia di file; l'effettuazione dell'archiviazione anche in assenza di quei metadati identificativi che quasi mai i produttori dei materiali si preoccupano di creare secondo le buone pratiche richieste dalla conservazione; la conversione tra formati dei documenti condotta in modalità non interattiva; l'effettuazione in forma ripetuta di procedure volte ad estrarre diverse tipologie di metadati dai dati esistenti; la visualizzazione della struttura dei *file-system* dell'archivio conferito, che riflette in qualche modo la struttura che l'autore ha inteso dare al proprio archivio e che può fornire notizie essenziali sulle pratiche di scrittura del medesimo; l'applicazione di restrizioni all'accesso ai dati, modulate in base agli accordi contrattuali con il conferente; la gestione di occorrenze multiple di un medesimo file (ossia con lo stesso *checksum*) all'interno di cartelle diverse, che occorre documentare per garantire la struttura originaria dell'archivio, ma rendere al contempo accessibile allo studioso in modalità tale da non confonderlo con l'effetto eco; la gestione di conferimenti successivi di porzioni diverse dell'archivio o dell'archivio nella sua totalità, ma in stadi di evoluzione diversi.

Le principali soluzioni commerciali e open source sono state quindi analizzate alla luce dei requisiti così individuati e alcune aziende sono state interpellate circa la possibilità di condividere la responsabilità dello sviluppo di un prodotto confacente. Nessuno dei sistemi presi in esame è parso rispondere adeguatamente a tutti o alla maggior parte dei requisiti, non soltanto dal punto di vista dei fattori tecnici (metadati, standard, tracciati, strumenti di controllo, ecc.), ma anche da quello della sostenibilità economica del progetto nel corso del tempo.

<pad.unipv.it/media>. Altre iniziative sono state concertate insieme con la libreria Il Delfino di Pavia. Il software PadManager divenne inoltre un importante strumento didattico per i laboratori del corso di Biblioteche digitali tenuti da Paul Gabriele Weston (a.a. 2015/16 e 2016/17); in quell'occasione gli studenti ebbero modo di effettuare una descrizione di alto livello del fondo di Franco Buffoni. Gli studenti vennero anche coinvolti in incontri sulla preservazione a lungo termine del digitale. Segnaliamo infine il ricco lavoro di ricognizione di Michele Stefani nella sua tesi di laurea *Gli archivi degli scrittori nativi digitali: verso la conservazione e oltre* (Stefani 2017).

²⁰ L'argomento meriterebbe una trattazione più ampia di quanto non consenta lo spazio di questo saggio. Ci si limita in questa sede a qualche indicazione metodologica, in relazione ai requisiti dei quali si è tenuto conto nella realizzazione dell'infrastruttura informatica.

Si è stabilito, pertanto, di procedere allo sviluppo *in house* del software realizzato interamente in Java, per quanto riguarda l'architettura, e con FileMaker per le applicazioni occorrenti alla consultazione. Ci si sarebbe potuti attendere una soluzione OS. La scelta è stata motivata da una serie di fattori: la disponibilità nel gruppo di lavoro delle competenze richieste per uno sviluppo tanto complesso; la gestione in automatico dell'interoperabilità tra sistemi; l'effettuazione in tempo reale di modifiche richieste dalla progressiva messa a punto dello strumento; infine, i costi davvero modesti per le licenze. D'altro canto, la soluzione OS avrebbe imposto tempi lunghi, non sempre prevedibili, per lo sviluppo e il funzionamento del sistema di archiviazione. Aver privilegiato in questa prima fase una soluzione proprietaria non preclude la possibilità che, a regime, se ne possa implementare una OS verso la quale i dati potrebbero facilmente migrare.

L'architettura degli archivi di PAD prevede sette aree ciascuna delle quali dedicata all'esecuzione di una fase di lavorazione, dal riversamento massivo dell'insieme dei file conferiti dall'autore, all'esecuzione delle procedure di controllo e di conversione tra formati, dallo svolgimento di routine per l'estrazione dei metadati, alla strutturazione, organizzazione e descrizione dell'archivio, dall'archiviazione definitiva all'erogazione di servizi di consultazione per le parti autorizzate del fondo.

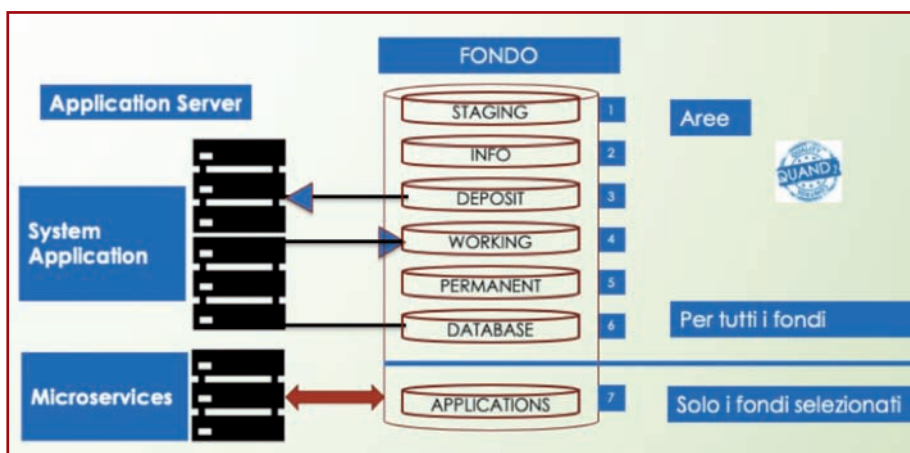


Figura 2. Le aree in cui si articola la struttura del sistema di archiviazione

Per ciascun conferimento si crea un nuovo fondo, il che non esclude la possibilità che documenti in relazione tra di loro localizzati in fondi diversi possano essere collegati e resi accessibili contestualmente.

Nella progettazione di PAD si è fatto riferimento a diverse tecnologie, le cui caratteristiche precipue sono state coniugate secondo le esigenze del sistema da realizzare. In particolare, ai criteri dell'informatica forense si è fatto ricorso nel definire le procedure utilizzate per l'identificazione e l'analisi dei dati e specificamente di grandi volumi di

dati non strutturati²¹. Big Data, Digital Asset Management (per organizzare, descrivere, archiviare, organizzare e dare accesso a risorse digitali collegate a modelli di meta-dati) e *repository* istituzionali sono stati gli altri ambiti ai quali si è fatto riferimento.

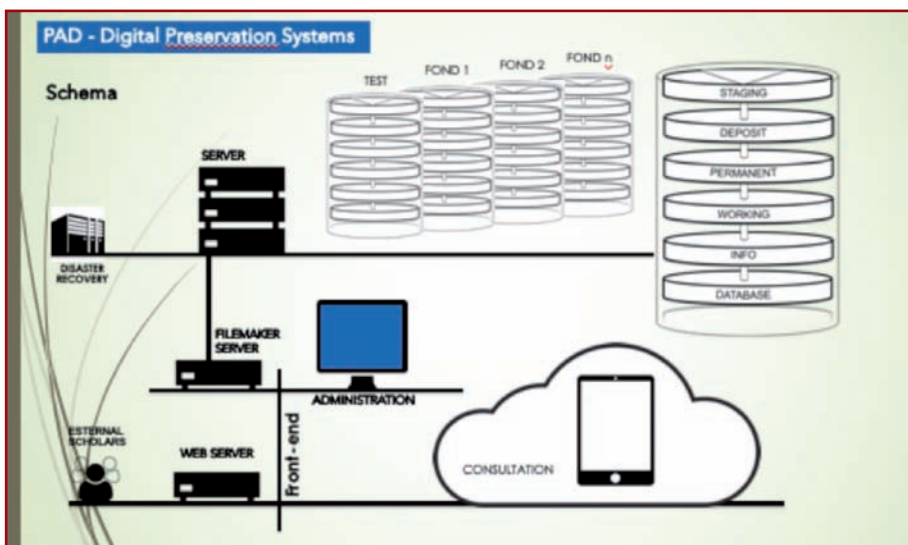


Figura 3. Le componenti del sistema di archiviazione e di accesso

Di queste considerazioni ci si è serviti per progettare un sistema di preservazione di utilizzo molto semplice (ad esempio, nella fase di *ingest* l'unico dato richiesto è il codice di identificazione del conferimento), a prova di errore (a questo scopo sono state definite aree diverse e si è stabilito di tenere separati i database) e orientato a gestire funzionalità innovative (*Data Mining*, *Machine Learning*, ecc.). Evidentemente si tratta di una struttura ideata su misura per le finalità di PAD e, di conseguenza, mal si presterebbe a servire in modo egualmente efficace in altri contesti.

