

# Digi/Italia

ISSN 1972-6201

Anno XIV, Numero 1 - 2019



R I V I S T A  
D E L D I G I T A L E  
N E I B E N I  
C U L T U R A L I

ICCU-ROMA

# 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

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 avvicinabile 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](mailto:digitalia@iccu.sbn.it)  
<http://digitalia.sbn.it>

In attesa di registrazione al Tribunale di Roma



ICCU

**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 Tammaro

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

**Il Portale delle biblioteche  
e degli istituti culturali italiani.  
Presentazione del progetto**

Roma, 11 aprile 2019, Sala Spadolini, MiBACT





# Il Portale delle biblioteche e degli istituti culturali italiani. Presentazione del progetto

Roma, 11 aprile 2019, Sala Spadolini, MiBACT

---

## Relatori:


- Paola Passarelli** - *Direttore generale Biblioteche e istituti culturali*  
**Simonetta Buttò** - *Direttrice dell'Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche (ICCU)*  
**Giovanni Solimine** - *Direttore del Dipartimento di lettere e culture moderne, Sapienza Università di Roma*  
**Claudio Leombroni** - *Responsabile del Servizio Biblioteche, archivi, musei e beni culturali, IBC Regione Emilia-Romagna*  
**Alberto Petrucciani** - *Professore ordinario di Biblioteconomia e bibliografia, Sapienza Università di Roma*  
**Gino Roncaglia** - *Professore associato di Editoria digitale e informatica umanistica, Università della Tuscia*  
**Marino Sinibaldi** - *Direttore di Rai Radio Tre*

## Paola Passarelli

Buongiorno a tutti. Per me è davvero un grande piacere essere qui oggi per la presentazione del progetto del nuovo *Portale delle biblioteche e degli istituti culturali italiani*, un'iniziativa portata avanti con tenacia e passione dalla Dott.ssa Simonetta Buttò in qualità di dirigente dell'ICCU.

Sarebbe dovuto essere con noi anche il Ministro Alberto Bonisoli, ma purtroppo, per concomitanti impegni istituzionali, non è potuto essere qui di persona. Come segno tangibile della sua partecipazione *in spiritu*, mi ha inviato il seguente messaggio proprio per sottolineare il suo pieno sostegno e plauso all'iniziativa odierna. Leggo testualmente dalla nota che ho ricevuto:

«Gentilissima Dott.ssa Passarelli, impegni istituzionali mi impediscono di intervenire alla presentazione del progetto per il nuovo Portale delle biblioteche e degli istituti culturali italiani, che sono ben consapevole essere evento di straordinaria importanza per il mondo delle biblioteche italiane: la biblioteca che si rinnova e diventa piattaforma multimediale, che mette a disposizione di tutti un incommensurabile patrimonio di libri nuovi e antichi, giornali, riviste, filmati, immagini, registrazioni audio. Il MiBACT che si proietta verso il futuro sulle spalle del suo passato: questo è quello che immagino per tutto il nostro patrimonio e che vedo realizzarsi per i beni librari. A tutti voi il mio ringraziamento e i più fervidi auguri di buon lavoro».



Senza retorica, sono io a ringraziare il Ministro per la sua vicinanza e per il suo interessamento che, come avete potuto constatare, ha manifestato in più occasioni: un interesse per il mondo dei libri e, non secondariamente, per le professioni che ruotano attorno a questo mondo e lo fanno esistere, penso in tal senso soprattutto ai bibliotecari.

Oggi si presenta questo progetto che ovviamente è stato sostenuto dalla Direzione generale che mi onoro di dirigere. Si tratta di un progetto che, rendendo interoperabili le diverse banche dati attualmente esistenti che fanno capo all'ICCU, punta a rendere più agevole l'accesso al sapere e alla conoscenza. L'utente, l'utilizzatore finale, sia esso lo specialista o il vasto pubblico, ha la possibilità di esplorare, attraverso un unico punto di accesso e un'unica interrogazione, varie piattaforme e varie banche dati.


Quali sono queste banche dati che finalmente diventano interoperabili? In primo luogo la madre di tutte loro, SBN e poi EDIT16, che riguarda il libro italiano del Cinquecento, Manus, il Portale dei manoscritti italiani, Internet Culturale che è la *digital library* delle biblioteche, attiva dal 2005, il Portale 14-18, nel quale confluisce tutta la documentazione relativa alla Prima Guerra Mondiale e il Portale Movio, dedicato alle mostre virtuali.

Un progetto che coniuga in maniera efficace e meritoria i criteri della conservazione, dell'accesso e della consultabilità con quelli della valorizzazione attraverso l'incremento della fruibilità e dell'usabilità. In questa prospettiva, il *Portale* si offre come uno strumento di conoscenza duttile, stratificato e di grande potenzialità, dal momento che è stato progettato per consentire continue implementazioni del patrimonio censito e delle risorse messe a disposizione. Elementi che ne fanno un vettore *omnibus* di sapere adatto a venire incontro a esigenze anche profondamente diverse.

Mi permetto di aggiungere con soddisfazione che il *Portale* rappresenterà una svolta epocale anche sul crinale dell'efficienza amministrativa. Esso, infatti, fa suo un obiettivo che da sempre mi sta molto a cuore, vale a dire la realizzazione di adeguamenti tecnologici e di progetti di implementazione dei sistemi informativi che oltre a ottimizzare e razionalizzare la gestione di procedure e servizi, sia in grado di rendere quest'ultima più economica e sostenibile in termini di risorse.

Più in particolare, il progetto che viene oggi presentato è stato reso possibile dalla gara europea bandita dall'ICCU, che consentirà di mettere insieme le tecnologie e le esperienze più avanzate per la realizzazione del Sistema di Ricerca Integrato e del relativo Portale. Procedura, questa, che permetterà di migliorare l'efficacia dei servizi e, al contempo, di razionalizzare le spese grazie all'accorpamento delle commesse di manutenzione delle diverse banche dati specialistiche e del sistema di *digital library*.

Quanto detto ci fa già sentire in un mondo futuribile, tuttavia occorre ricordare che la tecnologia è una realtà in continuo, nevrotico divenire, quindi possiamo dire



che la sfida è solo agli inizi. Si pone già da oggi in prospettiva l'ulteriore ambizioso traguardo di mettere a punto un sistema tecnologico che consentirà forme di interoperabilità su scala europea ed extraeuropea. La scommessa del domani è un mondo unico ma molteplice, un mondo sempre più interconnesso per favorire la conoscenza delle reciproche diversità e di quegli scrigni di saperi e modi d'essere che sono le culture.

Vedo un pubblico assai nutrito, tangibile testimonianza del reale interesse suscitato dal progetto. Ringraziando tutti voi per essere presenti, direi a questo punto di passare ai contenuti; cedo dunque la parola agli addetti ai lavori.

## **Simonetta Buttò**

Grazie al signor Ministro per il suo messaggio e alla Dott.ssa Passarelli per la sua introduzione.


Il progetto che presentiamo oggi è il punto di arrivo di una lunga attività di studio che l'ICCU ha iniziato oltre tre anni fa, a partire dall'avvio di una politica di sempre maggiore ascolto della vasta comunità che si raccoglie intorno al Servizio Bibliotecario Nazionale, e di sempre maggiore apertura del catalogo a tipologie di materiali diversi, su supporti differenti.

SBN vanta una attività ininterrotta ultratrentennale e che interessa l'intero territorio nazionale: oggi le biblioteche e gli istituti culturali che aderiscono a SBN sono oltre 6.400, di ogni tipologia e di ogni appartenenza amministrativa: ci sono le nostre biblioteche dei beni culturali, quelle delle università, le numerosissime di ente locale, le ecclesiastiche, le scolastiche, quelle degli istituti culturali, di enti privati, associazioni, fondazioni. Tutte insieme partecipano alla catalogazione collettiva del patrimonio bibliografico italiano.

Per il numero e la varietà delle amministrazioni partner, per il volume delle informazioni che offre, per la sua solidità e per la lunga durata che lo contraddistingue, SBN oggi rappresenta una realtà imprescindibile per l'accesso al patrimonio bibliografico, analogico e digitale, del nostro Paese, e la più grande "vetrina" del patrimonio culturale nazionale in rete: ogni anno si conta una media di oltre 60 milioni di ricerche effettuate dal pubblico e decine di milioni di prestiti, per 18 milioni di titoli con quasi 100 milioni di localizzazioni.

Non è un caso che SBN sia stato inserito dall'AGID nel *Piano triennale per l'informatica 2019-2021*, recentemente rilasciato, fra le basi dati di interesse nazionale rilevanti per lo svolgimento delle funzioni istituzionali delle Pubbliche amministrazioni<sup>1</sup>. Di questo risultato dobbiamo ringraziare anche i buoni auspici degli amici dell'AIB presso la Funzione pubblica.

<sup>1</sup> <https://pianotriennale-ict.italia.it/assets/pdf/2019-2021/Piano-Triennale-ICT-2019-2021.pdf>.



In SBN, grazie alla catalogazione partecipata e a campagne di recupero del digitale disperso, circa un milione di record bibliografici sono arricchiti oggi da un link diretto alla risorsa digitale esistente.

Ma oggi SBN è ben più di un database bibliografico: grazie alla flessibilità del sistema e alla lungimiranza delle nostre regole di catalogazione, le REICAT, sono infatti presenti in base dati varie tipologie di materiali non librari, sicuramente di grande interesse per il pubblico: ritratti di persone; immagini (carte, stampe, fotografie, cartoline) di luoghi; risorse audio e video (interviste, discorsi registrati, voci, musiche, canti, documentari); autografi, carteggi e documenti; mostre virtuali; oggetti d'arte e cimeli e perfino materiali non effettivamente pubblicati, come conferenze, appunti/dispense, esecuzioni musicali presentate nei festival o in piazza, saggi e performance.


Questa caratteristica "pluralità" di SBN, relativa sia all'ampiezza e varietà della cooperazione che alla diversità dei contenuti, consente di dare voce al policentrismo del patrimonio culturale italiano, che a questo livello non ha riscontro in nessun altro paese d'Europa, e contemporaneamente alla creatività degli istituti culturali italiani, che producono – come laboratori della cultura – pubblicazioni, convegni, mostre, ricerche, progetti.

Ma accanto a SBN l'ICCU ha sviluppato altri servizi dotati delle loro specifiche qualità, che vanno anche al di là di SBN, per il prestigio dei partner coinvolti e per la capacità di apertura alle collaborazioni con l'estero, recentemente di molto accresciute.

EDIT16, il censimento delle edizioni italiane del XVI secolo (quasi 70.000 schede e altrettante immagini) coinvolge 1.591 biblioteche partner, fra cui la Biblioteca Apostolica Vaticana, la Biblioteca di Stato della Repubblica di San Marino, la Biblioteca del Senato, la Biblioteca dell'Università cattolica di Milano, molte biblioteche di istituzioni ecclesiastiche oltre a una quarantina di collezioni private. Dal 2017 vi partecipa la British Library, che possiede oltre 9.000 cinquecentine italiane e dal 2018 anche la Biblioteca della Fondation Barbier-Mueller di Ginevra, con 354 edizioni italiane del Cinquecento. EDIT16 indicizza altre 24 biblioteche straniere (non partecipanti a EDIT16) detentrici di oltre 1.900 esemplari digitali, per un totale di 14.500 link a copie digitali di esemplari italiani e stranieri.

Anche il censimento dei manoscritti antichi e moderni, Manus OnLine, può contare su un vasto partenariato, anche al di fuori di SBN, che ammonta a 411 enti di conservazione e ricerca, che hanno popolato la base dati di 176.000 record e 32.000 immagini. Manus OnLine è in grado di accogliere le descrizioni di manoscritti conservati all'estero che compaiono in progetti tematici sviluppati nell'ambito di ricerche specialistiche.

Tutta questa ricchezza di documentazione è molto apprezzata dagli specialisti di vari settori anche a livello internazionale, ma non è ancora abbastanza conosciuta e utilizzata da parte di un pubblico più generale, che poi è quello che frequenta i



luoghi della cultura, che ci tiene a essere informato, ma che è rappresentato anche dal mondo della scuola, degli istituti italiani di cultura all'estero, dalle Società Dante Alighieri e da tutte le istituzioni in prima linea per la promozione della lingua italiana nel mondo.

Nel nuovo Portale verrà valorizzata anche una banca dati specialistica e interdisciplinare come *14-18: Documenti e immagini della grande Guerra*, che contiene 630.000 oggetti digitali (diari, giornali di trincea, cimeli, fotografie, fogli matricolari dei soldati, monumenti e lapidi, opuscoli, canti, voci registrate), forniti da oltre 200 partner e il sito MOVIO, uno strumento semplice e agile per la realizzazione di mostre virtuali, che contiene oltre 200 mostre prodotte da 150 istituzioni diverse.

Il caposaldo del nuovo Sistema di Ricerca Integrato – per quanto riguarda il digitale – è rappresentato da Internet Culturale, la *Digital Library* delle biblioteche e degli Istituti culturali italiani che ha un milione di documenti digitali (per 15 milioni di risorse) e al quale partecipano 179 istituti.


Grazie all'aggregatore nazionale Culturalitalia, che insieme a *Internet Culturale* costituisce il ponte che collega le istituzioni culturali del nostro paese con le grandi infrastrutture europee, l'ICCU garantisce il flusso di risorse digitali legate al patrimonio culturale italiano verso il portale europeo *Europeana*, con cui condivide standard e *policy* di pubblicazione dei dati.

Nel nuovo modello informativo, i dati presenti in Indice SBN, nella biblioteca digitale di Internet Culturale e nei singoli database specialistici, saranno tutti contemporaneamente accessibili a partire da un'interfaccia unica che rappresenterà il vero e proprio portale delle biblioteche e degli istituti culturali italiani offerto alla cittadinanza, una piattaforma di servizi informativi in grado di offrire al pubblico una nuova esperienza di navigazione tra record di catalogo arricchiti e risorse digitali, anche di natura diversa da quella bibliografica.

Le caratteristiche che dovrà avere questo Portale che presentiamo oggi sono riassumibili in poche righe: dovrà integrare informazioni che oggi sono separate in diverse basedati, costruite sulla base di sistemi tecnologici diversi; dovrà integrare l'informazione con gli oggetti digitali; dovrà comunicare con un linguaggio chiaro, amichevole, evitando tecnicismi e in particolare il nostro gergo professionale, per rivolgersi in buon italiano a tutto il pubblico; dovrà accompagnare l'utente, attraverso un'interfaccia intuitiva, a svolgere ricerche anche sofisticate come quelle di solito riservate ad una utenza specializzata.

La semplificazione della ricerca nel nuovo portale, nel quale sarà possibile interrogare contemporaneamente tutte le basi dati gestite dall'ICCU, a partire da alcune indicazioni chiave (Persone, Luoghi, Mostre, Biblioteche...), oppure a partire da alcune tipologie di materiali (manoscritti, fotografie, musica...), non sarà il varco per la banalizzazione dei risultati, ma restituirà la complessità dei contenuti culturali nelle loro relazioni semantiche.

Non solo. La prospettazione delle risposte avverrà in modo ordinato, rispecchian-



do la varietà delle risorse trovate: libri (analogici o digitali) di un certo autore, libri su quell'autore, la sua fotografia o ritratto, le sue lettere manoscritte, una intervista radiofonica, un video, una mostra (reale o virtuale) a lui dedicata, tanto per fare un esempio, distribuiti con chiarezza – per categorie – sulla pagina dei risultati della ricerca.

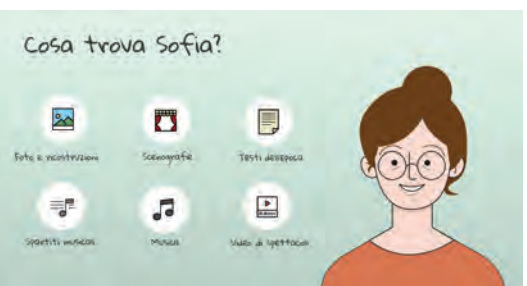
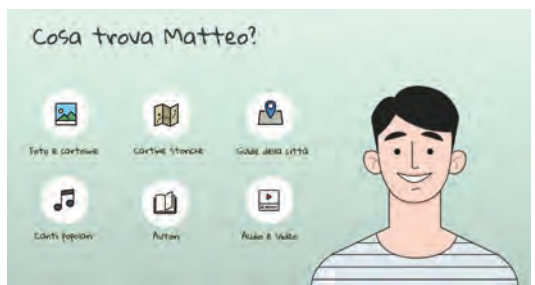
Quello che vogliamo ottenere con il Sistema di Ricerca Integrato – cioè il motore tecnologicamente avanzato che sta dietro le quinte e il nuovo Portale che consente al pubblico di accedervi – è una operazione che potremmo chiamare di *empowerment*, che consenta al pubblico di utilizzare facilmente le funzionalità più avanzate e complesse che sono presenti nel sistema informativo e che i nostri dati di qualità – che sono dettagliati, uniformati, standardizzati e interoperabili – permettono ampiamente.