3. La catalogazione del Fondo Buffoni

Prima fase: assegnazione delle categorie di rischio

Per quanto riguarda la catalogazione, il fondo Buffoni ha costituito per PAD un banco di prova di nuove procedure e tecnologie. Le esigenze e le caratteristiche dell'ambiente digitale, affatto diverse da quelle degli archivi tradizionali, hanno imposto lo studio di strategie che ne valorizzassero le specificità. Una tra le prime questioni affrontate è stata quella relativa all'accesso ai dati. La vicinanza cronologica dei materiali ha richiesto di verificare scrupolosamente che la loro pubblica-

²¹ Joe Sremack, *Big Data Forensics: Learning Hadoop Investigations*, Birmingham, UK-Mumbai, India: Packt Publishing, 2015.

zione non costituisca una violazione delle normative sulla privacy o sulla proprietà intellettuale. Copyright, contenuti o dati sensibili, corrispondenza e-mail, file privati sono tra le criticità che si presentano con maggiore frequenza. L'elevato numero di file che, in genere, costituisce l'archivio di un autore aumenta il rischio che tra di essi ve ne siano alcuni a carattere personale o la cui proprietà intellettuale appartenga ad altro soggetto. Per questa ragione, una volta fatto l'ingest dell'archivio, un elenco dettagliato dei documenti viene trasmesso al conferente. Ciò nonostante, l'archivista deve operare con grande cautela. Per distrazione o in buona fede, la loro presenza potrebbe passare inosservata. Perciò, durante la fase della catalogazione, ogni file viene vagliato per escludere che possa contenere dati sensibili o di proprietà intellettuale altrui. Talvolta è l'autore stesso a chiedere, per ragioni personali, l'esclusione dalla consultazione di particolari oggetti digitali, fino a una data concordata.

Per facilitare le operazioni di catalogazione e consultazione, ad ogni documento viene assegnata una categoria di rischio, che determina la possibilità che esso sia consultabile dallo studioso. Ad oggi, le categorie stabilite sono cinque:

1. Archiviare: file conferiti dall'autore, di sua creazione intellettuale, non contenenti dati sensibili e per i quali non esista esplicita richiesta di secretazione. In questa categoria vengono inserite anche le opere di altri autori, ma già edite (ad es. articoli di giornali, saggi, poesie). I documenti assegnati a questa categoria sono accessibili liberamente per ragioni di studio o didattiche.
2. Secretare: file che l'autore ha esplicitamente sottoposto ad embargo.
3. Sensibile: documenti per ora non accessibili all'utenza, in quanto violano la vigente legislazione in materia di dati personali o di diritto d'autore. La categoria include documenti sensibili, produzioni intellettuali altrui inedite o di cui non si possiede esplicito consenso alla consultazione, corrispondenza, imma-

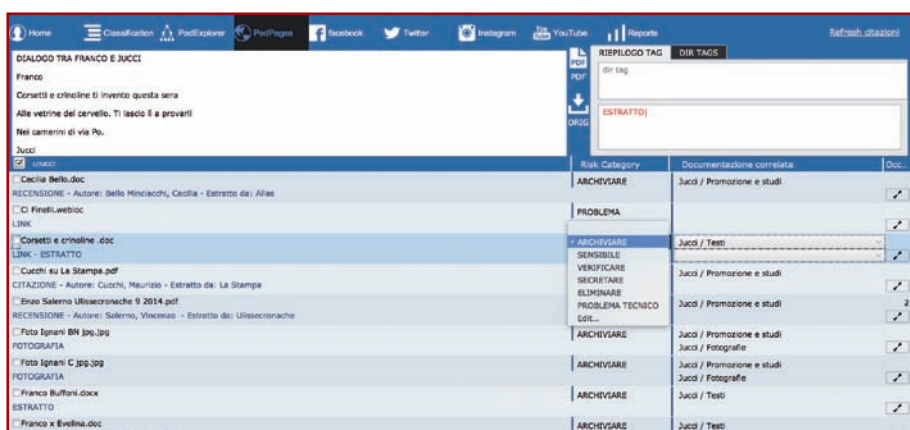


Figura 4. Attraverso il menu a tendina è possibile scegliere tra le categorie di rischio già preinserite. La scelta è facilitata dalla funzione di preview

- gini, video o audio inediti in cui compaiano persone diverse dal conferente. I tempi per il libero accesso a queste risorse sono molto lunghi, ma PAD si impegna alla loro conservazione per il tempo stabilito in termini di legge.
4. Disponibile: comprende file che sono accessibili, nonostante le restrizioni per copyright e dati sensibili, in presenza della liberatoria dei soggetti coinvolti (ad es. corrispondenza, fotografie, nonché produzione intellettuale inedita, per cui sia stato rilasciato esplicito consenso alla consultazione).
 5. Eliminare: file creati autonomamente dal software e dai sistemi operativi che sono ospitati nell'archivio all'insaputa del conferente.

L'assegnazione di queste categorie a ogni singolo documento è un'operazione fondamentale, da eseguire con estrema attenzione. Solo una volta stabilite queste sarà possibile concedere la consultazione dei file.

Seconda fase: creazione dell'albero

Questa fase mostra come le potenzialità dell'informatica contribuiscano a fornire nuove opportunità ad archivisti e studiosi. Come nel caso del riordino degli archivi cartacei, si è creata una struttura ad albero rovesciato che comprende le diverse sezioni alle quali vengono poi assegnati i file.

L'albero del fondo di Franco Buffoni è composto da 34 livelli alti, denominati Produzioni. Si è stabilito che a ogni Produzione corrispondesse una singola opera. In aggiunta, sono stati creati alcuni livelli Produzione generici per raggruppare materiale vario, di produzione di Buffoni medesimo (ad es. saggi o articoli di giornale), o di soggetti diversi (ad es. studi sulla sua produzione generale).

Alcuni di questi livelli alti comprendono a loro volta livelli intermedi, denominati Progetto. Il livello Progetto è stato utilizzato nel caso di materiale facente riferimento alla stessa opera dell'autore, ma in lingue o edizioni differenti, oppure per aggregare articoli pubblicati su una stessa rivista. In questo modo l'utilizzatore del catalogo del fondo Buffoni potrà trovare il materiale posseduto da PAD relativo alla sua opera *Con il testo a fronte* già suddiviso tra le due edizioni della medesima o trovare accanto alla raccolta dei testi di *Guerra* la versione della stessa in francese. I Progetti si basano, nella maggior parte dei casi, sulle opere editate dell'autore. Ciò spiega perché, per ogni Progetto, siano state introdotte categorie di pubblicazioni collegate a schede. Queste ultime riportano informazioni specifiche per le diverse tipologie, che possono essere di aiuto nel reperire agevolmente la fonte edita.

L'ultimo livello dell'albero è quello della Documentazione correlata, ossia macrocategorie che riuniscono il materiale su un'opera distinguendolo per tipologie. Durante la fase di catalogazione il materiale viene reso identificabile mediante l'assegnazione a una di queste categorie. Le tipologie più comuni tra queste sono ovviamente quelle che comprendono i testi, ma anche le copertine, il materiale promozionale e le recensioni. È possibile trovare anche fotografie, traduzioni, audio o video.

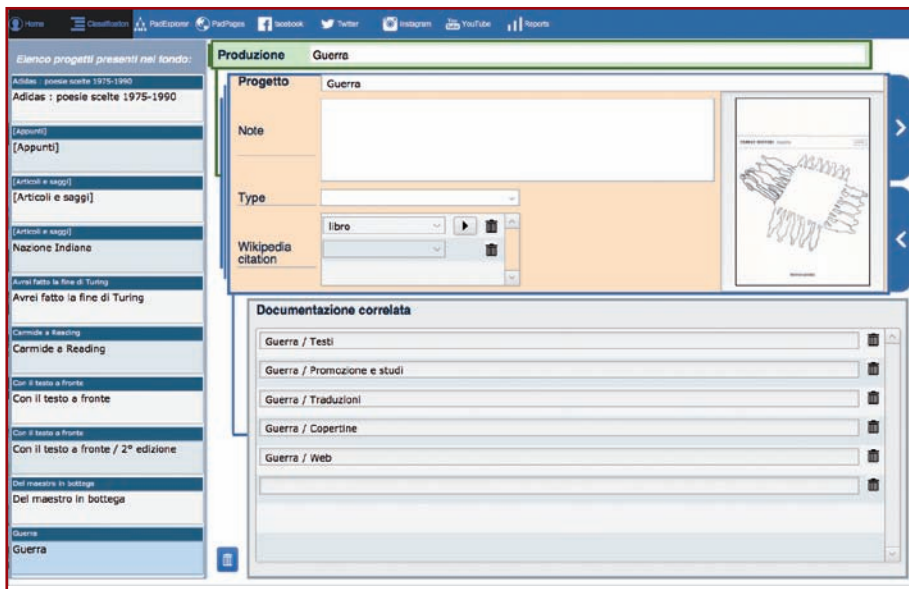


Figura 5. Albero a livelli della produzione documentale dell'autore. Ogni documento del fondo è inserito tra una particolare Documentazione correlata, legata alla Produzione e al Progetto. Nella parte a sinistra è visualizzato l'elenco sintetico dei progetti presenti nel fondo

L'albero prodotto può essere modificato in ogni momento e ogni documento riassegnato a una diversa Documentazione correlata. La possibilità di assegnare un singolo file a più di uno di questi raggruppamenti rappresenta un elemento di novità e un grande vantaggio rispetto a quanto avviene in ambito cartaceo. Se in quest'ultimo il riordino archivistico comporta la modifica dell'ordine (o disordine) esistente per uno più funzionale alla conservazione e alla ricerca, nel digitale ordine originale e riordino possono coesistere. Trattandosi di archivi letterari è possibile che uno studioso sia anche interessato a conoscere il modo in cui un autore abbia ordinato i propri documenti, l'ordine da lui seguito, quanto possa essere stato preciso o eclettico. Ovviamente questo non può andare a discapito della praticità nel collocare o trovare un documento. Facendo coesistere le due visioni, i due approcci, lo strumento informatico - PadManager - rende più agevole il lavoro sia dell'archivista che dello studioso.

Terza fase: catalogazione o descrizione dei file

La procedura richiede a questo punto che ogni singolo file venga esaminato ai fini della sua identificazione e assegnazione a un Progetto. Il modulo di catalogazione di PAD mostra l'intero archivio dell'autore, suddiviso in cartelle o sottocartelle, nello stesso ordine attribuitogli dal suo creatore. Grazie alla funzione di *preview* è possibile visualizzare immediatamente il testo contenuto nel documento. Una vol-

ta che l'archivista abbia individuato l'opera, la tipologia di materiale e le responsabilità, l'interfaccia permette l'inserimento di queste informazioni mediante tabelle appositamente predisposte.

Per consentire di individuare e comprendere la natura del documento, l'archivista si avvale di tag costituiti da parole chiave, che vengono aggiunti secondo le esigenze dell'archivio che sta descrivendo e sono facilmente modificabili. Essi possono anche essere combinati tra di loro. Nel caso del fondo Buffoni, per identificare specifiche categorie di documenti, sono stati usati principalmente tag relativi al contenuto testuale del file quali:

- bozza o bozza di stampa (testo a opera dell'autore, non ancora predisposto o con prove di impaginazione per la stampa);
- appunti o commento (abbozzi su un testo o idee generali per altri scritti, oppure critica del testo opera di altri);
- lettera (ogni forma di corrispondenza personale, inviata o ricevuta);
- articolo (soltanto articoli scritti dall'autore stesso);
- poesia (testi di poesie singole, edite o inedite);
- estratto (parte tratta da un'opera più ampia in presenza di particolari finalità o per tematica);
- inedito (qualsiasi testo non ancora pubblicato);
- introduzione o premessa (testo scritto da qualcuno che non sia il produttore del fondo).

Tenendo conto che in un archivio personale il materiale è molto spesso prodotto da soggetti diversi dal conferente, sono stati usati tag diversi per identificare questi file:

- recensione (testo a opera di altro autore di recensione e commento al testo di Franco Buffoni; viene indicato, dove possibile, chi l'ha scritta e, se edita, dove è stata pubblicata, sia una testata sia un sito web);
- presentazione (materiale informativo relativo a una presentazione di un libro o di un'opera dell'autore, oppure per un testo preparatorio per l'intervento, anche redatto da qualcuno che non sia lo scrittore, come nel caso di articoli di giornale che lo pubblicizzano);
- locandina (immagine pubblicitaria o promozionale);
- intervista;
- convegno (atti di un convegno o un singolo intervento sull'autore o al quale abbia partecipato);
- traduzione (qualunque testo tradotto in una lingua differente da quella originale, sia che si tratti di un testo dell'autore tradotto, sia che il traduttore sia l'autore stesso e in questo secondo caso viene indicato, se possibile, il nome di chi ha composto il testo in lingua originale, mentre nel primo si può collegare il nome del traduttore);

- testo originale (testo successivamente tradotto, elaborato o commentato dall'autore);
- citazione (testi di terzi, nei quali l'autore o una sua opera vengano citati);
- tesi (tesi di laurea o ricerca, oppure studio preliminare ad un articolo o saggio in cui vengano trattate una singola opera o l'intera produzione dell'autore);
- contratto (copia del contratto stipulato dall'autore con l'editore);
- biografia;
- sinossi (testo di presentazione del libro);
- diario (trascrizioni di memorie personali).

Negli archivi d'autore non mancano quasi mai documenti che riguardano confezione ed impaginazione del volume stampato, nonché la coperta. Per consentire l'identificazione di questi materiali sono stati utilizzati copertina (immagini, bozze o prove di stampa per la copertina di un volume) e quarta di coperta (componente testuale scritta appositamente o estratta dall'interno della pubblicazione).

Infine, si è reso necessario qualificare talune tipologie di file diverse da quelle testuali e presenti in numero minore, che rappresentano tuttavia una parte significativa dell'archivio, quali: video; link; fotografia o immagine; audio o musica (quest'ultimo soltanto per i contenuti musicali).

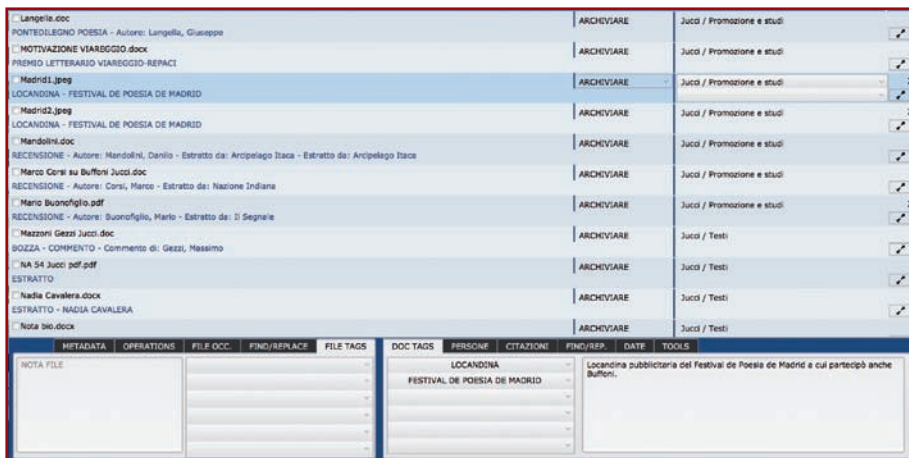


Figura 6. Nella parte bassa dello schermo sono presenti dei menu a tendina da cui è possibile selezionare dei tag preinseriti oppure inserirne di nuovi. Inoltre ogni file può essere corredato da note a campo libero

I suddetti tag, definiti per le esigenze del fondo Buffoni, sono ovviamente utilizzabili anche per altri archivi ed è possibile implementarne di altri in relazione a caratteristiche differenti di questi ultimi.