In più il portale offrirà anche informazioni sulle risorse nuove che via via arrivano ad arricchire il servizio, sui nuovi progetti e sulle iniziative delle biblioteche e degli istituti culturali che potranno segnalare in modo semplice e semiautomatico mostre, eventi e altre iniziative che compariranno in un calendario continuamente aggiornato. Un modo per dare visibilità e incoraggiare i partner a fare sempre di più. Non si tratta di creare una nuova *digital library* ma di mettere al sicuro, in un luogo certo e facilmente accessibile, il prodotto del grande lavoro culturale svolto in Italia da biblioteche e istituti culturali presenti in tutti gli angoli del territorio nazionale.


Durante la realizzazione del progetto l'ICCU sarà impegnato infatti nel recupero all'interno dei sistemi nazionali delle risorse digitali di ogni tipo esistenti nel Paese, ma spesso conosciute solo in ambito locale oppure "sommese", o meglio annegate, in centinaia di *repository* e siti web difficilmente raggiungibili dagli utenti, per farli emergere e per dare loro la migliore e più ampia visibilità.

Il passo successivo del nostro progetto è rappresentato dalla creazione di un nuovo ambiente gestionale al servizio della comunità professionale: una piattaforma in *cloud* e rigorosamente basata su *software open source* che consentirà ai bibliotecari di descrivere il patrimonio e caricare il digitale, parziale o integrale, che hanno, senza uscire dalla consueta procedura di catalogazione. Allora il cerchio si chiuderà, e il Digitale sarà direttamente allegato alla scheda di catalogo, senza più rischio di dispersione.

La flessibilità e l'apertura del sistema che stiamo costruendo prevedono anche un potenziamento delle relazioni esterne alla comunità di SBN: è nostra intenzione stringere accordi mirati con importanti istituzioni nazionali che hanno prodotto nel tempo siti culturali di grande successo che – attraverso il nuovo Portale – potranno suggerire nuovi spunti di ricerca a un pubblico più vasto (Treccani, Istituto Luce, TecheRAI...), e magari nuove scoperte. Come si vede in questo *artboard* realizzato in collaborazione con la RTI aggiudicataria della gara europea e composta da DM Cultura, Inera e GruppoMeta.







La gara europea per la realizzazione del Sistema di Ricerca Integrato e del relativo Portale che servirà al pubblico per esplorare il patrimonio culturale, che è stata appena aggiudicata, consentirà dunque all'ICCU di migliorare l'efficacia dei suoi servizi, ma anche di ridurre le spese grazie all'accorpamento delle commesse di manutenzione delle diverse banche dati specialistiche e del sistema di *digital library*, fino ad oggi oggetto di forniture diverse.


## Giovanni Solimine

Immagino di essere stato invitato a questa presentazione per il fatto che il nostro dipartimento ha da tempo una collaborazione con l'ICCU: la convenzione in atto con l'Istituto prevede la nostra partecipazione a diverse iniziative, in alcuni casi anche con altri partner: per esempio, assieme all'Istat stiamo lavorando con l'ICCU sull'Anagrafe delle biblioteche per far partire da lì quella che forse sarà la prima indagine a tappeto sulle biblioteche italiane.

All'interno di questa convenzione c'è anche una collaborazione che riguarda il progetto del portale che si presenta questa mattina: nel merito del nostro contributo a questa iniziativa interverrà con maggiore precisione più avanti il collega Petrucciani.

Per noi questa collaborazione è importante, non solo perché la Biblioteconomia è una disciplina applicata; quando si ha la possibilità di partecipare ad attività, progetti e iniziative di punta come SBN e a quello che si muove attorno all'ICCU nel panorama delle biblioteche italiane questa è la migliore occasione, credo, per noi e per i nostri studenti, di poter essere al centro di quello che accade, di ciò che si muove nel quadro degli istituti bibliotecari del nostro paese; ringrazio quindi il Direttore generale e il Direttore dell'ICCU per l'invito e per questa opportunità di collaborazione.

Voglio soltanto riprendere un paio di cose che sono state dette nei due interventi di apertura e che credo in qualche modo possano essere motivo di riflessione e anche di apprezzamento dell'iniziativa che sta partendo: diceva prima la Dottoressa Buttò che l'obiettivo di questo portale è di consentire un accesso semplificato, ma non banale, e questo è uno dei problemi, cioè la semplificazione. È sicuramente necessario, è indispensabile: però un accesso semplificato non deve sacrificare la complessità, come ha detto Simonetta Buttò, non deve in alcun modo impoverire i contenuti. Questo è il punto di equilibrio che va ricercato, perché effettivamente può essere uno dei modi attraverso i quali poi un portale come questo trova anche la sua identità all'interno dell'offerta della rete. Rivolgendosi a un pubblico eterogeneo trovare il giusto modo di proporsi non è facile, ma credo che questo debba essere l'obiettivo principale, perché sicuramente la rete, con la sua grande semplicità e con quella grande quantità di in-



formazioni e di documenti ai quali ci consente di accedere, rischia di creare anche un effetto di sazietà che può farci perdere il gusto di andare un po' oltre nella ricerca. Una studiosa importante nel campo degli effetti che il digitale provoca all'intelletto umano – mi riferisco a Maryanne Wolf – ha evidenziato che la grande facilità e amichevolezza di alcuni strumenti di ricerca rischiano di farci perdere la “pazienza cognitiva”, cioè la necessità di dedicare un po' più di tempo a una ricerca, di fare un po' più di fatica, per poter andare oltre l'immediato. Un altro aspetto che credo sia emerso chiaramente è l'integrazione, non soltanto tra diverse tipologie di documenti, ma tra fonti e provenienze diverse di questi documenti, quindi fra documenti e informazioni nati per scopi differenti, ma che vengono orientati al raggiungimento di una medesima finalità. Si tratta di strumenti diversi, che usano linguaggi diversi, che propongono approcci differenti, che vanno in qualche modo ricondotti a unità, con l'obiettivo di stimolare la curiosità e anche quella pazienza cognitiva di cui parlavo prima, spingendo gli utenti a ricercare contenuti di qualità.

Questo è un ultimo aspetto che vorrei sottolineare: la qualità che viene da iniziative come queste consente all'offerta bibliotecaria di distinguersi all'interno dell'offerta del web, esercitando una propria specifica funzione. Ciò è possibile perché si tratta di contenuti validati, e perché dietro l'allestimento di questa offerta c'è un lavoro enorme.


Credo che il compito di tutti noi sia quello di sostenere il grande lavoro che l'ICCU ha fatto in 30 anni e che ora si appresta a fare con questo portale.

Buon lavoro!

## Claudio Leombroni


Il progetto di cui Simonetta Buttò ha illustrato il *rationale* è molto importante e travalica in realtà la semplice progettazione e realizzazione di un portale.

Ciò per almeno tre ragioni che vorrei brevemente argomentare. In primo luogo, per l'aspetto tecnologico, che commenterò senza entrare nel dettaglio utilizzando alcuni concetti e alcuni strumenti interpretativi proposti da uno dei più innovativi teorici dei sistemi informativi, purtroppo scomparso qualche anno fa, Claudio Ciborra. Progettare un punto di accesso unificato alle tante banche dati gestite dall'ICCU con modalità di ricerca capaci di restituire la complessità delle relazioni senza banalizzazioni, affrontare il tema, centrale da almeno quindici anni, dell'integrazione con le risorse digitali, disegnare un cloud SBN e una teca SBN del digitale, prevedere una interoperabilità con servizi internazionali, rappresentano indubbiamente un salto tecnologico paragonabile agli altri due o tre compiuti da SBN nell'ultimo ventennio. Certo il progetto non affronta alcuni aspetti che a mio avviso sono oggi necessari, come una revisione della struttura dell'Indice SBN e




del record bibliografico SBN, o non affronta affatto, credo volutamente e per precise scelte biblioteconomiche, alcuni temi emersi nel dibattito della comunità SBN; né dal punto di vista architeturale sembra sfruttare appieno filosofie, paradigmi e soluzioni del web attuale. Tuttavia, se consideriamo SBN come una infrastruttura, le scelte progettuali effettuate possono far conto su argomenti di plausibilità o ragionevolezza.

In quanto infrastruttura, SBN non coincide infatti con un insieme di hardware e applicazioni software, ma investe necessariamente anche i processi di business della cooperazione. Da questo punto di vista le scelte manageriali devono tener conto del contesto strategico di SBN, dell'analisi dei bisogni di migliorare o trasformare i processi esistenti e dei possibili cambiamenti in termini di ruoli e responsabilità e di una sorta di assioma che Michel Boisset formulò così: «il n'y a pas de coopération, si l'on n'a pas des partenaires autonomes». In altre parole, la gestione dell'infrastruttura deve affrontare il problema dell'allineamento strategico fra l'architettura ICT, i processi fondamentali e i requisiti informativi; fra le scelte locali anche in termini ICT e la necessità di preservare centralmente una logica di business coerente, se non una intelligente integrazione, pur garantendo un'ampia autonomia dei poli/reti/comunità che compongono il Servizio Bibliotecario Nazionale. In questa prospettiva il progetto non potrà non confrontarsi con i profili istituzionali della cooperazione SBN, ancorché attualmente priva degli organi di governo, peraltro previsti da patti probabilmente da riscrivere, e conseguentemente fortemente limitata nella possibilità di condividere ad alto livello progetti e scelte. Così come non può non confrontarsi con la gestione di una infrastruttura che coinvolge molteplici attori, con interessi in qualche caso divergenti e quindi potenzialmente in grado di evolvere secondo direzioni diverse fra centro e periferia. Inoltre, l'infrastruttura SBN, come ogni altra grande infrastruttura, non può essere cambiata all'istante: la nuova versione deve essere compatibile con l'infrastruttura preesistente. In un certo senso si può dire che ogni nuova grande infrastruttura non è mai sviluppata dal nulla. Ciò è tanto più vero se si considerano le caratteristiche di molte componenti di SBN, complessivamente etichettabili come *legacy*, e fortemente ancorate a soluzioni tecniche caratterizzate da meccanismi di *self-reinforcing* o più semplicemente di *lock-in* e di *path dependency*. Con un po' di benevola ironia potremmo anche dire che l'Indice da mero dispositivo di smistamento delle comunicazioni come era stato concepito alle origini è diventato affine alla nozione heideggeriana di *Gestell*. In queste condizioni le scelte manageriali devono sfruttare finestre di opportunità, ossia sfruttare gli intervalli di tempo in cui si determinano condizioni e opportunità per mettere in campo interventi efficaci, o restringere il perimetro degli interventi sull'infrastruttura a ciò che presenta maggiori probabilità di successo in virtù di una minore complessità, di una minore *path dependency*, di una maggiore sostenibilità e così via, come ad esempio le componenti di front-end della infrastruttura SBN, centrale o periferica.



Progettare o gestire una infrastruttura, insomma, è molto diverso dal progettare uno specifico sistema informativo. Se poi consideriamo l'infrastruttura come una sorta di rete sociotecnica, a sua volta concepibile come insieme di reti sociotecniche (nel caso di SBN potrebbero essere i Poli), il grado di complessità diviene davvero notevole, sia perché tali reti oltre alle componenti tecniche includono per definizione individui e istituzioni, sia perché le tecnologie in questo contesto si definiscono in quanto operanti in ambienti popolati da elementi non tecnologici. Inoltre, le infrastrutture incorporano le rappresentazioni del funzionamento delle organizzazioni che supportano e quindi anche modalità organizzative, immagini cognitive, ordinamenti istituzionali. Interagiscono, in sostanza con situazioni complesse e alimentano "ecologie cognitive, organizzative e istituzionali". SBN, pertanto, non è un dispositivo puramente tecnico, ma include tali ecologie e qualunque intervento migliorativo o evolutivo dell'infrastruttura deve inevitabilmente tenere conto di tale complessità valutandone attentamente vincoli e opportunità. Le scelte manageriali effettuate devono essere pertanto inquadrare e valutate tenendo presente questo contesto, che impone un giudizioso pragmatismo progettuale e la restrizione del perimetro degli interventi possibili a vantaggio della loro probabilità di realizzazione entro i termini programmati. Da questo punto di vista una scelta manageriale quanto mai opportuna in considerazione dell'insufficienza di risorse umane con competenze di dominio presenti attualmente in ICCU è rappresentata dall'impiego, mediante accordi e convenzioni, di competenze esterne allocate presso i partner della cooperazione.

La seconda ragione è rappresentata dal servizio o, per meglio dire, dai nuovi servizi resi possibili. Da questo punto di vista il progetto si innerva su una traiettoria che ha inizio idealmente nel 1981, quando Angela Vinay e Michel Boisset sciolsero consapevolmente l'acronimo SBN in "Servizio bibliotecario nazionale" piuttosto che in "Sistema bibliotecario nazionale" con l'intento di definirne la missione – forse l'unica missione possibile – in un contesto culturalmente e istituzionalmente policentrico. Dopo tanti anni, SBN – dovremmo ricordarlo ai nostri interlocutori e a noi stessi – pur fra difficoltà e ritardi è ancora uno dei pochi progetti della P.A. che abbia prodotto risultati di lungo periodo, con un tasso di capillarità sul territorio nazionale paragonabile agli uffici postali. Certo i servizi bibliotecari hanno a che fare in buona parte con il catalogo, ma molto anche con la cultura organizzativa, col "servire il pubblico" come amava dire Luigi Crocetti. Da questo punto di vista il progetto attribuisce adeguata centralità alle componenti di front-end per consentire al pubblico una migliore esplorazione delle risorse della cooperazione e l'accesso a una maggiore ricchezza di contenuti. Il progetto prevede anche nuovi servizi per le biblioteche (SBN cloud e teca digitale) che mi paiono, per certi aspetti, la parte più significativa proprio per il possibile impatto sulle "ecologie cognitive, organizzative e istituzionali" di SBN. Ed è questa la terza ragione di importanza – forse la più rilevante – del progetto.




Immaginare servizi erogati direttamente dal centro dell'infrastruttura alle biblioteche è una scelta che qualche anno fa avrebbe suscitato molte discussioni e resistenze, perché avrebbe significato turbare i delicati equilibri dell'ecosistema istituzionale di SBN. Anche ora, pur a fronte dello sfilacciamento degli organi di governo SBN, si tratta di una scelta ardua, perché può entrare in competizione o in collisione con servizi che le Regioni, ma non solo, stanno realizzando o che intendono realizzare per i propri territori. Si tratta però di una scelta salutare, perché impone finalmente di affrontare due temi sinora rimasti ai margini dell'infrastruttura SBN. Il primo è quella della riconfigurazione dell'infrastruttura come sistema territoriale (nazionale) multiscala o multilivello; il secondo riguarda la riorganizzazione della cooperazione locale.

Pensare ad un sistema territoriale multiscala significa immaginare che l'ordinamento istituzionale dell'infrastruttura SBN dovrà seriamente riflettere sulla distribuzione dei servizi e della relativa gestione fra i diversi livelli istituzionali dell'infrastruttura; distribuzione che non potrà più avvenire come in passato in base a principi gerarchici, ma sulla base dei principi di convenienza (dove costa meno gestire i servizi) e di adeguatezza (dove è più efficace gestire i servizi). Significa anche immaginare l'azione degli organi di governo SBN nell'ambito di un'autentica governance multilivello. La riallocazione, inoltre, non dovrà riguardare solo le componenti tecniche dell'infrastruttura, ma anche le azioni a supporto delle politiche pubbliche riconducibili all'infrastruttura stessa.

La riallocazione dei servizi dell'infrastruttura SBN potrà consentire ai poli/reti locali di focalizzarsi maggiormente sui servizi non tecnologici, ma decisivi per assicurare la sostenibilità e il riconoscimento sociale delle biblioteche. Nonostante le diversità che contraddistinguono i territori italiani potremmo affermare che localmente occorre perseguire i seguenti obiettivi: a) rispondere alla necessità di far fronte con nuovi servizi, in un quadro di sostenibilità, alle attese delle comunità locali cercando di spostare gradualmente l'asse delle reti territoriali dai tradizionali servizi SBN ai servizi pretesi dalla contemporaneità; b) irrobustire il legame con le comunità locali, quale tratto identitario della biblioteca pubblica, attraverso la condivisione e la partecipazione; c) realizzare progressivamente ecosistemi regionali del patrimonio capace di dar luogo localmente a sistemi culturali integrati. Non è cosa da poco, ma dobbiamo crederci.

## **Alberto Petrucciani**


Sono lieto di essere qui anche in veste di componente del Comitato tecnico scientifico per le biblioteche e gli istituti culturali, di cui sono stato presidente nell'ultimo mandato e che si insedierà nella nuova composizione nelle prossime settimane, al termine delle procedure di rinnovo in corso per tutti gli organi consultivi del



Ministero per i beni e le attività culturali. Del nuovo Comitato sono presenti qui oggi anche la dottoressa Annamaria Malato e il dottor Marino Sinibaldi, mentre il dottor Luca Bellingeri, direttore della Biblioteca nazionale centrale di Firenze, non è potuto intervenire perché trattenuto dai suoi impegni istituzionali: anche questo interesse del nuovo Comitato, quindi, mi sembra di buon auspicio.

Non posso intervenire qui come se sentissi parlare per la prima volta di questo bel progetto, perché in verità sono stato in vario modo informato, e in parte coinvolto, nel suo avvio già da diversi mesi. Il dipartimento a cui appartengo nella Sapienza, recentemente trasformatosi in un grande Dipartimento di lettere e culture moderne, ha da tempo una convenzione con l'Istituto centrale per il catalogo unico, con il quale collabora per varie attività fra le quali, appunto, il progetto che si presenta oggi. Del progetto quindi conosco le finalità e l'impostazione e confesserò subito francamente che è un progetto in cui credo molto. Penso di poter dire che ci credo in maniera ponderata, per una serie di ragioni che cercherò di riassumere rapidamente. Intanto, ho fiducia nel gruppo che ci sta lavorando, con molto entusiasmo: è un gruppo con esperienze e competenze diverse, in cui si impara moltissimo confrontando i propri punti di vista e, appunto, le proprie esperienze. Anche le imprese che lavoreranno a questo progetto sono imprese qualificatissime e di grande esperienza nel campo specifico. Le premesse, da questo punto di vista, sono sicuramente ottime.

Personalmente ritengo che nel comprendere questo progetto, nel portarlo avanti e nel comunicarlo all'esterno, non sia opportuno, e possa essere fuorviante, focalizzarlo semplicemente dal punto di vista delle biblioteche e dei bibliotecari. È un progetto che certamente ha radici profonde nella lunga storia di cooperazione e di servizio delle biblioteche e dei bibliotecari italiani – una storia a cui personalmente, come molti sanno, sono legato, tanto da averne fatto negli ultimi vent'anni la mia principale area di ricerca – ma che vuole assumere un orizzonte più ampio e, soprattutto, vuole porsi dal punto di vista delle esigenze e dei comportamenti di oggi, del *pubblico della rete* e non semplicemente del pubblico delle biblioteche. Prenderei come premessa, perciò, una considerazione che faceva poco fa Simonetta Buttò: il grande *policentrismo culturale italiano*, che non ha uguali negli altri grandi paesi europei, e che non è soltanto il policentrismo delle cento città, ma è in realtà, e in maniera non meno significativa, anche un policentrismo istituzionale. Non mi riferisco soltanto all'articolazione della Repubblica, dallo Stato alle Regioni ai Comuni, ma alla varietà delle istituzioni e delle organizzazioni che fanno cultura, dalle università agli istituti ecclesiastici, e particolarmente al numero quasi infinito di istituti di cultura, centri studi, fondazioni, associazioni, "archivi" – come spesso impropriamente si designano – di questo o quel fenomeno, ambito, territorio. Alle volte viene da pensare che forse siano anche troppi, rispetto alle possibilità reali, in termini di risorse umane e finanziarie, di garantirne la funzionalità nel presente e la persistenza nel futuro. Comunque è tipica




dell'Italia questa straordinaria proliferazione di iniziative di cultura e di istituti, non semplicemente con riferimento al territorio, ma anche sotto il profilo della specializzazione degli interessi.

A questo panorama si unisce una tradizione, indubbiamente importante, di più di trent'anni di cooperazione in una rete nazionale delle biblioteche, la rete del Servizio Bibliotecario Nazionale, e in altri progetti: insomma un'esperienza di cooperazione solida e radicata in maniera capillare. Una cooperazione, però, che per ragioni anche contingenti è stata in gran parte incentrata sulla catalogazione del patrimonio bibliografico e non si è pienamente trasferita, alla stessa maniera, sul piano dei contenuti digitali, che naturalmente sono venuti dopo e sono spesso rimasti poco o per nulla integrati con le attività bibliografiche e catalogafiche. In Italia, in particolare, non sono stati mai realizzati progetti centrali di digitalizzazione di grandissime dimensioni come è avvenuto, ad esempio, in Francia. C'è invece, a ben vedere, una quantità veramente notevolissima di collezioni digitali disseminate, per lo più piccole ma spesso di singolare interesse e realizzate anche da istituti praticamente sconosciuti: ad esempio, una delle migliori collezioni di fotografie che si può vedere catalogata in SBN si trova in un paese che non avevo mai neanche sentito nominare, né avevo mai sentito nominare la sua biblioteca.

C'è quindi una grande quantità di materiale digitale interessante sparso qua e là, non solo nei siti delle biblioteche ma anche in quelli di istituti culturali di ogni tipo e direttamente nei siti istituzionali di regioni ed enti locali: ma questo materiale è difficilmente raggiungibile dal pubblico e difficilmente può essere conosciuto, finché rimane disperso. La rete Internet, come sappiamo, ha questa caratteristica: da una parte è facilissimo affacciarsi in rete, aprire un sito e caricare delle pagine web o altre risorse, ma poi si rischia di rimanere completamente invisibili. La visibilità, di fatto, è appannaggio di pochi grandi servizi, i servizi *leader* nei diversi ambiti. Ciascuno di noi può capitare alle volte, per caso o tramite un motore di ricerca, in siti di cui ignorava completamente l'esistenza, ma per lo più si rivolge, per una determinata esigenza, a un grande servizio che conosce e che costituisce, per quella esigenza, il suo punto di riferimento primario.

In questo contesto, mi auguro che il progetto che si presenta oggi abbia un pieno successo, faccia compiere veramente un grande passo avanti, e questo a mio parere potrà dipendere principalmente da due fattori: uno che è maggiormente sotto il controllo dell'Istituto centrale per il catalogo unico e un altro che invece riguarda tutti.

Cosa è sotto il controllo dell'ICCU? Di questo progetto fa parte, ovviamente, una componente tecnologica, molto innovativa, molto avanzata: non si tratta, per capirci, di adottare l'approccio semplicistico che purtroppo caratterizza molti progetti, in cui si prende una sorta di scatolone informatico, ci si butta una copia di materiali eterogenei di varia provenienza, lo si frulla un po', con qualche algoritmo, e




si spera che ne sortisca un servizio utile e importante. Questo modo di procedere non funziona, e in Italia dovremmo esserne particolarmente consapevoli perché dovremmo conoscere l'esperienza insoddisfacente che è stata fatta nei primi vent'anni di attività del Centro nazionale per il catalogo unico, impostando in maniera poco efficace i lavori finalizzati a realizzare un grande catalogo collettivo delle biblioteche italiane. Ma purtroppo gli errori spesso si ripetono, perché non ci si documenta a sufficienza sulle esperienze del passato. Come diceva un filosofo: «quelli che ignorano il passato sono condannati a ripeterlo». Le basi tecnologiche del progetto, quindi, sono state disegnate in modo molto innovativo, con le metodologie più avanzate, puntando a integrare informazioni e contenuti proprio alla fonte, dove nascono, e non con interventi successivi, che non possono mai essere pienamente efficaci e che inevitabilmente perpetuano frammentazioni e disallineamenti. Da questo punto di vista, la base di partenza è quindi sicuramente ottima, dà a mio avviso pieno affidamento.

A ciò dovrà corrispondere, e non è tanto facile, altrettanta capacità innovativa e originale nel progettare la comunicazione, le modalità d'uso, di ricerca e di restituzione del nuovo portale. È la questione che ha già toccato pochi minuti fa Giovanni Solimine: rendere facile, semplice e intuitivo l'accesso ai contenuti senza banalizzarlo. Come ha spiegato molto bene Simonetta Buttò, l'obiettivo è quello di riuscire ad accompagnare l'utente a realizzare anche ricerche complesse e precise ("volevo proprio questo"), in una maniera semplice e chiara, e proporre i risultati in una forma comprensibile, strutturata, in cui ciascuno possa scegliere rapidamente e con sicurezza ciò che lo interessa. Pensate a un supermercato in cui i prodotti siano distribuiti in modo casuale nell'intera superficie, in cui ad esempio gli yogurt non stiano tutti nello stesso frigorifero, ma sparsi, uno qui uno lì, nell'intero supermercato. Questo, purtroppo, è lo scoglio su cui s'infrangono quasi sempre i grandi progetti di questo tipo: viene diffusa la notizia della loro realizzazione e si va a vederli, si clicca un po' qua e là, si pensa "toh, quante belle cose che ci sono", ma poi non ci torna più, perché con quei servizi non si riesce a lavorare, si possono guardare un po' ma non funzionano come strumenti di ricerca quando se ne ha bisogno.

Anche da questo punto di vista penso che dovremmo guardare con attenzione alle esperienze delle istituzioni culturali, che in genere si sono affacciate nell'ambiente della rete un po' più tardi rispetto alle biblioteche, ma che negli ultimi anni hanno realizzato servizi di grande interesse anche per un pubblico generale, come mostrano le statistiche rispetto all'utilizzazione dei loro siti. Mi limiterò a citare il successo del sito della Treccani, che è il sito culturale più utilizzato in Italia, e quello dell'Accademia della Crusca, che oltre a mettere a disposizione parecchi materiali offre nelle sue rubriche ottime spiegazioni, scientificamente rigorose ma anche presentate con una comunicazione semplice e chiara, su tanti dubbi che possono sorgere nell'uso della nostra lingua.






Questo su cui mi sono soffermato è il primo elemento che serve per il successo di questa iniziativa. Ma il secondo è anche più importante. In un altro campo della biblioteconomia, quello della classificazione, si dice che un buon sistema di notazione, di codifica, non fa una buona classificazione, ma un cattivo sistema di notazione la compromette, la guasta. Analogamente, possiamo dire che se gli strumenti tecnologici, le interfacce ecc. non saranno validi, il servizio ne sarà pesantemente condizionato: ma il fatto che siano validi non basta, perché quello che effettivamente conterà saranno soprattutto i contenuti.

Questo, è evidente, dipende da tutti, dipende cioè dalla capacità delle biblioteche e degli istituti culturali di produrre e di mettere a disposizione quei contenuti, nella più grande quantità possibile e a un buon livello di qualità.

In questa prospettiva mi sembra opportuno ricordare alcune criticità importanti. Per prima, una lacuna informativa molto grave, ossia il fatto che, a oltre trent'anni dall'avvio della rete SBN, buona parte del patrimonio delle grandi biblioteche non è ancora registrato in Internet, non tanto per arretrati di catalogazione, ma perché non è stato ancora recuperato nei sistemi informatici: è ancora registrato soltanto in cataloghi a schede, o perfino manoscritti, o in registri ed elenchi che si possono consultare soltanto recandosi materialmente nelle sale delle singole biblioteche.

Un altro importante ritardo riguarda la messa a disposizione in rete di materiale digitale di cui le biblioteche dispongono già (e che può essere, dal punto di vista legale, liberamente messo a disposizione di tutti): tanto materiale digitale giace nei computer delle biblioteche, o in cassette e armadi di CD e altri supporti, materiale realizzato per i motivi più diversi (per mostre e cataloghi, per richieste degli utenti, per copie di sicurezza, per progetti vari) e poi non riutilizzato e non messo a disposizione di tutti. Alcune biblioteche hanno fatto un'operazione semplicissima, ma preziosa e che tutti dovrebbero imitare: l'ufficio riproduzioni della biblioteca ha recuperato e messo in rete, nei limiti ovviamente delle norme sul diritto d'autore, tutto quanto aveva realizzato, negli ultimi anni, su richiesta di singoli utenti o per i servizi interni.

Terza e ultima criticità che vorrei segnalare, di non minore importanza, è che il materiale digitale che è stato messo a disposizione, nel sito della biblioteca o di qualche progetto speciale, molto spesso non viene registrato nel catalogo, e quindi sfugge proprio alle ricerche di chi ne ha bisogno, di chi lo sta cercando. È in primo luogo a chi sta cercando nella rete SBN un certo libro, un certo periodico o un altro materiale, che dovremmo segnalare subito, in evidenza, che è disponibile in forma digitale, accessibile per tutti a distanza tramite la rete. Troppo spesso, invece, le collezioni digitali non sono registrate nei cataloghi generali, e in molti casi è difficile trovarle anche se, invece dei cataloghi, si ricorre ai motori di ricerca: i motivi di queste difficoltà di accesso sono diversi e di natura tecnica (software proprietari chiusi ai motori di ricerca, o che comunque non si prestano alla loro esplo-




razione, uso del solo formato immagine senza testo incorporato, frammentazione degli oggetti digitali, ecc.) e non è possibile ovviamente soffermarsi qui.

Queste criticità vanno tenute presenti e vanno affrontate, ma sono superabili senza eccessive difficoltà, tenendo presenti gli esempi migliori: si potrà in questo modo potenziare notevolmente l'offerta di materiali che il nuovo servizio metterà a disposizione.

Vorrei sottolineare, a questo proposito – lo indica il titolo dell'iniziativa di oggi e lo diceva anche la dottoressa Buttò – che dobbiamo abituarci a ragionare non solo in termini di biblioteche, ma di *biblioteche e istituti culturali*. E dovremmo anche cercare di superare, a mio parere, questa visione un po' troppo verticale, per compartimenti: c'è il sistema delle biblioteche, con i suoi oltre trent'anni di esperienza, c'è il sistema archivistico, anch'esso molto dinamico e organizzato, c'è il sistema museale, anche se ai suoi primi passi, d'accordo. Ma in realtà l'Italia è piena di istituti, che spesso sono tra quelli che producono di più anche in rete, che non sono semplicemente riconducibili a uno dei tre settori, istituti che hanno in genere un archivio, una biblioteca e spesso anche collezioni d'interesse museale, ma non sono né una biblioteca, né un archivio, né un museo: sono un centro, un istituto, un'associazione che fa ricerca, ha anche libri e altri materiali. Dobbiamo appunto ragionare in questi termini: di istituti culturali che producono contenuti, che fanno e non semplicemente conservano, e che non sono classificabili semplicisticamente in uno di quei tre settori.

Dobbiamo allargare la nostra visuale anche in termini di materiali. Per noi il libro è sempre il libro, resta uno strumento assolutamente fondamentale per la formazione dell'uomo, e cioè dell'uomo e della donna, del bambino e della bambina, del ragazzo e della ragazza. Però le biblioteche trattano anche altro materiale, e non da oggi: già nel Seicento le biblioteche erano piene di disegni, di carte geografiche e spesso anche di monete e medaglie, di reperti naturalistici, di strumenti scientifici, insomma delle cose più diverse che si riteneva potessero essere utili per le loro funzioni al servizio degli studi. Le biblioteche, insomma, non sono multimediali dalla settimana scorsa, lo sono sempre state: ma dobbiamo valorizzare di più e meglio questo aspetto, che a mio avviso è stato un po' trascurato. Non so quante persone sappiano, ad esempio, che patrimonio enorme di materiale sulla canzone italiana si raggiunge già adesso tramite SBN: probabilmente il pubblico lo ignora, altrimenti andrebbe a vederlo.

Ci auguriamo, perciò, il migliore sviluppo di questo progetto, il raggiungimento dei risultati che si propone: ci auguriamo, soprattutto, che tutti gli istituti, le biblioteche e le istituzioni di cultura, tutti quelli che producono contenuti culturali, contribuiscano al progetto in maniera che il cittadino possa trovare ciò che cerca e ciò che desidera, ciò che può interessarlo e anche ciò a cui non aveva ancora pensato. Perché la finalità dei servizi culturali non è semplicemente di soddisfare dei bisogni, delle esigenze, ma di stimolarli, di farli sorgere, di farli crescere.



Di questa finalità vorrei sottolineare, in conclusione, anche il significato politico, non solo per la formazione culturale del cittadino, ma anche per un principio di correttezza: il cittadino ha il diritto di trovare, rivolgendosi a un servizio nazionale integrato, a una sorta di “sportello unico”, tutto quello che è stato realizzato con denaro pubblico, e che non deve rimanere chiuso all’interno degli istituti o, anche quando viene riversato in rete, praticamente invisibile e introvabile. Il cittadino deve poter trovare facilmente tutto ciò che è stato realizzato con denaro pubblico o dalle istituzioni pubbliche, e tutte le cose molto belle e interessanti che sono state realizzate da un grande numero di istituti di ogni genere e in tutto il paese, per lo più poco note e difficili da conoscere, devono poter essere facilmente trovate da tutti, come chi le ha prodotte certo desidera. La finalità complessiva del progetto è quella di far incontrare queste due esigenze.


## **Gino Roncaglia**

Grazie innanzitutto per l’invito. Vorrei intervenire brevemente partendo da un punto di vista specifico, che è quello dell’integrazione di risorse e contenuti attraverso strumenti di ricerca e scoperta, i cosiddetti “discovery tool”: mi pare infatti che questa sia l’idea potenzialmente più rilevante fra quelle alla base del vostro progetto.

Tra gli obiettivi fondamentali dei discovery tool, oltre a quello di aiutare nella ricerca dei contenuti, c’è quello di connettere risorse informative fra loro e collegarle alle esigenze informative degli utenti: il motto “semplicemente connettere” («only connect») di E.M. Forster sintetizza efficacemente una delle raccomandazioni che tutti i lavori di digitalizzazione nell’ambito del mondo della pubblica amministrazione dovrebbero sempre tenere presente. Anche perché un limite che purtroppo ha spesso caratterizzato questi lavori, e che auspicabilmente stiamo imparando a superare, era proprio quello della frammentazione: frammentazione degli strumenti e non solo dei progetti.