Poiché recensioni, traduzioni o studi, immagini delle copertine ed altri materiali sono in genere spesso prodotti da terze persone, vi è la possibilità di collegare a

ogni file i nomi di coloro che hanno a diverso titolo contribuito alla sua realizzazione con l'indicazione della natura della responsabilità (ad es. prefatore, traduttore, illustratore, fotografo).

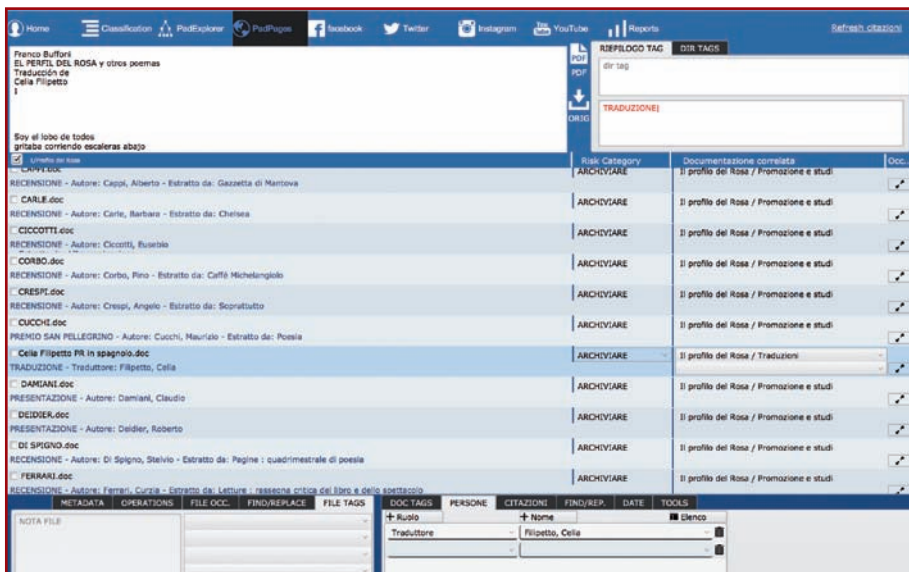


Figura 7. La scheda *Persone* permette di legare al file il nome di una persona diversa dall'autore principale, specificandone il ruolo

Siti web e social network

Il progetto, che inizialmente si è focalizzato sulla conservazione a medio-lungo termine degli archivi digitali in remoto, tenendo conto del crescente uso che della rete fanno gli autori, ha ritenuto di sviluppare un sistema per la salvaguardia dei siti web e delle pagine dei social media. L'intento di PAD non è competere con analoghi progetti internazionali, ma nel rispetto della sostenibilità offrire un servizio di qualità. Le risorse culturali che si trovano in rete sono, per una molteplicità di fattori, tra le più esposte al rischio di perdita dei dati. I malfunzionamenti della rete, degli *storage* o dei software, i problemi tecnici interni, nonché gli attacchi degli *hacker*, il *malware*, l'*account hijacking* sono solo alcuni tra i pericoli a cui le preziose informazioni dell'archivio sono esposte. Altri fattori di rischio sono le riorganizzazioni interne dei siti, che potrebbero modificare la struttura delle pagine, o, all'opposto, la scarsa manutenzione della risorsa. Occorre tener conto del fatto che una crescente percentuale delle testimonianze culturali della nostra epoca sia ormai presente solo sul web, ciò che rende la loro preservazione tanto fondamentale quanto quella di biblioteche o archivi cartacei e digitali²².

²² Julien Masanès, *Web archiving: issues and methods*, Berlin - Heidelberg: Springer, 2006.

Il sito web di Franco Buffoni²³ ha fatto da banco di prova per la messa a punto delle procedure di trasferimento in locale e di salvaguardia. Con il consenso dell'autore, il sito è stato copiato sul server di PAD. Poiché i siti web sono assai dinamici e vengono modificati ed aggiornati frequentemente, è stata concordata l'effettuazione di salvataggi a cadenza prestabilita. Ogni versione successiva è così messa a disposizione dello studioso che la può liberamente navigare. Attraverso un software per il web scraping, il sito dell'autore viene riprodotto in locale, in modo da garantire il browsing offline. Talvolta una versione del sito, specie se non più disponibile nel web, è già stata salvata dallo scrittore e pertanto fa parte del suo conferimento.

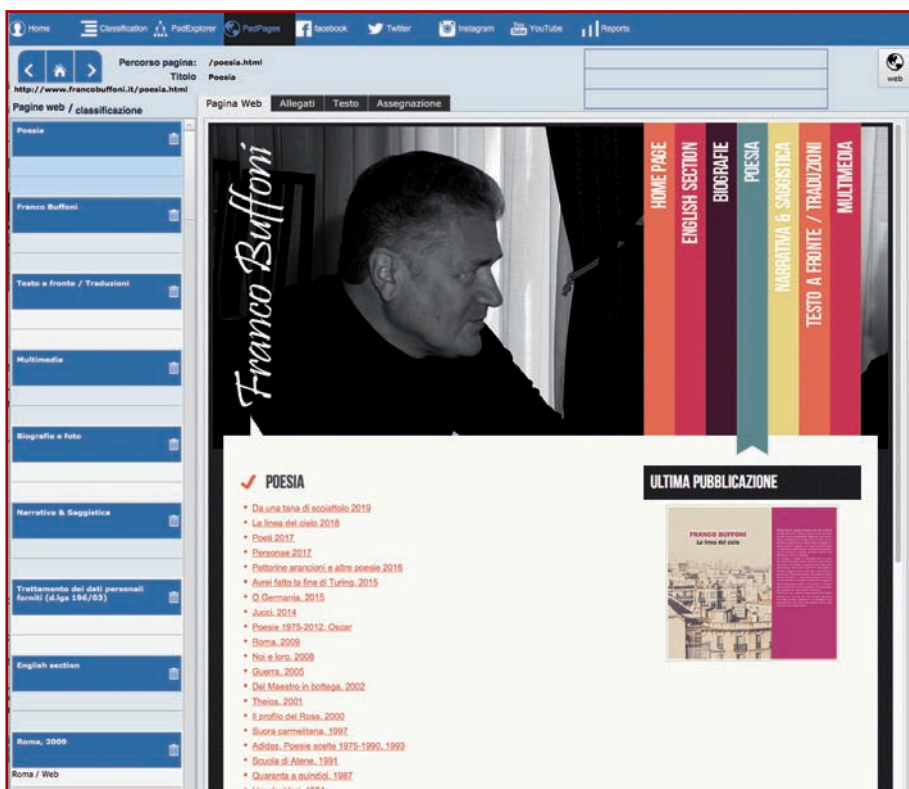


Figura 8. Navigazione offline del backup del sito web

I siti presentano talvolta strutture molto complesse, con numerosi riferimenti ad altre pagine, interne o esterne nel web, oppure contenenti documenti di varia natura (testo, immagini, video o audio). Per questa ragione di ogni pagina che compone il sito, PAD memorizza, oltre alla pagina stessa, i link alle pagine esterne, con

²³ www.francobuffoni.it.

un'immagine della pagina a cui il link porta, nonché i documenti allegati. In questo modo si può tenere meglio traccia dei percorsi - *path* - che il creatore del sito ha voluto valorizzare. Qualora un link ad una pagina esterna non fosse più funzionante, si perderebbe una parte di ciò che l'autore voleva comunicare. I tempi di questa operazione sono piuttosto lunghi, a causa della quantità di dati da salvare. Una volta acquisito tutto il materiale, viene generata un'immagine disco, che viene salvata nell'area di conservazione permanente del sistema PAD. Questi dati vengono trattati alla stregua di quelli degli archivi nativi digitali e sono sottoposti alle medesime procedure messe in atto per garantire la *disaster recovery*. Il sito web di Buffoni è ricco di materiali di varia tipologia. Per il momento gli interventi hanno riguardato soltanto le risorse testuali, che ne costituiscono la parte preponderante. Descrivendo i file ci si è resi conto che spesso l'autore conservava testi che poi venivano riversati sul web, come file allegati alla pagina o come link.