Costruire discovery tool efficaci, strumenti di ricerca che permettano di collegare fra loro i contenuti e di lavorare su basi dati integrate, recuperando risorse in maniera il più possibile semplice e amichevole, è fondamentale anche dal punto di vista dell’uso e del riuso dei contenuti che sono stati digitalizzati; e quindi è anche un modo per valorizzare la spesa che c’è dietro i progetti di digitalizzazione. Una spesa spesso rilevante, ma che risulta sensata, più che sensata, se i risultati di questi progetti possono essere facilmente recuperati e riutilizzati.

Per fornire un esempio, possiamo guardare a questo aspetto da un punto di vista particolare: quello del mondo della scuola, che ha un enorme bisogno di poter recuperare in maniera facile contenuti culturali affidabili e di qualità, per poterli riutilizzare in forme il più possibile organizzate. I discovery tool partono dall’idea della ri-




cerca del contenuto, ma sempre più spesso dobbiamo imparare a passare dalla ricerca del singolo contenuto alla capacità di riusare il contenuto che abbiamo trovato all'interno di percorsi strutturati, di attività formative, di progetti di ricerca complessi. E questo è particolarmente importante proprio per il mondo della scuola: passare dalla singola ricerca alla possibilità di costruire percorsi. Un passaggio che non deve essere fatto necessariamente dallo stesso discovery tool: non è detto che lo strumento di ricerca debba essere immediatamente anche lo strumento per costruire percorsi; ma il discovery tool deve fornire contenuti in una forma adatta a permettere e facilitare il lavoro di costruzione di percorsi, anche su piattaforme diverse e utilizzando strumenti diversi.

Fra le attività che ho svolto per Rai Cultura, ho lavorato per un certo periodo a un progetto al quale tengo molto, e che spero possa essere ripreso e ulteriormente sviluppato in futuro: la realizzazione di una piattaforma per la costruzione di *lesson plan*, utilizzabile direttamente da parte dei docenti, che renda possibile recuperare risorse e contenuti da fonti autorevoli e validate, e costruire, a partire da queste risorse, percorsi strutturati che possono servire per impostare e costruire la lezione. Piattaforme di questo tipo hanno bisogno, per poter funzionare bene, di agganciarsi a discovery tool adeguati.

Oggi si parla molto di *flipped classroom*, una metodologia didattica che presuppone la capacità di dare agli studenti dei materiali validi, accessibili e ben organizzati per studiare e sui quali sviluppare in seguito il lavoro di classe: discussioni, *debate*, produzione di contenuti integrativi... E ancora una volta, per poter fare bene questo tipo di attività (su cui tanto aveva scritto, ad esempio, Tullio De Mauro), occorre avere alle spalle buoni strumenti di scoperta dei contenuti, assieme a strumenti che permettano di organizzarli e naturalmente assieme a competenze specifiche nel campo dell'information literacy; così come occorre avere all'interno delle scuole degli spazi e dei tempi destinati allo svolgimento di attività didattiche orientate agli interessi e non solo ai curricula.

Mi sono impegnato molto sul fronte delle biblioteche scolastiche perché sono convinto che la biblioteca scolastica debba essere oggi proprio una delle sedi in cui fare queste attività. E la biblioteca scolastica ha bisogno naturalmente di libri e risorse informative, ma ha anche bisogno di strumenti, tradizionali e digitali, e ha bisogno di spazi e di buona organizzazione degli spazi. Le nostre scuole sono fondamentalmente organizzate intorno a due tipologie di spazi e di tempi: da un lato la classe e il gruppo classe, dall'altro le discipline, accompagnate dalla rigida scansione di tempi e spazi disciplinari. Siamo invece molto indietro sul fronte di quelli che dovrebbero essere anche nel mondo della scuola i terzi spazi e i terzi tempi, legati all'approfondimento degli interessi individuali, alla ricerca e alla documentazione, alla capacità di recuperare informazioni e contenuti validati, accompagnati dai relativi metadati e dunque da informazioni affidabili e standardizzate sulla fonte da cui sono stati tratti.



I terzi spazi e i relativi strumenti, compresa la disponibilità di buoni discovery tool, sono legati appunto non solo alla didattica disciplinare ma anche a quella trasversale, alla capacità di rispondere efficacemente a bisogni informativi legati agli interessi individuali, alla ricerca estemporanea, all'approfondimento di una lettura: penso sempre alla biblioteca scolastica e a molte delle attività che si possono fare a scuola in ambito culturale come legate in primo luogo agli interessi, fondamentali per la motivazione allo studio e capaci anche di scardinare almeno in parte le gabbie troppo strette del gruppo classe o delle singole discipline. Un lavoro interdisciplinare che può mettere insieme studentesse e studenti di classi diverse, entro limiti ragionevoli anche di età diverse, intorno a un lavoro di ricerca comune.

Per lavori di questo tipo oggi il mondo della scuola ha anche strumenti di sostegno organizzativo ed economico: in primo luogo i PON, che potrebbero e dovrebbero rappresentare anche un veicolo importante di innovazione metodologica. Fra i PON recenti, ad esempio, quello sui beni culturali (ho seguito le attività svolte da alcune scuole che hanno ottenuto il relativo finanziamento) ha offerto un'occasione straordinaria di scoperta dei beni culturali del territorio, di mappatura, di riuso: ma questo presuppone sempre la disponibilità di strumenti efficaci.

Credo e spero che strumenti come quello a cui state lavorando possano essere pensati anche per facilitare questo tipo di lavoro da parte delle scuole, e più in generale da parte del cittadino. Perché quando si parla della possibilità di approfondire interessi personali costruendo percorsi, non limitandosi quindi al recupero della singola risorsa ma facilitando la capacità di collegare tra di loro risorse diverse, si parla anche di uno sviluppo delle competenze informative e degli stessi strumenti di rete nella direzione che considero oggi più importante: la capacità di superare l'eccessiva frammentazione, andando progressivamente a costruire contenuti più complessi, più strutturati, più autorevoli e, dal punto di vista della formazione, anche più utili, in particolare rispetto alla costruzione delle competenze di cittadinanza, che dovrebbero sempre rappresentare una delle nostre preoccupazioni fondamentali.

**Marino Sinibaldi**, direttore di Rai Radio Tre, ha concluso l'incontro formulando gli auguri per il progetto avviato dall'ICCU, che va nella direzione della maggiore diffusione, nello spazio della rete, di contenuti di qualità e nello stesso tempo comunicati in modo accessibile, sorretti da cataloghi integrati, autorevoli e credibili.

**Saggi**



# 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 bilanci<sup>1</sup>

**E**ra 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 digitale<sup>2</sup>. 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 merita<sup>3</sup>.

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 progetto<sup>4</sup>. 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 comunicazione<sup>5</sup>.

<sup>1</sup> 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.

<sup>2</sup> 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.

<sup>3</sup> 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 attuato 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.

<sup>4</sup> 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.

<sup>5</sup> 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](http://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 Bruttocao 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 server<sup>6</sup>. 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 Mac<sup>7</sup>. Venne messo a punto un primo contratto standard per il conferimento dei materiali<sup>8</sup>. 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 Carofiglio<sup>9</sup>, dimostrarono immediatamente che gli scenari di acquisizione e le tipologie di fondi potevano essere difformi 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

<sup>6</sup> 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.

<sup>7</sup> 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.

<sup>8</sup> 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.

<sup>9</sup> 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).

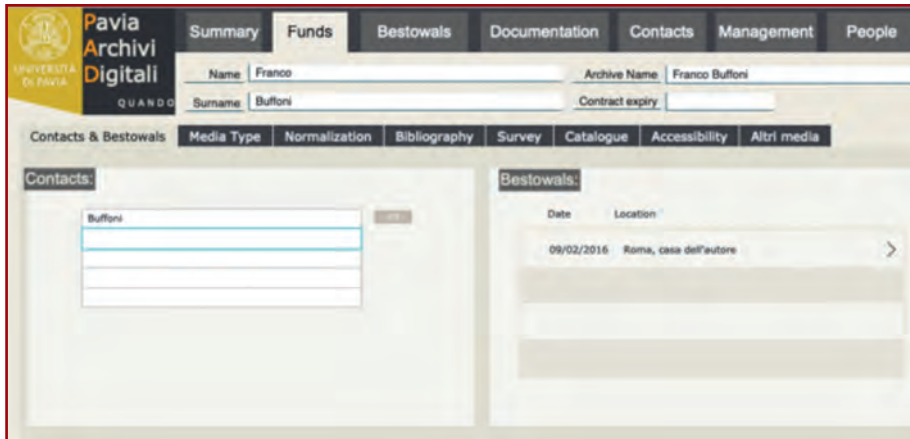


Figura 1. Il gestionale QUANDO

qualità e delle fasi operative (QUANDO)<sup>10</sup>, 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 archivi<sup>11</sup>.

È con questo nuovo assetto che vennero realizzati, a partire dal 2015, i nuovi conferimenti<sup>12</sup>. Il primo ad aderire fu Paolo Di Paolo, seguì poco dopo Francesco Pecoraro<sup>13</sup>. 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 complesso<sup>14</sup> 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 contenuto<sup>15</sup>.

<sup>10</sup> Cfr. Weston – Carbé – Baldini 2016.

<sup>11</sup> Cfr. Weston – Carbé – Baldini 2017. Si segnala inoltre Feliciati et al. 2018.

<sup>12</sup> 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.

<sup>13</sup> 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.

<sup>14</sup> Carbé 2017.

<sup>15</sup> 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 fondi<sup>16</sup>.

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 Manoscritti<sup>17</sup>, 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 Pavia<sup>18</sup>.

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 cittadinanza<sup>19</sup>.

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.

<sup>16</sup> 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.

<sup>17</sup> Il conferimento avvenne il 9 febbraio 2016 presso la casa dell'autore, in presenza di Paul Gabriele Weston e Emmanuela 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](https://lombardiarchivi.servizirl.it/groups/UniPV_CentroManoscritti/fonds/47480)>.

<sup>18</sup> 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).

<sup>19</sup> 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 file<sup>20</sup>

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).

<sup>20</sup> 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.

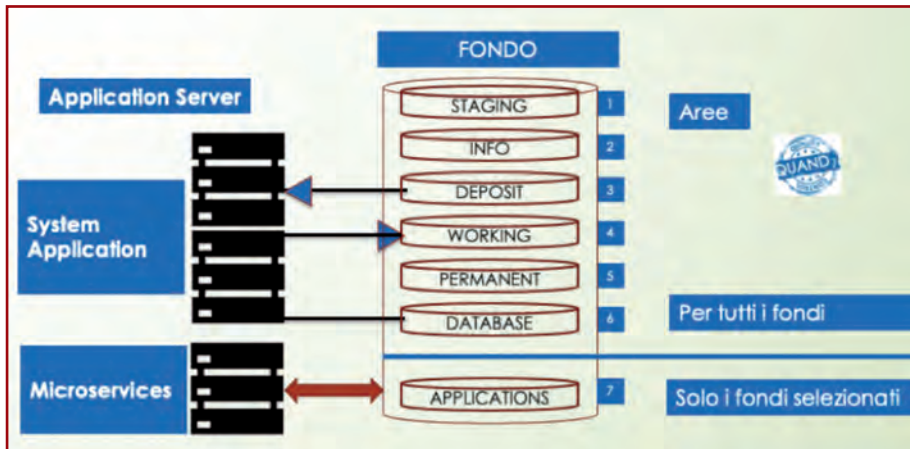


Figura 2. Le aree in cui si articola la struttura del sistema di archiviazione

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 strutturati<sup>21</sup>. 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.

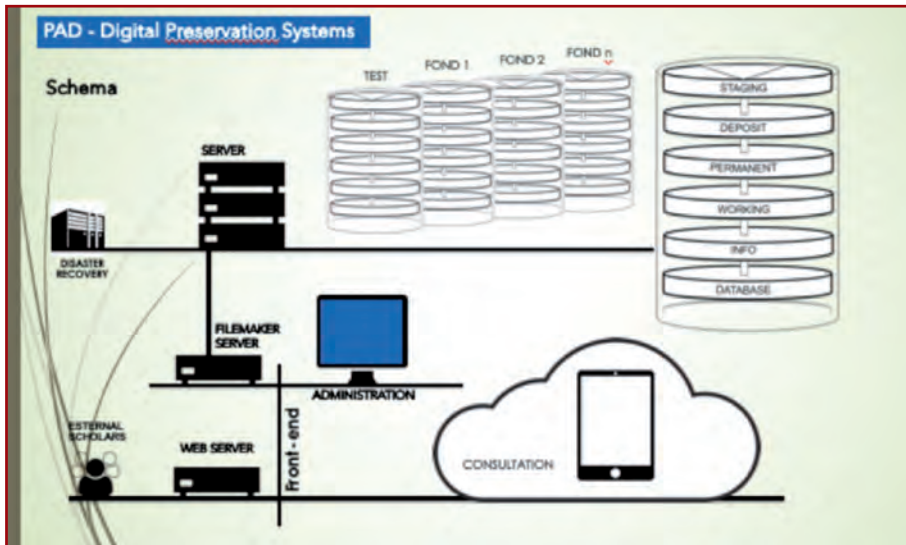


Figura 3. Le componenti del sistema di archiviazione e di accesso

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 (*DataMining*, *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-

<sup>21</sup> 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-

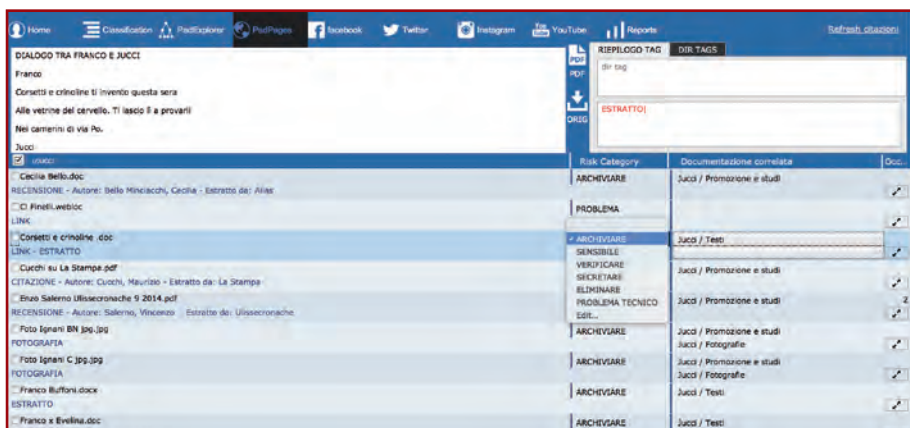


Figura 4. Attraverso il menu a tendina è possibile scegliere tra le categorie di rischio già preinserte. La scelta è facilitata dalla funzione di preview



- 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.

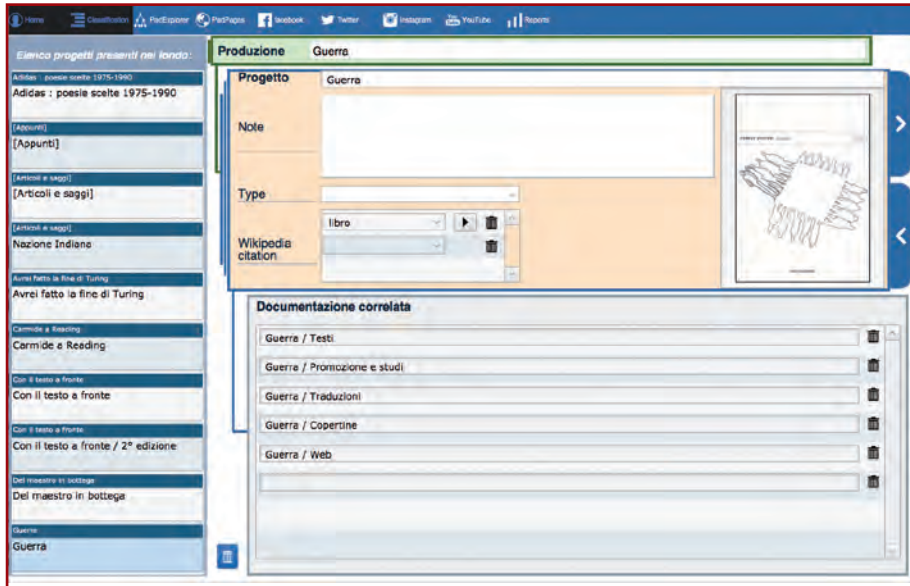


Figura 5. Albero a livelli della produzione documentale dell'autore. Ogni documento del fondo è inserito tra una particolare Documentazione correlata, legata alla Produzione e al Progetto. Nella parte a sinistra è visualizzato l'elenco sintetico dei progetti presenti nel fondo

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).

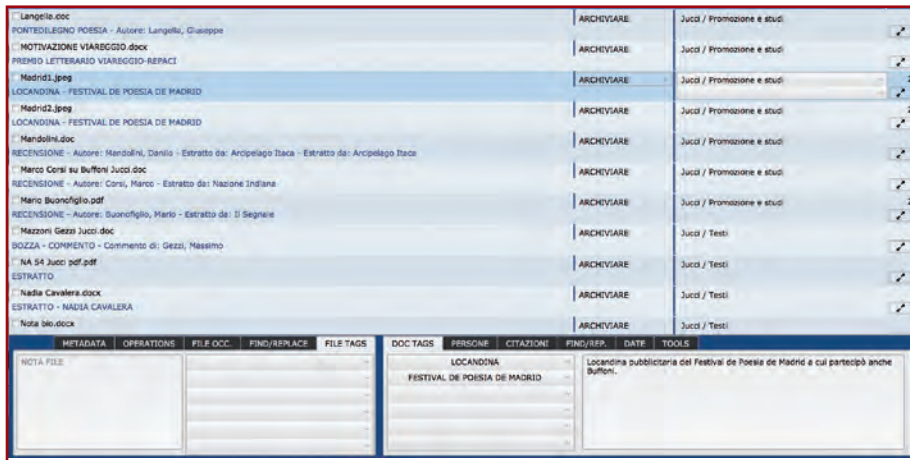


Figura 6. Nella parte bassa dello schermo sono presenti dei menu a tendina da cui è possibile selezionare dei tag preinseriti oppure inserirne di nuovi. Inoltre ogni file può essere corredato da note a campo libero

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).

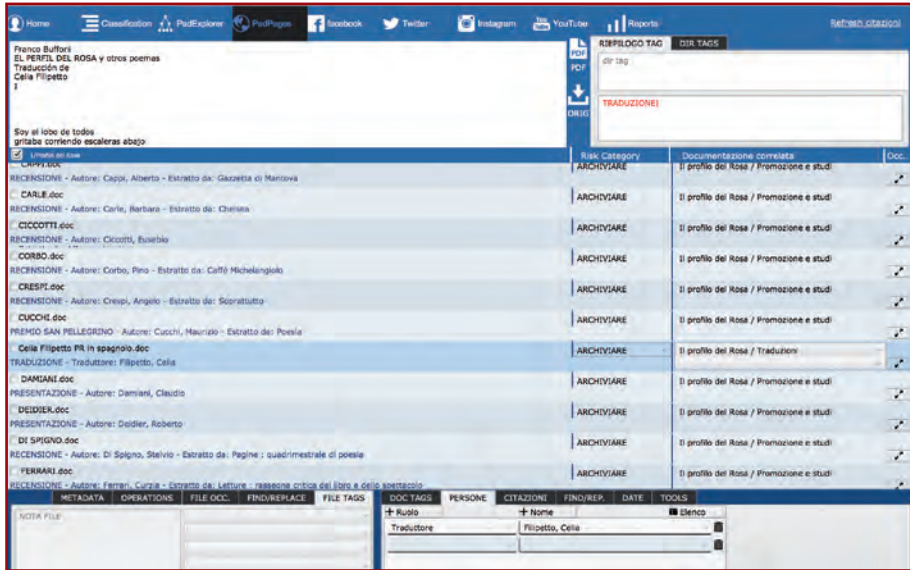


Figura 7. La scheda *Persone* permette di legare al file il nome di una persona diversa dall'autore principale, specificandone il ruolo

### *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 digitali<sup>22</sup>.

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

Il sito web di Franco Buffoni<sup>23</sup> 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.

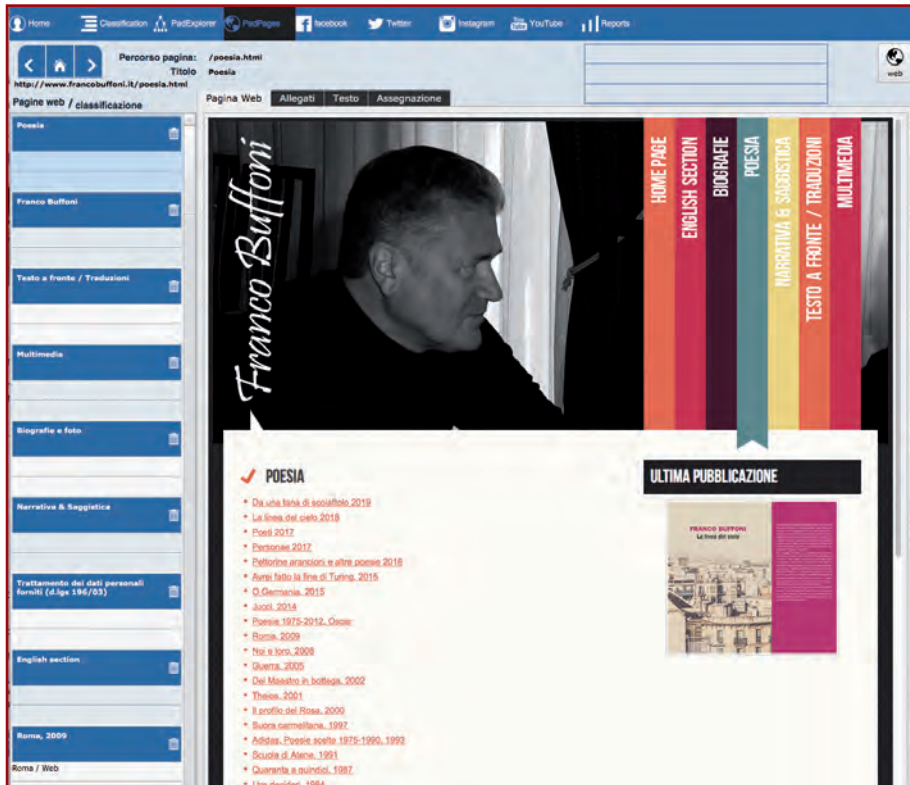


Figura 8. Navigazione offline del backup del sito web

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

<sup>23</sup> 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 il *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.

The screenshot shows a web browser window displaying a digital archive. The address bar shows the URL <http://www.francobuffoni.it/poesia/roma.html>. The page title is "Roma, 2009". The interface is organized into several sections:

- Left Sidebar:** A list of files with their titles and categories. Visible items include "Roma, 2009", "I tre desideri, 1984", "Del Maestro in bottega, 2002", "Adidas, Poesia scelta 1975-1990, 1993", "Nui e loro, 2008", "Buona carnefatta, 1997", "Aucci", "Riflessioni sull'opera", "Quaderni di Poesia Italiana Contemporanea", and "Il profilo del Roma, 2000".
- Main Content Area:** A table-like structure showing file details and previews.
  - File 1:** "Pardonanotte-Dedalus" (file: `files/pdf/dedalus_clas_1209.pdf`). Preview: "Il. Lappia. Riportando tutto a casa. Enaudi. 52 vetri. L. Pugno. Quando verrà. minimum Etc."
  - File 2:** "Valerio Magrelli, prefazione a ROMA" (file: `files/pdf/roma_valerio_magrelli_prefazione.pdf`). Preview: "Valerio Magrelli, prefazione a ROMA. L'ultima raccolta di Franco Buffoni (nato a Gallarate nel 1948 e considerato fra i maggiori poeti italiani della sua generazione) nasce da uno spostamento geografico e antropologico, lo spostamento con cui questo 'vecchio lombardo esente' 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 transito in una capitale insieme classica e inventiva, pagana e islamizzata, cristiana e disomnista. Proprio aggiungendo quei suoi due dialetti, egli dedicherà alla Città Eterna pagine che costano all'incirca quanto alle circostanze letterarie della riveducibilissima dell'antico."
  - File 3:** "L'anticipazione apparsa su Nuovi Argomenti 43, luglio-settembre 2008" (file: `files/pdf/roma_anticipazione_nuovi_argomenti.pdf`). Preview: "ROMA - L'ANTICIPAZIONE APPARSA SU NUOVI ARGOMENTI 43. LUGLIO-SETTEMBRE 2008. Nel gioco decorativo la battaglia. Diaccia dagli arazzi. Al rianzo di sistema dei ricavi."
  - File 4:** "L'anticipazione su Il Parlar franco con testo di M. Cohen" (file: `files/pdf/manuel_cohen_su_roma.pdf`). Preview: "FRANCO BUFFONI IL CIELO CHE SI MUOVE SOPRA ROMA. di Manuel Cohen, Il Parlar franco, gennaio 2009."
  - File 5:** "L'anticipazione su Metamorfosi con testo di F. Matteani" (file: `files/pdf/matteani_su_roma.pdf`). Preview: "Su ROMA di Franco Buffoni. Vivere in una grande, antica città significa spesso attraversare le epoche. Immaginate gli edifici antenati nella parte dei corredi e di cui sui loro tronconi, sconosciuti abitanti. Camminando, per le strade le voci che ascoltate escono direttamente dalle mura, dai fregi e perfino dal sottosuolo dove dormono altre città, con lo stesso nome della vostra. Nei testi che compongono Roma il poeta Franco Buffoni si appia per la città eterna, dove l'eterno più che un tempo senza fine sembra indicare la coincidenza delle ere. L'imprevedibile di ciò che è stato, che si trasforma in questi ultimi momenti. Così, così sono antichissimi, sopravvissuti in modo diverso, con le loro, costruzioni in."
  - File 6:** "L'anticipazione nella traduzione in spagnolo di Caterina Castellary" (file: `files/pdf/castellary.pdf`). Preview: "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 appesantendosi. Come il lavoro precedente, mi resi conto che l'antichità e il ribatte di una città mitica e attuale, dove alcuni personaggi approdati nella Roma di ieri - da Pasolini a Galilei al Pinturicchio, da Leopardi a Sandro Penna a Keats - interagivano 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-

Verificato: X Sì No

**WEB** Similarity: 30% **NATIVO**

Titolo del link:  
Bando "Premio Marazza 2016"

Filename  
files/pdf/bando\_marazza\_2016.pdf

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 1946 e considerato fra i maggiori poeti italiani della sua generazione) nasce da uno spaesamento geografico e antropologico, lo spessamente con cui questo "vecchio 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 deciderà 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, colti), 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.

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 Levenshtein<sup>24</sup>, 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 termine<sup>25</sup>.

### 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'archivio, 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	
<b>Godwin Club Italia</b>	
<b>Estratto da</b>	
1/Carmide a Reading	
2003 - D'Elia, Gaetano - Letteratura e ipocrisia... (Godwin Club Italia - apr 2003).doc	
<b>Avanti!</b>	ISSN 0391-6685
<b>Estratto da</b>	
1/Casa Palestro	
Broda Minello su Casa Pal.docx	
1/UVCI	
Intervista Avanti ABM.docx	
1/RDMA	
alberto toni pdf.pdf	
1/7hs-Mstr-Sci-Sra-Addis-Q-3D-AO	
TONI.doc	
<b>Le Cronache del Salernitano</b>	
<b>Estratto da</b>	
1/Casa Palestro	
Enzo Salerno.pdf	
<b>Le Monde diplomatique</b>	ISSN 1127-3243
<b>Estratto da</b>	
1/Casa Palestro	
Finelli Casa pal.pdf	
1/LAICO ALFABETO	
FANELLI .doc	
1/LAICO ALFABETO	
MONDE.jpg	
1/Q Germania	
Finelli su D Germania.jpg	
1/SERVO BVRON	
Finelli Monde .pdf	
1/Turing	
Finelli su Turing.pdf	

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

<sup>24</sup> 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](https://www.researchgate.net/publication/246951886_Dictionary_of_Algorithms_and_Data_Structures)>.

<sup>25</sup> 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	
<b>O Germania</b>			
<b>O Germania</b>			
<b>O Germania / Copertine</b>			
1/O Germania / cop_ogermania_1.jpg	[JPEG]	COPERTINA	
1/O Germania / O Germania   Cover.pdf	[PDF]	COPERTINA	
<b>O Germania / Promozione e studi</b>			
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
<b>Abbà, Marta</b>	
<b>Autore</b>	1/Casa Palestro Marta ABBÀ.docx
<b>Accolla, Dario</b>	VIAP ID: 118146094160700330568
<b>Autore</b>	1/Casa Palestro Accolla su Casa.docx
<b>L'OSCAR</b>	scrittore v fu.doc
<b>Adolgisio, Armando</b>	
<b>Intervistatore</b>	1/PIU look, padre Nascosto a Adolgisio
<b>Agostinalli, Mattia</b>	
<b>Autore</b>	1/776-Mat-Scl-Sra-Adda-Q-3D-AD AGOSTINALLI.doc
<b>Albisani, Sauro</b>	VIAP ID: 69029137
<b>Autore</b>	1/Brollo de Rosa ALBISANI.doc
<b>Alfano, Giancarlo</b>	VIAP ID: 64225180
<b>Autore</b>	1/L'ART - SONGS ALFANO L'ART.doc
<b>Amerio, Andrea</b>	VIAP ID: 108073710
<b>Autore</b>	1/PIU look, padre

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	
<b>Avrei fatto la fine di Turing</b>			
<b>Avrei fatto la fine di Turing</b>			
<input type="checkbox"/>	Avrei fatto la fine di Turing / Copertine		
<input type="checkbox"/>	Avrei fatto la fine di Turing / Promozione e studi		
<input type="checkbox"/>	Avrei fatto la fine di Turing / Testi		
<input type="checkbox"/>	Avrei fatto la fine di Turing / Web		
<b>Carmide a Reading</b>			
<b>Carmide a Reading</b>			
<input type="checkbox"/>	Carmide a Reading / Promozione e studi		
<input type="checkbox"/>	Carmide a Reading / Testi		
<input type="checkbox"/>	Carmide a Reading / Web		
<b>Con il testo a fronte</b>			
<b>Con il testo a fronte</b>			
<input type="checkbox"/>	Con il testo a fronte / Copertine		
<input type="checkbox"/>	Con il testo a fronte / Promozione e studi		
<input type="checkbox"/>	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 studio 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. Allegranza, 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](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\\_](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*. «Bibliothecae», 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.

# Through the Looking Glass. Cultural Heritage Custodians to Populate the Mirrorworld

**Susan Hazan**

Senior Curator of New Media and Head of the Internet Office - The Israel Museum, Jerusalem

*This contribution is a wake-up call to custodians of cultural heritage to get ready and prepare for the next platform – the 3D and 4D experiences of the nascent Mirrorworld. Built upon the first generations of online experience, the sprawling communication platform of the World Wide Web, and driven by the ubiquitous reach of social networks and online gaming, the Mirrorworld will breach the glass ceiling of 2D experiences to draw us into a fully functioning 3D environment. Public institutions that manage cultural heritage for visitors, such as galleries, libraries, archives and museums (GLAMS), are ideally placed to seed the spatial communications network that will drive the new platform. Despite the formidable challenges to be surmounted, it is up to custodians of cultural heritage to become proactive in building the Mirrorworld rather than to wait for third party actors to advance commercial interests with imitation culture, simulated histories, or made-up stories. As gatekeepers of cultural artifacts, GLAMs have the capability, the experience, and even the responsibility to tell their stories forcefully, accurately and in fully-functioning 4D.*

## What is the Mirrorworld?

The Mirrorworld doesn't yet fully exist, but it is coming. Someday soon, every place and thing in the real world - every street, lamppost, building and room - will have its full-size digital twin in the Mirrorworld. For now, only tiny patches of the Mirrorworld are visible through AR headsets. Piece by piece, these virtual fragments are being stitched together to form a shared, persistent place that will parallel the real world<sup>1</sup>. According to Shoshanna Zuboff

«Among high-tech leaders, within the specialist literature, and among expert professionals there appears to be universal agreement on the idea that everything will be connected, knowable, and actionable in the near future: ubiquity and its consequences in total information are an article of faith»<sup>2</sup>.

<sup>1</sup> Kevin Kelly, *AR Will Spark the Next Big Tech Platform - Call It Mirrorworld*, «Wired», 12 February 2019, <<https://www.wired.com/story/mirrorworld-ar-next-big-tech-platform/>>.

<sup>2</sup> Shoshanna Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, New York: Public Affairs-Hachette Book Group, 2019, p. 223.



Fascinating? Yes. Scary? Very. But, while Zuboff has identified this as an unprecedented form of power which she calls «surveillance capitalism» this chapter is concerned with what this will mean for the virtual global patchwork of human creativity and for custodians of cultural heritage. How will the thousands of institutions scattered around the globe, who together hold disconnected slivers of our shared history and globally distributed traces of our common artistic endeavor, become active players? How will all these fragments come together, and how will they become accessible through the looking glass?