File	Allegati individuali
Roma, 2009 Pordenonelegge-Oedalus files/pdf/oedalus_class_1208.pdf	Allegati individuali: Il Lago Riprendo tutto a casa Einaudi 52 voti L. Pugno Quanto vorrà minimum day
Valerio Magrelli, prefazione a ROMA files/pdf/roma_valerio_magrelli_prefazione.pdf	Valerio Magrelli, prefazione a ROMA L'ultima raccolta di Franco Buffoni (nato a Gallarate nel 1948 e considerato fra i maggiori poeti italiani delle sue generazioni) nasce da uno spostamento geografico e antropologico, lo spostamento con cui questo "vecchio lombardo assente" narra la sua scoperta di Roma. Al pari di Penna, sceso dal cuore dell'Umbria, al pari di Pasolini, calato da Bologna e dai Friuli, lo scrittore racconta del trapianto in una capitale insieme classico e barocca, pagana e islamizzata, cristiana e consumista. Proprio seguendo quei suoi due discorsi, egli dedicherà alla Città eterna pagine che saranno dell'incisa scrittura allo sfruttamento lavorativo della riproduzione dell'opuscolo
L'anticipazione apparsa su Nuovi Argomenti 43, luglio-settembre 2008 files/pdf/roma_anticipazione_nuovi_argomenti.pdf	ROMA - L'ANTICIPAZIONE APPARSA SU NUOVI ARGOMENTI 43, LUGLIO-SETTEMBRE 2008 Nel gioco decorativo la battaglia Dicesse dagli arazzi Al ramo di natività dei reattori
L'anticipazione su Il Parlar franco con testo di M. Cohen files/pdf/manual_cohen_su_roma.pdf	FRANCO BUFFONI IL CIELO CHE SI MUOVE SOPRA ROMA. di Manuel Cohen, Il Parlar franco, gennaio 2009
L'anticipazione su Metromorfosi con testo di F. Matteoni files/pdf/matteoni_su_roma.pdf	Su ROMA di Franco Buffoni Vivere in una grande, antica città significa spesso attraversare le epoche. Immaginate gli edifici silenziosi nella prima luci dell'alba o chiusi sui loro temporanei, sconosciuti abitanti. C'è un tempo per le strade le voci che ascoltate escono direttamente dalle mura, dai frangi e perfino dal sottosuolo dove dormono altre città, con lo stesso nome della vostra. Nei testi che compongono Roma il poeta Franco Buffoni si aggira per la città eterna, dove l'eterno più che un tempo senza fine sembra indicare la coincidenza delle ere, l'irrimediabile di ciò che è stato, che si trasforma in aspetti spesso contraddittori. Così, dieci secoli antichissimi coesistono in molte Rome, una ad una, sovrapposte le une alle altre.
L'anticipazione nella traduzione in spagnolo di Caterina Castellary files/pdf/castellary.pdf	Franco Buffoni Il mio nuovo libro di poesia di si intitola ROMA. Lombardo quale sono, dopo dieci anni di permanenza nella capitale, ho sentito la necessità di mettere ordine in un materiale poetico che ormai andava ingessandosi. Come il lavoro procedeva, mi resi conto che funzionava il ritratto di una città antica e attuale, dove alcuni personaggi approdati nella Roma di ieri - da Pasolini a Galilei al Ritrucchio, da Leopardi a Sandro Penna a Keats - intraggano anche con le contraddizioni

Figura 9. Elenco dei file caricati sul sito, con l'indirizzo web, la funzione di preview e un suggerimento di catalogazione automatica

In altri casi, i testi della pagina web sono collegati a quelli dell'archivio a motivo di una comune ispirazione. Estrahendo i metadati e poi considerando quelli più importanti per identificare in modo puntuale la risorsa web, si possono cercare correlazioni tra nativo digitale e web. I processi di questa operazione non sono molto differenti da quelli che vengono normalmente effettuati per l'elaborazione dei conferimenti di PAD.

In primo luogo viene creata una mappa del sito, vengono estratti i metadati, che sono poi ridotti con l'eliminazione di quelli di scarsa utilità, e infine i documenti vengono sottoposti a procedure di normalizzazione. Conclusa questa fase, i documenti del web sono pronti per essere descritti.

La descrizione si avvale di una procedura automatica; data la grande quantità di materiale trasferita dal web, il procedimento richiederebbe tempi eccessivamente lunghi e dispendiosi, se ad occuparsene fosse una persona soltanto. Il software Web Analyzer confronta ogni pagina web con i documenti presenti nell'archivio, prendendo in considerazione sia il testo vero e proprio della pagina, sia i file allegati. Ogni testo del web viene quindi visualizzato accanto al testo del nativo digitale che presenta maggiori elementi di similarità nei contenuti. Per ciascuno di essi viene stabilito un indice di similitudine, che indica quale percentuale delle parti testuali sul web trovino corrispondenze nei documenti presenti nell'archivio nativo digitale. Qualora l'indice fornisca un risultato basso, ossia sotto il 50%, la corri-

Home Classificazione PadExplorer PadPages Stampa

Verificato: X SI [] NO

Titolo del link:
Bando "Premio Marazza 2016"

Filename
files/pdf/bando_marazza_2016.pdf

WEB Similarity: 30% NATIVO

PREMIO NAZIONALE DI POESIA E DI TRADUZIONE POETICA ACHILLE MARAZZA
BANDO 2016 - XX EDIZIONE

La Fondazione Achille Marazza, in collaborazione con la Regione Piemonte e con il Comune di Borgomanero, bandisce la ventesima edizione del Premio Nazionale di poesia e traduzione poetica

"Achille Marazza" con il seguente regolamento:
PREMIO ACHILLE MARAZZA
- sezione traduzione

Premio destinato a una traduzione poetica da lingue antiche o moderne edita tra il 1 ottobre 2014 e il 31 gennaio 2016, con una dotazione di euro 2.000.

- sezione poesia

Tre premi da euro 1000 ciascuno a tre libri di poesia editi tra il 1 ottobre 2014 e il 31 gennaio 2016.

La Fondazione acquisterà 25 copie di ciascuno dei tre libri finalisti. Una giuria popolare composta da studenti, insegnanti e lettori della biblioteca voterà - all'interno della tema stabilita dalla Giuria tecnica - il vincitore al quale verrà assegnato un ulteriore premio di euro 1000.

Le opere candidate al premio dovranno essere inviate in sei copie entro il 31 gennaio 2016 a:
Eleonora Bellini, Segreteria del Premio Marazza - Fondazione Achille Marazza, Viale Marazza 5, 28021 BORGOMANERO - NO. Farà fede la data del timbro postale. I volumi devono essere

Filename
1/ROMA
Magrelli su ROMA.doc
Roma / Testi

L'ultima raccolta di Franco Buffoni (nato a Gallarate nel 1948 e considerato fra i maggiori poeti italiani della sua generazione) nasce da uno spaesamento geografico e antropologico, lo spaesamento con cui questo "vochlo longobardo assente" narra la sua scoperta di Roma. Al pari di Penna, sceso dal cuore dell'Umbria, al pari di Pasolini, calato da Bologna e dal Friuli, lo scrittore racconta del trapianto in una capitale insieme classica e levantina, pagana e islamizzata, cristiana e consumista. Proprio seguendo quei suoi due discorsi, egli dedicherà alla Città Eterna pagine che spaziano dall'epica sportiva allo sfruttamento lavorativo, dalla rivendicazione dell'amore omosessuale alla meditazione artistica (come nel bel capitolo sulla pittura caravaggesca).

Ecco allora che al senso del peccato predicato dalla Chiesa, si oppone il sogno di una Grecia "troppo lontana", e mentre "disorganizzata pulsa Roma anonima", "Roma di corsa, Roma disperata", si compie un'ideale, profana staffetta fra passato e presente, con l'opus alexandrinum tramutato nell'opus novum "di un odierno / evasore totale".

Erano tante Rome, recita una sezione del volume. E l'Urbe in effetti si presenta al contempo come tragico scenario della lotta partigiana (dall'Ardeatina a via Rasella), dolente teatro di umilissime vite (fra extracomunitari, anziane rapinate, emarginati, colf), o infine piaga "desertica" (quale si mostra allo sguardo di un Leopardi suddito pontificio nichilista e dissidente). Del resto, conclude Buffoni, dove altro se non qui, le campane delle basiliche arrivano a risuonare "anche in cripta di banca"?

Valerio Magrelli

Figura 10. Il sistema di confronto tra documenti web e documenti d'archivio

spondenza non viene ritenuta significativa; in caso contrario, all'operatore viene segnalata una potenziale corrispondenza. In presenza di corrispondenze attorno al 95% le probabilità che si tratti di uno stesso documento crescono considerevolmente, mentre con il 98% la corrispondenza è certa. Occorre considerare, infatti, che i documenti nell'archivio possono essere in formati diversi da quelli pubblicati o allegati sul web (ad es. un file Word solitamente viene convertito in PDF per essere pubblicato sul sito) e ciò spiega lo scostamento tra i due valori. Ovviamente la macchina non ha le capacità di discernimento degli umani, ai quali è infatti richiesto subentrare per un controllo definitivo. La possibilità di effettuare un controllo diretto sui due testi a confronto rende questa procedura assai più efficace. Nel momento in cui il procedimento automatico ha generato un valore elevato, al computer viene fatta proporre l'assegnazione del file a uno dei progetti già inseriti in seguito alla descrizione manuale dell'archivio. Nel caso opposto, quando cioè è stato riscontrato un valore di indice basso, la casella delle assegnazioni resta vuota e sarà poi compito dell'operatore compilarla nel modo che riterrà più opportuno.

Verificato: Sì No

WEB Similarity: 97% **NATIVO**

Titolo del link:
DALL'INTERVISTA DI PULSONI A BALDI

Filename
files/pulsoni_baldi.pdf

Studi sulla produzione di Franco Buffoni / Testi

www.insulaeuropea.eu 13 - 01- 2015

Dall'intervista di Carlo Pulsoni (Università di Perugia) a Valentino Baldi (University of Malta)

Valentino Baldi (1982) si è laureato a Siena e insegna letteratura italiana e teoria della letteratura presso la University of Malta. È redattore di «Allegoria», del blog «Laletteraturaenoi» e collabora con riviste di italianistica e di letterature comparate fra cui «Moderna», «Strumenti critici», «Intersezioni», «Esperienze letterarie», «Studi novecenteschi» e «Between». Ha pubblicato uno studio su Pirandello e Gadda (Reale invisibile, Marsilio, Venezia, 2010) e un lavoro sulla critica letteraria psicoanalitica (Psicoanalisi, critica e letteratura, Pacini, Pisa, 2014). Ha appena pubblicato un libro dedicato alla teoria letteraria di Francesco Orlando che si intitola Il sole e la morte (Quodlibet, Macerata, 2015). Sta curando, assieme ad Alessandra Ginzburg e Romano Lupolini, un numero monografico di «Moderna» dedicato a letteratura e teoria delle emozioni.

(...)

Quali sono gli autori italiani più letti o più richiesti nel tuo corso?

Mi occupo soprattutto di letteratura del Novecento e di teoria letteraria, dunque due nomi di scrittori spiccano su tutti: Pirandello e Montale. Per la teoria della letteratura, invece, trovo che le pagine di Giacomo Debenedetti siano molto apprezzate dagli studenti. In questi ultimi anni sto provando a far leggere alcune pagine di Francesco Orlando, Sebastiano Timpanaro e Gianfranco Contini: ovviamente con necessarie introduzioni. Però devo anche menzionare Pasolini e Fortini: ogni volta

Filename
1/Turing
Pulsoni Baldi.docx

Studi sulla produzione di Franco Buffoni / Testi

www.insulaeuropea.eu 13 - 01- 2015

Dall'intervista di Carlo Pulsoni (Università di Perugia) a Valentino Baldi (University of Malta)

Valentino Baldi (1982) si è laureato a Siena e insegna letteratura italiana e teoria della letteratura presso la University of Malta. È redattore di «Allegoria», del blog «Laletteraturaenoi» e collabora con riviste di italianistica e di letterature comparate fra cui «Moderna», «Strumenti critici», «Intersezioni», «Esperienze letterarie», «Studi novecenteschi» e «Between». Ha pubblicato uno studio su Pirandello e Gadda (Reale invisibile, Marsilio, Venezia, 2010) e un lavoro sulla critica letteraria psicoanalitica (Psicoanalisi, critica e letteratura, Pacini, Pisa, 2014). Ha appena pubblicato un libro dedicato alla teoria letteraria di Francesco Orlando che si intitola Il sole e la morte (Quodlibet, Macerata, 2015). Sta curando, assieme ad Alessandra Ginzburg e Romano Lupolini, un numero monografico di «Moderna» dedicato a letteratura e teoria delle emozioni.

(...)

Quali sono gli autori italiani più letti o più richiesti nel tuo corso?

Mi occupo soprattutto di letteratura del Novecento e di teoria letteraria, dunque due nomi di scrittori spiccano su tutti: Pirandello e Montale. Per la teoria della letteratura, invece, trovo che le pagine di Giacomo Debenedetti siano molto apprezzate dagli studenti. In questi ultimi anni sto provando a far leggere alcune pagine di Francesco Orlando, Sebastiano Timpanaro e Gianfranco Contini: ovviamente con necessarie introduzioni. Però devo anche menzionare Pasolini e Fortini: ogni volta

Figura 11. Percentuale di somiglianza. Documenti con un basso e un alto similarity index

È ovviamente possibile che il testo sia stato prodotto direttamente sulla rete oppure che il file originale sia stato eliminato o non conferito, e che perciò non si individuino corrispondenze. Per ottenere questi risultati PAD ha implementato all'interno del software l'algoritmo della Distanza di Levenshtein²⁴, adattandolo alle proprie esigenze. Attualmente PAD sta sperimentando descrizione delle pagine personali dei più comuni social network (Facebook, Instagram, Twitter e Youtube). I backup delle pagine sono stati consegnati dall'autore e sono già entrati nel processo di conservazione a lungo termine²⁵.

La stampa del catalogo

Una volta terminato il lavoro sui singoli file, il risultato finale può essere visualizzato attraverso la stampa della descrizione dell'archivio. La stampa può essere generata secondo modalità diverse. Il rapporto può riprodurre la struttura dell'archi-

vio, ossia l'albero completo dei suoi vari livelli. Selezionando, invece, un rendiconto delle assegnazioni si otterrà un elenco più dettagliato dove, oltre alla struttura dell'albero, viene mostrato ogni file assegnato a un Progetto. Accanto al nome, che resta sempre quello dato dall'autore, viene indicata la cartella in cui è inserito il file nella struttura originale dell'archivio, l'estensione, i tag assegnati ed eventualmente l'autore e il titolo della pubblicazione da cui è stata estratta. Prevedendo che un autore potrebbe fare conferimenti successivi, man mano che il suo archivio digitale si accresce di nuovi file, viene premesso un numero che identifica il conferimento dal quale il file sia arrivato a PAD.

Elenco documenti per assegnazione	
Godwin Club Italia	
Estratto da	
1/Carmide e Reading 2003 - D'Elia, Gaetano - Letteratura e ipocrisia... (Godwin Club Italia - apr 2003).doc	
Avanti!	ISSN 0391-6685
Estratto da	
1/Casa Palestro Broda Finello su Casa Pal.docx	
1/UVCCI Intervista Avanti ABM.docx	
1/ROMA alberto toni pdf.pdf	
1/7hs-Matr-Scl-Sra-Adda-Q-3D-AD TONI.doc	
Le Cronache del Salernitano	
Estratto da	
1/Casa Palestro Enzo Salerno.pdf	
Le Monde diplomatique	
ISSN 1127-5243	
Estratto da	
1/Casa Palestro Finelli Casa pel.pdf	
1/LAICO ALFABETO FINELLI .doc	
1/LAICO ALFABETO MONDE.jpg	
1/O Germania Finelli su O Germania.jpg	
1/SERVO BYRON Finelli Monde .pdf	
1/Turing Finelli su Turing.pdf	

Figura 12. La stampa del catalogo. La stampa dell'albero del posseduto

²⁴ Paul E. Black, *Levenshtein distance*, in: *Dictionary of Algorithms and Data Structures [online]*, U.S. National Institute of Standards and Technology, 2008, <https://www.researchgate.net/publication/246951886_Dictionary_of_Algorithms_and_Data_Structures>.