We already use automatic recognition systems to identify objects in the world around us as we move through our streets and visit parks, museums, and zoos. All we need to do is to download the right app and open the camera on our mobile phones. Point and shoot the plants and flowers in your garden or on an exotic island while on holiday and you can instantly identify their species and genus at the click of a button. A simple search through the many apps available will turn up several options – *FlowerChecker*, *NatureGate*, *PlantSnapp*, *Plantifier*, *LikeThat Garden*, *Leafsnap*, *iPflanzen* for starters. *NatureGate* will also help you identify birds, fish and butterflies<sup>3</sup>. The same goes for artworks – all instantly recognisable at the push of a button. It is all already at your fingertips. Even as I write this article, *Word* is analysing the illustrations in my own text and prompting me to use suggested captions. It is surprisingly accurate but actually makes me feel kind of redundant and somewhat inadequate – a taste, perhaps, of things to come. «The major difference in our lives will occur» – according to Yuval Noah Harari – «as more and more machines read our potential “when tech knows you better than you know yourself”<sup>4</sup>» – now that is really something to worry about! «Everything connected to the internet will be connected to the Mirrorworld. And anything connected to the Mirrorworld will see and be seen by everything else in this interconnected and highly compelling environment<sup>5</sup>».

We can begin to sense the nascent world when we move through Google Earth<sup>6</sup>. Accessible to all, we “walk” through this wallpaper-world, crisscross the digitally reincarnated streets of cities, and gaze at the facades of buildings. We “wander” the breathtaking rolling landscapes of our natural environment and “visit” places we have only dreamed of. This is just one of the stepping stones that lead us to the Mirrorworld, yet, while we are still thrilled by the *smorgasbord* of virtual space

<sup>3</sup> See *The Best Plant Identification Apps*, «The Plant Guide», 23 August 2017, <<https://theplant-guide.net/2017/08/23/best-plant-identification-apps/>>.

<sup>4</sup> Nicholas Thompson, *When tech knows you better than you know yourself*, interview with Tristan Harris, «Wired», 10 April 2018, <<https://www.wired.com/story/artificial-intelligence-yuval-noah-harari-tristan-harris/>>.

<sup>5</sup> K. Kelly, *AR Will Spark the Next Big Tech Platform - Call It Mirrorworld*, cit.

<sup>6</sup> <https://www.google.com/earth/>.

– albeit in 2D – the Mirrorworld is still not quite here. We still need to make one more conceptual step – the magic leap from a 2D to 3D world in order to reach the full potential of the world that is awaiting us.

At the same time, social networks have an equally impressive global reach. The frontrunners, usually listed as Facebook, Twitter, Instagram, WeChat and Snapchat – as well as all the myriad wannabees – now connect people 24/7, usually from the palm of our hand, to a seamless, global place. While some claim this cuts out extraneous travel, thereby reducing our carbon footprint, this does give us the option to avoid unessential face-to-face meetings. At the same time, social networks are usually perceived as a celebration of social interaction, enabling online shopping, disseminating both traditional and innovative news, promoting music and sports events, as well as enticing us to travel and dine out. Sadly, especially over recent months, we are witnessing the extreme toxicity of some of these exchanges through terrorism and hate crimes down to the subversion of democracy gaining traction from social networks. In May 2019, Facebook went as far as to ban individuals deemed as “dangerous”<sup>7</sup>, causing vehement global discussion on whether Facebook should be classified as a platform or rather a publisher with the resulting implications, and questions about who should be responsible for identifying who is “dangerous”. The complete lack of guidelines was lamented by all, and Facebook, once again, failed to rise to the challenge of online toxicity so prevalent on its platform – even on its new, purified, “white” platform<sup>8</sup>.

Once the Mirrorworld has become ubiquitous (possibly even in the near future), we might also retrace its evolution as a master-platform that shared its provenance both with social networks as well as with online gaming. Since the debut of *Myst* in 1993, the first online game that shook up the world<sup>9</sup>, hundreds, if not thousands of games migrated from early handheld consoles or PCs to online gaming. Today, with 250 million registered players, *Fortnite*, developed by Epic Games, has become one of the most popular online video games ever created since it was first released in 2017. Participants move through the world with four players in a squad, cooperating towards a common objective on various missions<sup>10</sup>.

<sup>7</sup> Dave Lee, *Facebook bans “dangerous individuals”*, «BBC News Technology», 3 May 2019, <<https://www.bbc.com/news/technology-48142098>>.

<sup>8</sup> Annie Palmer, *Facebook ditched its classic blue layout for new all-white design with a “cleaner look and feel”, as it unveils Messenger desktop app and a tool to help you flirt with your “secret crush” in “privacy focused” overhaul*, «Daily Mail Online», 30 April 2019, <<https://www.dailymail.co.uk/sciencetech/article-6975479/Zuckerberg-explain-Facebook-gets-privacy-focused.html>>.

<sup>9</sup> <https://myst.com>.

<sup>10</sup> <https://en.wikipedia.org/wiki/Fortnite>.

Anyone noticing young and not so young people spontaneously dancing “the Floss” in the real world is witnessing the impact of protracted battles in the fictitious world, laced with its oh-so-comforting dances for respite – “the Floss” has its own Wikipedia page<sup>11</sup> – and has become emblematic of the Fortnite-effect emerging from the many, many hours spent in-world to discreetly seep into real-life behaviour.

Fortnite’s developers made some strategic decisions in devising the player’s interactions that were driven by the dark clouds of social media. «For example», Fortnite’s CEO Tim Sweeney explains, «there’s only chat between you and your squadmates. There’s no proximity chat, so it’s not possible for one person to grieve another person they don’t know, out in the open world»<sup>12</sup>. Lessons are being learned from the problematics of social networks and are influencing decisions in the development of new, more progressive experiences for players. At the same time, Fortnite’s designers were attentive to their committed players, and continued to develop new possibilities for the game. After trialing a new mode with players, they created a dedicated space where players could build their own islands – in much the same way that *Second Life*<sup>13</sup> did when they launched their persistent virtual world in 2003. Tim Sweeney further points out:

«Creative Mode, where players can go and create their own island and build their own work, including collaboratively with others, is now driving a large number of players. They’re building maps, they’re sharing them with friends, and we’re seeing more and more new types of gameplay emerge from that. It’s taken on a life of its own, and I think that’s going to be the driving force for Fortnite in the future»<sup>14</sup>.

Many more functionalities are in the pipeline as the developers continuously respond to player’s needs. The agency that drives social networks, in which everyone can simultaneously be both a consumer and producer, is similarly driving gaming. According to young players Eviatar and Elai (age 11), «once we enjoyed running along the *Temple Run*<sup>15</sup> path, jumping as we went, but that is no longer interesting. We need the space, and freedom to build our own world; invite our friends to develop our own island and know that we can return there to continue to develop it each time we log in»<sup>16</sup>. They are advocating for an open-ended, non-linear persistent world, where they are assured that the fruits of their labor remain stable, even after they log out.

<sup>11</sup> [https://en.wikipedia.org/wiki/Floss\\_\(dance\)](https://en.wikipedia.org/wiki/Floss_(dance)).

<sup>12</sup> Peter Rubin, *It’s a short hop from Fortnite to AI best friend*, interview with Tim Sweeney and Kim Libreri, «Wired», 21 March 2019, <<https://www.wired.com/story/epic-games-qa/>>.

<sup>13</sup> <https://secondlife.com/>.

<sup>14</sup> P. Rubin, *It’s a short hop*, cit.

<sup>15</sup> <<https://templerun3.co/>>. *Temple Run* is developed and published by Imangi Studios.

<sup>16</sup> Eviatar Hazan and Elai Rofe.



Figure 1. Eviatar and Elai's island in Fortnite

The notion of online collaboration required for a fully functional Mirrorworld demands a total global reach, in which participants are individually responsible for their own action and making their own contribution to the shared and persistent platform. The key to success here – and yet again a lesson to be learned from previous platforms – is the unequivocal rule of scalability. In the words of CTO Kim Libreri, founder of Fortnite:

«To truly see the future of collaborative gameplay and social experiences, you need a large-scale community to help that happen. This idea that a computer programmer in a basement invents the metaverse, it's just not true. We need a planet full of people to really help guide these things»<sup>17</sup>.

The vision that Libreri is suggesting here resonates with that foreseen by David Gelernter in his *Mirror Worlds* in 1992<sup>18</sup>. Over the years this vision has been called by many names: the “Metaverse”<sup>19</sup>, “Supranet”<sup>20</sup>, “Matrix”<sup>21</sup>, and now “Mirrorworld”. “Mirror Worlds” – a term first popularized by Yale computer scien-

<sup>17</sup> P. Rubin, *It's a short hop*, cit.

<sup>18</sup> David Gelernter, *Mirror Worlds or the day software puts the universe in a shoebox: how it will happen and what it will mean*, e-book, New York; Oxford: Oxford University Press, 1992.

<sup>19</sup> The term “Metaverse” was coined in Neal Stephenson's 1992 science fiction novel *Snow Crash*.

<sup>20</sup> The “Supranet” is a term coined at the turn of the 21st century by information technology analysis firm Gartner to describe the fusion of the physical and the digital (virtual) worlds, a concept that embeds the “Internet of things” as one of its elements: see <<https://en.wikipedia.org/wiki/Supranet>>.

<sup>21</sup> “The Matrix” is a virtual-reality dataspace in the 1984 science-fiction novel *Neuromancer* and in the 1999 science-fiction movie entitled, indeed, *The Matrix*.

tist Gelernter – brings additional and critical dimensions into the next phase of online experience – context, meaning, and, most importantly, action.

«Mirror Worlds? What are they? They are software models of some chunk of reality, some piece of the real world going on outside your window. Oceans of information pour endlessly into the model (through a vast maze of software pipes and hoses): so much information that the model can mimic the reality's every move, moment-by-moment»<sup>22</sup>.

Each and every one of us will be able to interact in the Mirrorworld; not simply by building a map of our own island, and protecting it from marauders, and not just “wandering” around the facades that we know from Google Earth. We will still do our shopping on Amazon and all its clones and will continue to swap images and experiences as we know from social networks; the major difference is that it will all be happening in 3D. This means that daunting software and hardware solutions still need to be resolved. But they are coming. 3D technologies are emerging all around us. A compelling solution could be floating screens we evoke in the space surrounding us – the kind of interaction we watched as Tom Cruise's iconic role as Chief John Anderton in the movie *Minority Report* (2002)<sup>23</sup> whose sci fi world was populated with gesture-activated screens.



Figure 2. Screenshot from the movie *Minority Report*, 2002

<sup>22</sup> D. Gelernter, *Mirror Worlds*, cit.

<sup>23</sup> <https://www.imdb.com/title/tt0181689/>.

There are, in fact, impressive innovations already evolving in reality to introduce 3D worlds into our lives. Perhaps we will evoke screens to float ghostly-like in front of us, gesture tracking and finger recognition to drive our programs. There are already solutions that pump data directly into our eyes through retina scanners<sup>24</sup> and, of course the ultimate solution: brain-driven computer interaction<sup>25</sup>. Some call this spatial computing, and, while most of the solutions are currently beyond our expertise and purse, we can buy Google Glasses, Magic Leap Goggles<sup>26</sup>, and many similar products online (at a price). Be assured that within a very short time the new generation of 3D-interaction will be appearing on a screen near you.



Figure 3. Takehiro Nakagawa, from Kusatu City (Japan) walking around the National Park in Kyoto with his family during the Saukura spring

### Taking the (3D) Leap of Faith

While social networks, online gaming, and Google Earth are reaching the boundaries of 2D experiences, the true leap of faith will occur once we move into a 3D environment. Every company that builds these platforms, comprising both the software and the hardware developers that will be driving them, is already in a

<sup>24</sup> [https://en.wikipedia.org/wiki/Retinal\\_scan](https://en.wikipedia.org/wiki/Retinal_scan).

<sup>25</sup> Brain Driven Interface, <[https://en.wikipedia.org/wiki/Brain%E2%80%93computer\\_interface](https://en.wikipedia.org/wiki/Brain%E2%80%93computer_interface)>.

<sup>26</sup> <https://www.magicleap.com/>.

fierce battle to be the one to get there first. While this seems very futuristic, it is actually very contemporary. We have already been there! Even if you had not downloaded Pokémon Go<sup>27</sup> yourself onto your smartphone, you probably know someone who has. Even museums were pulled into these frenetic adventures.

Pokémon Go's phenomenal success story started with Ingress<sup>28</sup>, the augmented-reality multiplayer game that was launched as a beta-version by Niantic in 2011. Ingress – currently with some 7 million players actively playing around the world – is an original IP (Intellectual Property) based in a near future science-fiction setting. The backstory is almost irrelevant, but it does drive the action. Players are informed how CERN discovered the Higgs Boson; a new substance called “exotic matter” which was abruptly released into the world. This exotic matter seeps into the world through portals and has the ability to affect humans. What is interesting for custodians of cultural heritage, however, is that these portals are in fact real-life cultural and historical locations such as parks and monuments, places where people traditionally congregate. What better a way to exploit this unassuming pastime than to turn these landmarks into nodes in a global game?

There are two factions of humans in the game – the enlightened and the resistance. According to the Ingress scenario, the enlightened believe “exotic matter” will bring about the next evolutionary step for mankind. The resistance sees the matter as coming from a possible alien force and are trying to control it so that it does not find its way into the wrong hands and destroy humanity. But who cares about the backstory? The point is to get out and catch them! This app is not that kind you play while slumped on your chair at home. It requires actual movement and encourages players to get outside in order to unlock new features, weapons and challenges: for example, to capture and control “portals” which spew matter from cultural places, such as museums and libraries. You have to actually go somewhere to play this game. Once you connect three portals, you form a region and claim a territory. Winner of the Grand Prize at the 2014 Japanese Media Arts Festival, Ingress was acclaimed for «transforming the real world into the landscape for a global game of mystery, intrigue, and competition»<sup>29</sup>.

Ingress's users were responsible for helping create the data pool that determined where “Pokéstops” and gyms appear in Pokémon Go. The game was first released

<sup>27</sup> <https://www.pokemongo.com/en-us>.

<sup>28</sup> <https://play.google.com/store/apps/details?id=com.nianticproject.ingress>.

<sup>29</sup> See <<https://gadgets.ndtv.com/photos/best-android-apps-of-2013-google-play-users-choice-awards-16648#7>>. About the prize-winning game Ingress see also Rene Ritchie, *Ingress, Google's augmented reality game, hits iPhone and iPad*, «iMore», 13 July 2014, <<https://www.imore.com/ingress-googles-augmented-reality-game-hits-iphone-and-ipad>>.

exclusively for Android devices on 15 November 2012 and was made available for Apple's iOS on July 14, 2014. To play, you open up the game and start walking to prominent local landmarks – represented in the game as “Pokestops” – where you can gather supplies such as “Pokeballs”. Pokeballs are what you fling at online “pocket monsters”, or Pokemon, to capture them for training. At other locations called “gyms” – which may or may not be actual gyms while locations such as museums and libraries made visible in the game are very much in the real world – all places where Pokemon battle one another for supremacy. In the early days of Ingress, Niantic formed a beginning pool of portal locations for the game based on historical markers, as well as a data set of public artwork mined from geo-tagged photos on Google. «We basically defined the kinds of places that we wanted to be part of the game», CEO of Niantic, John Hanke said. «Things that were public artwork, that were historical sites, that were buildings with some unique architectural history or characteristic, or a unique local business»<sup>30</sup>.



Museums soon noticed that their physical locations were getting a lot more traffic in unusual places, such as in their galleries and gardens, especially from young people who were strangely engrossed in their smartphones. Many rose to the challenge and advertised their Pokemon potential through social networks, others were less enamored with the unwelcome intrusion into their spaces.

Figure 4. Screenshot: *Pokemon comes to the Israel Museum, Jerusalem, July 2016*

<sup>30</sup> Michelle Lulic, *Are Pokemon Based On Location, in "Pokemon Go"? Traveling Is The Best Way To Catch 'Em All*, «Bustle», 12 July 2016, <<https://www.bustle.com/articles/172192-are-pokemon-based-on-location-in-pokemon-go-traveling-is-the-best-way-to-catch-em>>.





Figure 5. Screenshot, US Holocaust Museum asks Pokemon Go players to stop

In July 2016 the US Holocaust Museum asked Pokemon Go players to stop. A spokesman for the museum in Washington said that «playing the game inside a memorial to victims of Nazism was “extremely inappropriate”»<sup>31</sup> The Arlington National Cemetery, just three miles away from the museum, had also warned off Pokemon players.

Cultural custodians must wake up to the fact that they are already identified on the map as “historical markers”. The footfall that has marked their territory for more than a decade is up for grabs, easily mined through ubiquitous geo-tags and their tradition of public welcoming easily exploited for commercial gain. The key here is not to lament the loss of agency but to step into the game and make sure the game is played by their rules.

Of course, augmented reality is not new to museums. There are several examples of excellent projects that add additional layers of interpretation for visitors.

Gallery One at the Cleveland Museum of Art launched in 2013 their ArtLens<sup>32</sup> with several functions that interpret the collection for the visitor. “Featured Artworks” combine interpretive media (movies, comparative images, text and au-

<sup>31</sup> *US Holocaust Museum asks players to stop*, «BBC News US & Canada», 13 July 2016, <<http://www.bbc.com/news/world-us-canada-36780610>>.

<sup>32</sup> Jane Alexander – Jake Barton – Caroline Goeser, *Transforming the Art Museum Experience: Gallery One*, <<https://mw2013.museumsandtheweb.com/paper/transforming-the-art-museum-experience-gallery-one-2/>>.

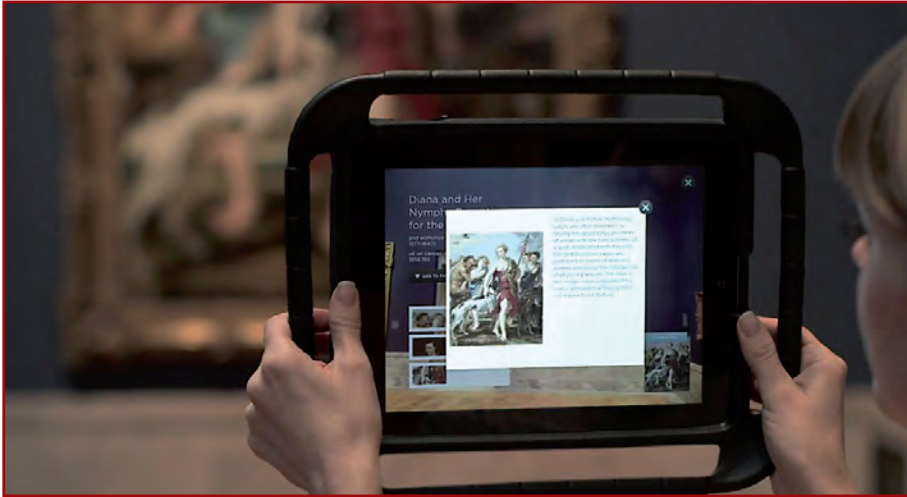


Figure 6. *ArtLens, Gallery One, Cleveland Museum of Art*

dio) with scanning image recognition, features that the curatorial team called “guided looking”. The scanning feature incorporates the device’s camera and image-recognition to provide an augmented reality experience for users on site. As visitors scan the artwork, the app recognises the object and provides context-sensitive content about the work. While the Cleveland Museum was one of the early institutions to introduce augmented reality (AR) in their galleries, many museums have followed in kind. AR apps are particularly popular in science and natural history museums, with dinosaurs and polar bears popping up all over the place to the delight of the visitors.

In 2016 the Seattle Art Museum took visitors out to their Olympic Sculpture Park (OSP) to visit the Gardens of the Anthropocene, an AR installation by Tamiko Thiel. Perhaps to the horror of the visitors, the farewell-to-spring plants had mutated to develop succulent drought-resistant leaves, as the species named there, *Clarkia irritabilis* or *Clarkia antenna*, had reacted to climate change in more than terrifying ways. According to the OSP, you could see two or three stages of development, in which the stamens and pistils had begun to fuse into an antenna-like form. In the final mutation, the flower petals had developed “marginal teeth” on their rims that detected the presence of mobile devices, and the flowers became engorged as they apparently fed off the electromagnetic emissions. The artist assured the public, however, «the behavior of the flowers is unnerving, but does not produce any known ill effects in humans»<sup>33</sup>.

<sup>33</sup> Tamiko Thiel, *Gardens of the Anthropocene*, <<http://tamikothiel.com/gota/m.html>>.



Figure 7. *Gardens of the Anthropocene, Augmented Reality Installation by Tamiko Thiel, 2016*



Figure 8. *Gardens of the Anthropocene, Augmented Reality Installation by Tamiko Thiel, 2016*

Another AR project launched in 2014 took visitors outdoors again. This time the host was the Museum of London and the AR experience took place across the Streets of London<sup>34</sup>. Streetmuseum 2.0, developed with creative agency Brothers and Sisters, invited visitors to visit sites across London where they could evoke

<sup>34</sup> Ellie Zolfagarifard, *Streets of London now... and then: Stand still and picture yourself in history with app that creates hybrid images of present and past*, «Daily Mail Online», 26 February 2014, <<https://www.dailymail.co.uk/sciencetech/article-2567739/Streetmuseum-app-creates-hybrid-images-London.html>>.

hidden histories of the city on the especially designed app through the geo-tagged points on the Google-map of the city. The alluring overlays of the contemporary live scene were superimposed side by side together with the historic images drawn from the Museum of London's extensive art and photographic collections. The narrative was further enhanced by the historical information from the curatorial interpretation. Sadly, Streetmuseum is no longer available at the moment but the Museum of London has several other excellent apps for visitors<sup>35</sup>.



Figure 9. Museum of London, StreetMuseum 2.0, 2014



Figure 10. Christina Broom, People and traffic in Oxford Street around the turn of the 20th century

<sup>35</sup> <https://www.museumoflondon.org.uk/discover/museum-london-apps>.

## A persistent place

Our initial experience with global, digital technology was with the World Wide Web. This sprawling communication platform consists of interconnected networks, using standardised communication protocols and serves up information driven by the power of algorithms. This text-driven platform eventually became dominated by Google. The second significant platform comprised social media, distributed predominantly over mobile phones. Social networks digitised people and encapsulated human behavior; even our personal relationships became data-points to be mined and exploited. These potentially democratising networks were primed to “connect people” but were eventually dominated by Facebook, Vkontakte and WeChat, which make vast profits and exercise almost unimaginable power over information, politics and even democracy itself. According to John Kelly the third platform is just around the corner:

«We are now at the dawn of the third platform, which will digitise the rest of the world. On this platform, all things and places will be machine-readable, subject to the power of algorithms. Whoever dominates this grand third platform will become among the wealthiest and most powerful people and companies in history, just as those who now dominate the first two platforms have. Also, like its predecessors, this new platform will unleash the prosperity of thousands more companies in its ecosystem, and a million new ideas – and problems – that weren’t possible before machines could read the world»<sup>36</sup>.

A machine-readable world promises a formidable reality, which at the same time is coupled with a momentous loss of human control. In essence, it will be the digital actors, the robots who will be seeing this world for us, in much the same way that self-driving cars see the world today in a combination of reality and computerised perception. When a robot looks, it will be both seeing for itself and providing a scan for other robots. The measure of success here is the way the world is modeled, the mapping of the road ahead of the vehicle, together with the on-going interpretation of real-time data that will be flooding in at speeds that humans could never dream of. This world will depend on simultaneous localisation and mapping (SLAM)<sup>37</sup>, the computational challenge of constructing and constantly updating a map of the potential environment while simultaneously keeping track of an agent’s location, and movement within the potential space.

This will create a radical paradigm shift in the ways we experience the world around us. Some are calling it “XR”, a catch-all term that describes a range of visual extensions of reality, or cross reality. The evolving aesthetic is driven by com-

<sup>36</sup> K. Kelly, *AR Will Spark the Next Big Tech Platform*, cit.

<sup>37</sup> [https://en.wikipedia.org/wiki/Simultaneous\\_localization\\_and\\_mapping](https://en.wikipedia.org/wiki/Simultaneous_localization_and_mapping).

puter vision, machine learning, and the increasingly collaborative intersection between human and machine. This growing capacity enables us to act with digital devices in a world that blurs the digital and biological. We can call up inexhaustible data to enrich and inform the action in real time in a formidably collaborative environment. The move towards AR vision is sneaking up all around us. Google is already taking its first steps in this direction with its Google Maps augmented reality<sup>38</sup> while Minecraft Earth and the Harry Potter Wizards Unite<sup>39</sup> AR apps launched in the summer of 2019 which they hope will all rival the Pokemon Go craze<sup>40</sup>. Significant steps are already underway whether we notice or not. Facebook has recently announced several of its robotics projects, which explore ways for robots to teach themselves from experience rather than data prepared for training purposes – a skill Facebook believes could have implications beyond the field of robotics. «The real world is messy, it's difficult», Roberto Calandra, a research scientist in Facebook's AI division explains. «The world is not a perfect place; it's not neat. So the fact that we are trying to develop algorithms that work on real robots [will] help to create [AI] algorithms that, generally speaking, are going to be more reliable, more robust, and that are going to learn faster»<sup>41</sup>.

At the same time, Google is frenetically engaged with Big Data, Machine Learning, Deep Learning and voluminous AR projects. Under their bespoke banner “Advancing the state of the art – We work on computer science problems that define the technology of today and tomorrow”, Google is already heading towards becoming one of the leading harbingers of the Mirrorworld. While we are concerned with the culture values of the Mirrorworld, at this moment in time we need to pause think why Google is laying all the groundwork for the Mirrorworld and are reminded of Zuboff's insightful, and dire warning of appropriating everything for the lucrative profits inherent in surveillance capitalism.

«[...] the everywhere, always-on instrumentation, datafication, connection, communication, and computation of all things, animate and inanimate, and all processes—natural, human, physiological, chemical, machine, administrative, vehicular, financial. Real-world activity is continuously rendered from phones, cars, streets, homes, shops, bodies, trees, buildings, airports, and cities back to the digital realm, where it finds new life as data ready for transformation into predictions, all of it filling the ever-expanding pages of the shadow text»<sup>42</sup>.

<sup>38</sup> Karissa Bell, *Google Maps augmented reality directions are here (and actually useful)*, «Mashable», 7 May 2019, <<https://mashable.com/article/google-maps-augmented-reality-walking-directions/>>.

<sup>39</sup> <https://www.harrypotterwizardsunite.com/>

<sup>40</sup> Peter Rubin, *Minecraft Earth wants to be the Next Pokemon Go – but Bigger*, «Wired», 17 May 2019, <<https://www.wired.com/story/minecraft-earth-wants-to-be-the-next-pokemon-go-but-bigger/>>.

<sup>41</sup> Kristin Houser, *Facebook finally unveils secret robotics projects*, 20 May 2019, <<https://futurism.com/the-byte/facebook-robot-projects>>.

<sup>42</sup> Shoshanna Zuboff, *The Age of Surveillance Capitalism*, cit, p. 202.

Predicting our actions is the *modus operandi* of marketing, targeting us to sell something, or even influence our political preposition for a particular political agenda. These progressions also clearly represent daunting challenges, to us as individuals, but especially to custodians of cultural heritage, those institutions who maintain historical and contemporary culture that is both tangible and intangible and is poised to flood the Mirrorworld with unique and archetypal content.

Clearly for the Mirrorworld to ripen to its full potential there are several significant, technological challenges to be surmounted, presumably by the leading companies such as Google. Ideally we will need digital prosthetics that offer six degrees of freedom (6DoF)<sup>43</sup>, the number of axes along which the body is able to freely move in three-dimensional space including the X, Y and Z axes as well as change orientation between those axes though rotation usually called pitch, yaw and roll. Somebody has to model the world in 3D, and, in addition, every object in this world will need a digital twin. Will consumers be doing this themselves with tiny head-mounted, phone-equipped, or street embedded cameras, or will the data be flooding in from the long-established, world-wide commercial pixel-pushers, Google, Apple, Microsoft, Ingress, Facebook and other global giants? What role will the custodians of culture be playing in this evolving matrix, and what will their responsibilities be to the public?

«History will be a verb. With a swipe of your hand, you will be able to go back in time, at any location, and see what came before. You will be able to lay a reconstructed 19th-century view right over the present reality. To visit an earlier time at a location, you simply revert to a previous version kept in the log. The entire mirror-world will be like a Word or Photoshop file that you can keep “undoing”. Or you’ll scroll in the other direction: forward. Artists might create future versions of a place, in place. The verisimilitude of such crafty world-building will be revolutionary. These scroll-forward scenarios will have the heft of reality because they will be derived from a full-scale present world. In this way, the Mirrorworld may be best referred to as a 4D world»<sup>44</sup>.

This is where the custodians of cultural heritage enter stage right. From museum basements, conservation laboratories and university research huddles, curators, developers and conservators have been compiling data cloud upon data cloud to preserve cultural objects, artworks, historical buildings and monuments for future generations. After the recent and disastrous fire at Notre-Dame Cathedral in Paris, for example, the international press reported how American historian Andrew Tallon, art professor at Vassar, and architectural historian and Columbia’s Paul Blaer spent five days in 2010 creating an incredibly accurate 3D laser scan of

<sup>43</sup> <https://www.techopedia.com/definition/12702/six-degrees-of-freedom-6dof>.

<sup>44</sup> K. Kelly, *AR Will Spark the Next Big Tech Platform*, cit.

Notre-Dame<sup>45</sup>. If researchers like Tallon and Blaer, and hundreds like them around the world, come together to share their data, these could serve as critical fragments of the future Mirrorworld. It is a question of collaboration and scalability. It sounds complicated – it needs not be so. All you need to do is to gather and combine the data points. Motivated by the terrifying loss of one of Europe's most beloved landmarks, within a few days more than 50 crowdfunding campaigns were launched to restore Notre-Dame on the website GoFundMe<sup>46</sup>. Even more sobering were the Notre-Dame crowdsourced photo/video photogrammetry repositories that were openly shared over several platforms, including shared Sketchfab<sup>47</sup>. The community is already there, and once motivated has much to contribute.

Typically, like an iceberg that is mostly hidden under water, most museums around the world can only showcase a small percentage of their collections at any given time, while other objects and artworks wait for their debut to make it up to the gallery for their moment of fame in the spotlight. However, once these hidden objects have been digitised, they can be called upon to tell their stories, and activated to compile the 4<sup>th</sup> dimension of the Mirrorworld: the scrolling back and forth in time. For custodians of cultural heritage, the added advantage in the Mirrorworld concerns their decades-long tradition of creating narratives. While most objects on display are exhibited to tell a specific story, in fact, from their own histories, they may have more than one story to tell. Typically, objects are accompanied by captions, labels or short wall texts, not always detailed enough to tell the full story. In a digital world, multi-narratives may be called up to recall the complex and sometimes contentious provenance of objects.

In 1995 a fiery controversy exploded when Enola Gay, the B-29 aircraft that was used to drop the bomb on Hiroshima, was put on display at the National Air and Space Museum (NASM)<sup>48</sup>. Exhibiting the plane fermented bitter dispute: the conflicting historical perspectives that were caught up in its tailwind ended up in upsetting so many communities – war veterans, Japanese survivors – and all parties were enraged by the staging of the narratives, especially regarding the decision to drop the bomb, that Martin Harwit, NASM Director, eventually quit the museum.

<sup>45</sup> Rachel Hartigan Shea, *Historian uses lasers to unlock mysteries of Gothic cathedrals*, «National Geographic», 16 April 2019: see Notre-Dame cloud point scan: <<https://news.nationalgeographic.com/2015/06/150622-andrew-tallon-notre-dame-cathedral-laser-scan-art-history-medieval-gothic/>>.

<sup>46</sup> *Over 50 crowdfunding campaigns launched to restore Notre-Dame*, «New Straits Times», 16 April 2019, <<https://www.nst.com.my/world/2019/04/479874/over-50-crowdfunding-campaigns-launched-restore-notre-dame>>.

<sup>47</sup> <https://sketchfab.com/3d-models/notre-dame-crowdsourced-photogrammetry-c6f03e2284e049f38f118378b012c979>.

<sup>48</sup> <https://www.atomicheritage.org/history/controversy-over-enola-gay-exhibition>.



The exhibition ran from January 1995 to May 1998. Despite all the controversies, it was one of the most popular special exhibitions in the history of the Air and Space Museum, attracting more than a million visits in its first year alone, and a total of nearly four million visitors by the time it closed. In the Mirrorworld there will be breadth and length enough for all narratives, and we will be able to scroll through competing stories in the batting of an eyelid.

The global players who have excelled in these kinds of big data projects are those countries that have amassed their cultural holdings at a national level and have created their own portals to promote their country's cultural assets to both local and international users. The portals that span over neighbouring countries, or even whole continents, are even more impressive. In the US the leaders have traditionally been the Smithsonian, and more recently the Digital Public Library of America (DPLA)<sup>49</sup> – modelled on the pan-European platform Europeana<sup>50</sup>. The latter currently has over 58 million objects, drawn from European galleries, libraries, archives and museums (GLAMs). The Europeana Foundation is already active in building the proto-Mirrorworld as one of the founding members of the initiative called Time Machine<sup>51</sup>, a consortium of GLAMs and computer science organisations who have made a case for a billion-euro investment package that aims to use AI to their advantage to mass digitise, simulate and interpret big data of the past. According to its website, Time Machine is promising to be one of the most advanced Artificial Intelligence systems ever built:

«The Time Machine will create advanced AI technologies to make sense of vast amounts of information from complex historical data sets. This will enable the transformation of fragmented data – with content ranging from medieval manuscripts and historical objects to smartphone and satellite images – into useable knowledge for industry. In essence, a large-scale computing and digitisation infrastructure will map Europe's entire social, cultural and geographical evolution»<sup>52</sup>.

The Time Machine plans to mine Europe's rich past through harnessing digital innovation and creating a living resource that allows you, according to their webpage, «to travel not only through space, but through time»<sup>53</sup>. The standards, frameworks and collections of the Europeana service will become important building blocks for the Time Machine for future generations. The pilot started in March 2019 and aims to deliver a project proposal for the Time Machine project in early 2020.

<sup>49</sup> <https://dp.la>.

<sup>50</sup> <https://www.europeana.eu/portal/en>.