²⁵ Per approfondimenti sull'argomento si veda il contributo intitolato *Long-term preservation of a web resource: PAD Web Archiving*, presentato da Laura Pusterla, Primo Baldini e Paul Gabriele Weston all'11th *Qualitative and Quantitative Methods in Libraries International Conference*, Istituto Universitario Europeo, 31 maggio 2019, i cui atti sono in corso di pubblicazione. È già stato pubblicato il libro degli abstracts Pusterla - Baldini - Weston 2019.

Pavia Archivi Digitali		Fondo Franco Buffoni
O Germania		
O Germania		
O Germania / Copertine		
1/O Germania /	cop_ogermania_1.jpg	[JPEG]
COPERTINA		
1/O Germania /	O Germania I Cover.pdf	[PDF]
COPERTINA		
O Germania / Promozione e studi		
1/JUCCI /	Pecora Fuis.doc	[DOC]
INTERVISTA - Intervistatore: Pecora, Elio - In: FUIS - Federazione Unitaria Italiana Scrittori		
1/O Germania /	Germania ADD.docx	[DOC]
ARTICOLO - Autore: Jubilee Debt Campaign - Traduttore: Faber Fabbris - Estratto da: Next quotidiano		
1/O Germania /	Artikel bassanelli.pdf	[PDF]
CITAZIONE - Autore: Marzoli, Valeria - Estratto da: Corriere d'Italia		
1/O Germania /	Artikel von Valeria (Automatisch gespeichert) (1).pdf	[PDF]
CITAZIONE - Autore: Marzoli, Valeria - Estratto da: Corriere d'Italia		
1/O Germania /	di Biasio facebook 4.webloc	[application/xml]
LINK - RECENSIONE - Autore: Di Biasio, Simone - Estratto da: Reader's Bench		
1/O Germania /	Finelli su O Germania.jpg	[JPEG]
RECENSIONE - Autore: Finelli, Claudio - Estratto da: Le Monde diplomatique		

Figura 13. La stampa del catalogo. La stampa dell'albero con l'elenco dettagliato dei documenti legati a ciascun progetto

Vi è, infine, la possibilità di chiedere al software un elenco delle assegnazioni dei file. Poiché quando i documenti presenti non sono stati prodotti da Franco Buffoni, i nomi dei responsabili sono stati collegati al file, questa tipologia di stampa fornisce un elenco di tutti i nomi che sono collegati a documenti nell'ar-

Nome / Ruolo		Elenco documenti per assegnazione
Abbi, Marta		
Autore		
1/Casa Palastro	Marta Abbi.docx	
Accolla, Dario		
VIAF ID: 118146094150700330568		
Autore		
1/Casa Palastro	Accolla su Casa.docx	
1/OSCAR	accolla su fb.doc	
Adolgho, Armando		
Intervistatore		
1/FG luca, paolo	Reportage su Adolgho	
Agostinali, Mattia		
Autore		
1/7th Main St - Sir - Adda - Q - 3D - AD	AGOSTINALI.doc	
Albani, Sauro		
VIAF ID: 69529127		
Autore		
1/Profilo del Rosa	ALBANI.doc	
Alfano, Giancarlo		
VIAF ID: 64225160		
Autore		
1/LIFT - SONGS	Alfano LIFT.doc	
Amerio, Andrea		
VIAF ID: 108073716		
Autore		
1/FG luca, paolo		

Figura 14. La stampa del catalogo. Elenco alfabetico delle persone produttrici di documenti dell'archivio, con l'elenco dettagliato dei documenti in cui hanno una responsabilità intellettuale

Pavia Archivi Digitali		Fondo: Franco Buffoni
Avrei fatto la fine di Turing		
Avrei fatto la fine di Turing		
Avrei fatto la fine di Turing / Copertine		
Avrei fatto la fine di Turing / Promozione e studi		
Avrei fatto la fine di Turing / Testi		
Avrei fatto la fine di Turing / Web		
Carmide a Reading		
Carmide a Reading		
Carmide a Reading / Promozione e studi		
Carmide a Reading / Testi		
Carmide a Reading / Web		
Con il testo a fronte		
Con il testo a fronte		
Con il testo a fronte / Copertine		
Con il testo a fronte / Promozione e studi		
Con il testo a fronte / Testi		

Figura 15. La stampa del catalogo. Elenco alfabetico delle pubblicazioni su cui sono apparse o da cui sono stati estratti i documenti dell'archivio

chivio, per ognuno indicando i file di cui sono autori e il ruolo intellettuale che hanno svolto.

Procedendo in modo analogo, il sistema produce l'elenco dei libri, delle riviste e dei blog da cui sono stati estratti i documenti dell'archivio.

4. Riflessioni conclusive

L'esperienza maturata nei dieci anni trascorsi dall'avvio della pianificazione del progetto PAD alla conclusione della descrizione del fondo Buffoni consentono oggi di avanzare una serie di riflessioni volte a metterne in evidenza le criticità rilevate e a gettare luce sulle opportunità che l'applicazione di uno strumento di conservazione a medio-lungo termine di archivi letterari digitali nativi può offrire a PAD stesso e, più in generale, ad altre esperienze della medesima natura in ambito nazionale ed internazionale.

Un primo elemento di riflessione è costituito dalla mole di dati che anche i fondi di minore estensione presentano, un numero tale di file da rendere improponibile la soluzione di affidarne l'ordinamento e la descrizione unicamente all'archivista e che richiede, al contrario, la messa a punto di algoritmi che si facciano carico di analizzare i contenuti, l'esistenza e la correttezza dei metadati, la corretta individuazione delle responsabilità, l'esistenza di relazioni tra file presenti in diverse sezioni del fondo, il trattamento dei file duplicati, di quelli rimossi o non più accessibili. Sono procedure ad alto impatto in termini di risorse che impegnano, tuttavia non eliminabili se si intende offrire un servizio di alta qualità allo studioso ed incrementare le possibilità che l'archivio digitale sopravviva al trascorrere del tempo. L'esperienza ha mostrato, inoltre, che l'applicazione di algoritmi, lungi dal mortificare la professionalità dell'archivista, richiede che essa sia sollecitata a delineare procedure che tengano conto dell'apporto dell'informatica e ad immaginare nuovi servizi per la consultazione e lo studio dei documenti.

Una seconda riflessione, in qualche modo collegata alla prima, riguarda la messa a punto di criteri di scarto calati sulle caratteristiche del digitale. Mi riferisco, in particolare, alla massa di dati generati dai sistemi utilizzati dagli autori per tracciare il loro comportamento in rete o per tenere memoria dei file rimossi. L'applicazione di procedure di informatica forense, occorrenti al trasferimento ottimale dei file e alla lettura dei documenti prodotti con software non più in uso o archiviati su supporti che abbiano subito danni a causa del tempo o di criteri di conservazione non ideali, fa emergere un gran numero di informazioni che l'autore conferisce a propria insaputa e che potrebbero venire utilizzati in modo poco etico. È questa una differenza fondamentale tra quanto un archivio cartaceo dice del soggetto che l'ha prodotto e quanto una invasiva analisi affidata all'informatica permette di conoscere della vita privata del medesimo soggetto. Un approccio eticamente corretto alla conservazione non può spingersi al punto di fornire

agli utenti di domani informazioni in grado di mutare uno studio scientifico in una storia “vista dal buco della serratura”.

Di segno completamente opposto è l'opportunità che offre il colloquio con l'autore che precede il riversamento dei suoi file dai suoi dispositivi all'archivio di PAD. Se, infatti, la motivazione del colloquio è predisporre le apparecchiature impiegate per la procedura, l'aspetto in definitiva più rilevante è disporre di un quadro approfondito delle sue pratiche di scrittura, del ruolo che il digitale ha nella sua produzione letteraria e nel suo rapporto con il contesto in cui opera (amici, colleghi, lettori, la società in generale e persino se stesso), una opportunità quest'ultima che sovente non offrono gli archivi cartacei, specie nel caso in cui il conferimento venga gestito dagli eredi che forse poco conoscono delle abitudini dello scrittore. I frequenti contatti con lo scrittore sono utili anche a sensibilizzarlo sulla necessità di organizzare il proprio archivio in modo tale da favorirne la conservazione e l'organizzazione dei dati per l'accesso. Evidentemente non si tratta di chiedere allo scrittore di rinunciare alle proprie abitudini a beneficio di una più efficiente sequenza di procedure informatiche; si tratta, invece, di informarlo dei rischi per la conservazione dei dati e per l'integrità del suo archivio che possono derivare da una insufficiente attenzione dedicata all'aggiornamento del software, all'allineamento tra i diversi dispositivi utilizzati, ai criteri utilizzati per denominare file e cartelle e per la produzione di metadati da parte del sistema che utilizza. Che queste informazioni siano utili lo dimostra il fatto che nella preparazione di conferimenti successivi alcune delle incongruenze registrate la prima volta sono assenti o significativamente ridotte.

Si è detto che quello di Buffoni rappresenta un caso emblematico di archivio ibrido. In realtà, verifiche effettuate su un certo numero di archivi cartacei ospitati dal Centro Manoscritti ha consentito di verificare l'esistenza di documenti digitali ospitati su floppy, cassette e compact disc dei quali, nel migliore dei casi, si è segnalata l'esistenza senza provvedere al riversamento dei file (lo stesso si può dire a proposito di documenti non digitali, ma egualmente bisognosi di attenzione, come registrazioni di conversazioni, interviste e letture ospitate su dischi in vinile e nastri magnetici). Gli inconvenienti riscontrati nell'accesso ai file consegnati dagli scrittori da un lato ha messo a fuoco l'esistenza del problema e dall'altro lato ha fornito l'occasione per dotarsi dei dispositivi hardware e software che possono contribuire a risolverlo.

Tra le criticità fin qui emerse almeno due meritano di essere almeno segnalate, in quanto presuppongono l'adozione di soluzioni e buone pratiche condivise tra più soggetti istituzionali: le questioni legali e la conservazione a lungo termine. Quanto a quest'ultima, infatti, non si può non rilevare che essa richiede ambienti informatici appositamente attrezzati e una stabilità nella gestione delle procedure che l'università, istituzioni nelle quali hanno preso avvio progetti di natura simile a PAD, non può in genere garantire. Se tra i compiti di un ateneo la ricerca, la di-

dattica e da qualche tempo quella che viene definita la “terza missione” rientrano a pieno titolo, la conservazione digitale presuppone la disponibilità a tempo indeterminato di risorse materiali e professionali che essa non ha. Non è infrequente che progetti anche di notevole interesse, avviati in seguito al finanziamento di una ricerca, nel momento in cui esso viene a mancare, cessino di essere popolati di nuovi dati e che il server che li ospita, una volta dismesso, venga ripulito per essere adibito ad altri servizi. Considerando le questioni di riservatezza l’alternativa non può consistere nell’archiviazione su un cloud. Occorre un *deep archive* gestito da una istituzione che garantisca professionalità, stabilità e continuità, come il servizio fornito in Italia da Magazzini Digitali.

L’ultima considerazione riguarda la visibilità dell’archivio e la sua integrazione nel tessuto culturale del Paese. In questa direzione si è adoperato fin dalla sua ideazione PAD, perseguendo l’adozione degli standard archivisti e biblioteconomici, l’implementazione di modelli di dati e di linguaggi di codifica che favoriscano l’interoperabilità e auspicando il coinvolgimento nella disseminazione dei dati tanto dei sistemi di aggregazione e ricerca predisposti dalle istituzioni culturali, quanto di quei *referrals* che sono in definitiva tra le risorse più utili ed utilizzate del web.

In 2009 the University of Pavia launched PAD – Pavia Archivi Digitali with the idea of creating an archive aimed at collecting digital born memories of writers, journalists and intellectuals. The archive was planned to include files and documents of a different nature, to ensure their long-term preservation and to make them accessible to scholars, in accordance with the author’s provisions.

Following the setting up of a first prototype, the project evolved into a sustainable system, capable of handling large amounts of data and keeping them safe. Having established that the products on the market were not able to support the complexity of the project, the decision was made to develop the system locally, ensuring that it would be scalable, flexible, compliant to standards and interoperable with other platforms.

Literary digital archives are mainly made of text documents (working papers, drafts at various stages of completeness, unpublished papers, pdf of gone to press volumes), and to a lesser extent of audio and video-recordings, photographs and images. PAD is currently extending the type of files to be treated, described and stored, searching solutions designed to support, host and make available to scholars digital material produced on the web (social networks, websites, blogs) and via mail. So far we have focused on files copied from physical media used by authors (their main computer hard disk, external hard drives, CDs, smartphones and mobiles) as well as from cloud storage systems (Dropbox, Google Drive and so on).

Due to the large size of these archives, often including thousands of files, the development team is studying new techniques for data extraction and AI procedures of file recognition. Franco Buffoni’s papers are used as a testbed for these developments.

L’ultima consultazione dei siti web è avvenuta nel mese di giugno 2019

RIFERIMENTI BIBLIOGRAFICI

- Antonelli 2010 Giuseppe Antonelli. *Archivi del giovane autore*. «Il Sole 24 Ore», 21 novembre 2010.
- Carbé 2017 Emmanuela Carbé. *Walking in my Shoes: A Case Study from a Born-digital Archive*. «CEUR», 2021 (2017), p. 137-146.
- Carbé 2018 Emmanuela Carbé. *Digitale d'autore: un archivio possibile?*. In: *Archivi editoriali: tra storia del testo e storia del libro*, a cura di V. Brigatti, A. L. Cavazzuti, E. Marazzi. Milano: Unicopli, 2018, p. 209-216.
- Feliciati et al. 2018 *Archivi digitali d'autore: ipotesi di lavoro*. In: *Book of Abstracts, 7. Convegno Annuale AIUCD 2018 (Bari, 31 gennaio – 2 febbraio 2018)*, a cura di D. Spampinato. Bologna: AIUCD, 2018, p. 256-259 (al panel hanno partecipato P. Feliciati, S. Allegrezza, E. Belisario, S. Vassallo, E. Carbé, P. Baldini, P.G. Weston).
- Pusterla – Baldini – Weston 2019 Laura Pusterla - Primo Baldini - Paul Gabriele Weston. *Long-term preservation of a web resource: PAD web archiving*. In: *Book of Abstracts. 11th Qualitative and Quantitative Methods in Libraries (QQML2019) International Conference, European University Institute, Florence, May 28-31, 2019*. Athens: ISAST, 2019, p. 114, <http://qqml.org/wp-content/uploads/2017/09/Book-of-Abstracts_Final_AfterConf_v1.pdf>.
- Stefani 2016 Michele Stefani. *Gli archivi degli scrittori nativi digitali: verso la conservazione e oltre*. Tesi di laurea in Storia e gestione del patrimonio archivistico e bibliografico. Relatore Paul Gabriele Weston, correlatore Riccardo Ridi. Università Ca' Foscari Venezia, a.a. 2015/16.
<<http://dspace.unive.it/bitstream/handle/10579/10134/855567-1204608.pdf?sequence=2>>.
- Weston – Carbé – Baldini 2016 Paul G. Weston – Emmanuela Carbé – Primo Baldini. *Hold it All Together: a Case Study in Quality Control for Born-Digital Archiving*. «Qualitative and Quantitative Methods in Libraries (QQML)», 5 (2016), p. 606-710, <http://www.qqml.net/papers/September_2016_Issue/5313QQML_Journal_2016_Westonetal_>.
- Weston – Carbé – Baldini 2017 Paul G. Weston – Emmanuela Carbé – Primo Baldini. *Se i bit non bastano: pratiche di conservazione del contesto di origine per gli archivi letterari nativi digitali*. «Bibliotheca», 6 (2017), p. 154-177.
DOI: <<https://doi.org/10.6092/issn.2283-9364/7027>>.
- Weston – Carbé – Baldini 2017a Paul G. Weston – Emmanuela Carbé – Primo Baldini. *Conservare e rendere accessibile un archivio letterario digitale. Il caso di PAD – Pavia Archivi Digitali*. Atti della giornata di studio AIB “Fondi e collezioni di persona e personalità negli archivi, nelle biblioteche, nei musei: una risorsa, una opportunità”, Bologna, Università di Bologna, 26 ottobre 2016, in corso di stampa.