<sup>51</sup> <https://timemachine.eu/>.

<sup>52</sup> <https://www.timemachine.eu/unleashing-big-data-of-the-past-europe-builds-a-time-machine/>.

<sup>53</sup> <https://www.timemachine.eu/>.

## Museum literacy to inform 4D literacy

We can look forward to a time when we will be able to search through physical space in much the same way we search by text on a browser. Perhaps we will be able to call up through gesture or verbal commands the future Siris and Alexas to help us find our way around the Mirrorworld – to scroll across histories and encounter a chorus of multi-narratives. This will demand a new kind of visual or spatial literacy that will enable us to simultaneously consume and produce objects and processes in a 4D environment. Michael Foucault drew attention to the mausoleum-like nature of modern museums that are connected to temporal discontinuities as «heterochronias – in which time never ceases to pile up and perch on its own summit»<sup>54</sup>. Where heterotopias may be seen as non-places, or physical spaces that have been geographically misplaced, Foucault recognizes how museums (and libraries) act as *heterochronias* – where the characteristic of Western culture in the 19<sup>th</sup> century dissolves both geography and time. He suggests that

«the idea of accumulating every-thing, the idea of constituting a sort of general archive, the desire to contain all times, all ages, all forms, all tastes in one place, the idea of constituting a place of all times that is itself outside time and protected from its erosion, the project of thus organizing a kind of perpetual and indefinite accumulation of time in a place that will not move – well, in fact, all of this belongs to our modernity»<sup>55</sup>.

## Conclusion

GLAMs are the ideal sand boxes that will help us master the new literacies of 4D *heterochronias*. Museums are ideally equipped to stage the kind of learning that we will need to safely navigate the Mirrorworld. It is actually not really that far in the future. As individuals acting in today's world, we are already generating gigabytes of data from the moment we wake up and our smartwatch tells us how much, and how well we slept last night, how many times we were in REM dream state and how restless we were between our dream cycles. If we check our pedometer we will be able to recall if we left our beds during the night, perhaps for a drink, or perhaps even sleepwalking. Only my pedometer can tell me. As soon as our fingers touch our keyboards we could be signalling our mood<sup>56</sup> and through a machine-readable selfie we might be posting, or even broadcasting, our emotional state. Perhaps in the near future we will be shouting out our mental state through a deviation from the rhythm of our normal keyboard strokes – that is if anyone cares to listen in.

<sup>54</sup> Michel Foucault, *The Essential Works of Foucault, 1954-1984*, v. 2: *Aesthetics*, edited by J. D. Foubion, translated by R. Hurley and others, London: Penguin, 1997, p. 182.

<sup>55</sup> *Ibidem*.

<sup>56</sup> Will Knight, *Your smartphone's AI algorithms could tell if you are depressed*, «MIT Technology Review», 3 December 2018, <<https://www.technologyreview.com/s/612499/your-smartphones-ai-algorithms-could-tell-if-you-are-depressed/>>.

CTV cameras already populate our cities. In China, facial recognition in public places allows individuals to simply smile in order to pay for a meal, as a pilot which is being tested in KFC in the Chinese city of Hangzhou does, or, on a different scale, it can help the police to catch criminals walking in the street<sup>57</sup>. Through tracking individuals over their mobile phones, China already has the capability to keep tabs on its population. Through the quickly evolving “social credit system”, surveillance technology is not only able to monitor, but even to regulate the everyday life of its population by automatically generating ratings for each citizen and affecting everything accordingly, from loan approvals to the permission to board flights<sup>58</sup>. The future is not coming soon – it is already here.

This call to action cannot halt the growing progression of the authorities to take over public space, even though some cities are taking a bold stance to staunch this relentless progression when they feel threatened by the potential of the digital gone wild<sup>59</sup>. But once these technologies are instituted, GLAMs can play an active role in perpetuating narratives of cultural heritage, through digital enterprise and the spatial and *heterochronic* nature of their institutions. Leaders such as Europeana have been promoting these kinds of initiatives since 2015:

«We believe that the digital cultural heritage available via Europeana can make Europe an even better place to live, learn, work or visit. That’s why, under the recent Presidencies of the Council of the European Union, we worked with experts, member state and EU policy-makers to forge a path and create Recommendations for the use of our digital cultural heritage in four sectors - research, education and learning, tourism and the creative industries»<sup>60</sup>.

Essentially it is up to custodians of cultural heritage to become proactive in building the Mirrorworld rather than wait for commercial actors to advance their own fabricated stories – and with the GLAM community blazing the trail in the Mirrorworld these unique and precious stories will be accessed, shared, and augmented in confidence.

<sup>57</sup> Bernard Marr, *The Fascinating Ways Facial Recognition AIs Are Used In China*, «Forbes», 17 December 2018, <<https://www.forbes.com/sites/bernardmarr/2018/12/17/the-amazing-ways-facial-recognition-ais-are-used-in-china/#767249c35fa5>>.

<sup>58</sup> Zhou Jiaquan, *Drones, facial recognition and a social credit system: 10 ways China watches its citizens*, «South China Morning Post», 4 August 2018, <<https://www.scmp.com/news/china/society/article/2157883/drones-facial-recognition-and-social-credit-system-10-ways-china>>.

<sup>59</sup> Kate Conger – Richard Fausset – Serge F. Kovalski, *San Francisco Bans Facial Recognition Technology*, «The New York Times», 14 May 2019, <<https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html>>.

<sup>60</sup> <https://pro.europeana.eu/post/getting-heritage-out-of-its-boxes-and-into-the-real-world>.

*Questo articolo vuole spronare i custodi del patrimonio culturale a farsi trovare preparati dalla nuova piattaforma in arrivo: le esperienze 3D e 4D del nascente Mirrorworld.*

*Costruito sulla base delle prime generazioni dell'esperienza online e sulla tentacolare piattaforma di comunicazione del World Wide Web e veicolato dalla portata universale di social network e gaming online, il Mirrorworld supererà le limitazioni intrinseche delle esperienze 2D per condurci in un ambiente 3D completamente funzionante.*

*Le istituzioni pubbliche che gestiscono il patrimonio culturale rivolto all'utenza (gallerie, biblioteche, archivi e musei - GLAM) sono nella posizione ideale per poter avviare la rete di comunicazioni spaziali che guiderà la nuova piattaforma.*

*Nonostante vi siano sfide notevoli da superare, è compito dei custodi del patrimonio culturale partecipare attivamente alla costruzione del Mirrorworld, invece di aspettare che soggetti terzi avanzino interessi commerciali producendo cultura imitativa, storie simulate o storie inventate.*

*Come guardiani degli oggetti culturali i GLAM hanno la capacità, l'esperienza e soprattutto la responsabilità di raccontare le loro storie in maniera potente, precisa e con una tecnologia 4D pienamente funzionante.*

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

# I MOOCs, opportunità per la formazione di base e l'apprendimento continuo: una storia (anche) italiana

**Matilde Fontanin** - Università di Trieste e Sapienza Università di Roma

**Eleonora Pantò** - Esperta di tecnologie digitali ed educazione

*I MOOCs – Massive Open Online Courses – sono corsi organizzati online e nati sotto il segno delle risorse educative aperte (OER), caratterizzati da grandi numeri di partecipanti diversi per nazionalità, formazione, cultura. Nati nel contesto accademico angloamericano sulla scia dello sviluppo dell'e-learning e della rivoluzione digitale, dai primordi al debutto vero e proprio, dalle prime reazioni entusiastiche alle successive difficoltà, sono maturati sviluppando modelli e approcci differenti all'apprendimento partecipativo, utilizzando svariate piattaforme, offerti da diverse istituzioni, dalle grandi università agli istituti di ricerca e formazione. Tutt'ora erogati in grande numero, costituiscono una preziosa opportunità per la formazione personale e professionale in un mondo che acquisisce sempre maggiore consapevolezza della necessità di un apprendimento che dura per tutto l'arco della vita. Dopo un rapido sguardo alla loro gestazione, si analizzeranno i modelli didattici, i ruoli dei portatori di interesse nel settore della formazione formale e non formale, le piattaforme dedicate in ambito internazionale, europeo ed italiano, con particolare attenzione all'ambito umanistico e alle discipline dell'organizzazione dell'informazione, della documentazione, della ricerca. Si riporteranno a titolo esemplificativo il caso di una piattaforma europea (EMMA) e quello di un corso di formazione professionale dei bibliotecari (AIBformazione).*

## I MOOCs: cosa sono e da dove vengono

Con l'acronimo MOOC – *Massive Open Online Course* – si designano corsi online rivolti a gruppi molto numerosi e diversificati (Deng – Benckendorff – Gannaway 2019, p. 48)<sup>1</sup>. Il fenomeno è recente e ovviamente legato alla rivoluzione digitale, ma non ne va sottovalutata la dimensione cognitiva.

Il termine fu coniato nel 2008 (Kenny – Gunter 2014) per il corso di George Siemens e Peter Downes<sup>2</sup> all'Università di Manitoba, in Canada, sul modello di

<sup>1</sup> La definizione completa parla di «open, large-scale web-based courses designed and delivered by accredited higher education institutions and organisations in which anyone with a smart device and internet connection can participate, regardless of age, gender, geographic location, or education background.»

<sup>2</sup> Il titolo del corso si può tradurre in “Connettivismo e conoscenza connettiva”, quindi un tema vicino ad un approccio partecipativo come quello del MOOC.

MMORPG (*Massively Multiplayer Online Role-Playing Game*), ambienti ad alta interazione a metà tra il gioco di ruolo e il videogioco in massa. Si trattava di un corso frequentato da 25 studenti in presenza che fu aperto alla partecipazione online, attirando oltre 2200 iscritti. Per l'accento posto sull'interazione e sulla partecipazione collettiva alla costruzione della conoscenza, il modello venne definito cMOOC, ovvero MOOC connettivista.

Un precursore di rilievo dei MOOCs era stato il corso su vita e pensiero di Sant'Agostino che James J. O'Donnell dell'Università della Pennsylvania aveva tenuto nel 1994 ad una dozzina di studenti in presenza e a 500 iscritti a distanza. Per definizione del docente stesso si trattava di un MOOC ante litteram (O'Donnell 2012), tuttavia è ben riconoscibile il modello didattico: i partecipanti, di aree e confessioni diverse, poterono discutere in maniera profonda e libera di pensiero religioso. Può sembrare curioso che un simile fenomeno muova i suoi primi passi in un ambito disciplinare che decisamente è quello delle *Humanities*, ma forse non è un caso: come vedremo, l'insegnamento e l'apprendimento attraverso qualsiasi mezzo – incluso il MOOC – necessitano di un approccio all'individuo; la pura tecnologia non è sufficiente a coinvolgere i partecipanti per motivarli a raggiungere gli obiettivi formativi.

Il termine "MOOC" si diffonderà solo dopo il 2011, con i corsi organizzati presso l'Università di Stanford (Siemens 2014, p. 32), ma a quel punto svilupperà un diverso modello. Dopo una sperimentazione di Andrew Ng nel 2007 – il progetto SEE (*Stanford Engineering Everywhere*) – nel 2011 vennero inaugurati tre corsi: il corso sul *Machine Learning* dello stesso Ng e quello sui *Databases* di Jennifer Widom, su una piattaforma che sarebbe diventata Coursera, e il MOOC sull'Intelligenza artificiale di Sebastian Thrun e Peter Norvig, sulla piattaforma che si sarebbe poi evoluta in *Udacity* (Ng – Widom 2014, p. 37). I numeri furono altissimi: il corso di Thrun ebbe 160.000 iscritti. Vennero tutti definiti xMOOCs, secondo un modello in cui l'esperienza didattica è fortemente diretta dal docente. Si proponevano come strumenti per abbattere le barriere ad una formazione strutturata di alta qualità disponibile per tutti.

A seguito dell'esperienza di Stanford un articolo del New York Times dichiarò il 2012 "L'anno dei MOOCs" (Pappano 2012). Si dice che coniugano istruzione, intrattenimento (inteso come *gaming*) e social networking, che, a differenza delle "vecchie" risorse educative aperte che danno la sensazione di spiare dal buco della serratura, offrono un'esperienza partecipativa "with you in mind". Non sfuggirà che quel *you* richiama la famosa copertina dello stesso New York Times<sup>3</sup> che a fine 2006 annunciava il Web 2.0, sembra si voglia fare riferimento ad una rivoluzione davvero epocale.

<sup>3</sup> «TIME Magazine», 25 dicembre 2006, <<http://content.time.com/time/covers/0,16641,20061225,00.html>>.

Già nel 2013 cominciarono ad apparire articoli di segno opposto<sup>4</sup>, dato che la percentuale di studenti che completava i percorsi si aggirava intorno al 7% (Gaggi 2013). Lo stesso Thrun ne attestava il fallimento, ma allo stesso tempo fondava *Udacity*, una piattaforma proprietaria che offre tutt'ora MOOC dedicati alle competenze dei settori scientifici - le cosiddette STEM - in stretta collaborazione con le aziende.

Un report di Stanford del 2014 (Hollands - Tirthali 2014) analizza la percezione dei MOOCs nelle università statunitensi, e rileva che questa modalità formativa suscita timore tra gli intervistati rispetto alla futura destinazione dei fondi federali, al ruolo futuro degli atenei e all'indipendenza della ricerca accademica, poiché determina uno spostamento verso il concetto che l'alta formazione si può fornire a costi contenuti. D'altro canto, se ne coglie il valore quali strumenti efficaci per aiutare a colmare il divario tra l'alta formazione - che richiede molto tempo e riflessione per la creazione di contenuto - e l'esigenza di risultati e competenze "qui e ora" del mondo dell'impresa. La CRUI nelle sue linee guida (Progetto MOOCs Italia - CRUI 2017) ne evidenzia l'utilità per la Terza missione e per l'orientamento.

Uno sguardo alla letteratura accademica restituisce numerosi articoli in diversi ambiti disciplinari - dalla tecnologia alla didattica, dall'impiego nell'educazione formale a quello nella formazione professionale - e punti di vista - quello del fruitore, del docente, delle istituzioni che li offrono e via dicendo. Una prima analisi degli articoli di taglio compilativo e riepilogativo conferma l'iniziale entusiasmo per i MOOCs come portatori di una nuova modalità di istruzione aperta, sulle ali degli obiettivi culturali dello sviluppo sostenibile, e ne rileva le battute di arresto. Preso atto che non potevano sostituire l'educazione formale, e che non erano sufficienti a far crescere il livello di istruzione mondiale, si è cercato di ridimensionarne gli obiettivi per coglierne le opportunità. I MOOC possono costituire una preziosa risorsa per l'arricchimento personale e professionale, per la collaborazione internazionale in rete e per la promozione delle istituzioni coinvolte. Un recente rapporto (Shah 2018) stima le università che offrono i MOOC a più di 800, il numero totale di corsi a 9.400, e il numero di studenti in 81 milioni: numeri tutt'altro che piccoli, nonostante le delusioni raccontate nella nostra premessa.

Vedremo ora su quali modelli si basano, come funzionano e su quali piattaforme si possono trovare, realtà che vanno da Coursera e EdX, da FutureLearn ad EMMA, e contano milioni di utenti e centinaia di corsi.

## I modelli di apprendimento

Dal momento che arrivano dalla tradizione dell'e-learning, i MOOC si basano sugli aspetti collaborativi, sull'acquisizione di competenze e sulla costruzione della co-

<sup>4</sup> Ad esempio si veda Wetterstrom 2014.

noscenza a partire dall'esperienza e dalla conoscenza pregressa. Ricordando le lezioni di Merrill (Merrill 2002), l'apprendimento è più efficace se si coinvolgono i discenti nella risoluzione di problemi, se si attiva la conoscenza pregressa e viene data la possibilità di applicarla e di integrarla nel proprio mondo. Un ambiente collaborativo online è certo favorevole a questo modello di progettazione.

Riprendiamo il confronto tra i due modelli di apprendimento cMOOC e xMOOC.

Il MOOC connettivista sviluppato da George Siemens pone l'accento sugli aspetti collaborativi e sulla condivisione sociale: attraverso il confronto tra pari i discenti creano contenuto degno di essere condiviso, e restano padroni del loro spazio di interazione. Il docente non è altro che uno dei nodi della conoscenza, gli stessi partecipanti diventano docenti: «Our interpretation of openness is that, if you learn transparently, you become a teacher, because people observe your learning; they see what you are doing and they can emulate or adapt it» (Siemens 2014, p. 31). Il contenuto non è tanto un prerequisito del corso quanto un suo prodotto. Siemens racconta come questo abbia favorito la nascita di veri e propri gruppi di ricerca dopo la conclusione del corso, inoltre sottolinea l'approccio aperto dei cMOOCs alle risorse didattiche aperte (OER)<sup>5</sup>, che non si limita a far riferimento a documenti ad accesso aperto, ma facilita la produzione di contenuti da parte dei partecipanti, e favorisce la condivisione e il riuso delle risorse prodotte anche dagli stessi organizzatori. Vale a dire, ad esempio, che le videolezioni o i test prodotti nell'ambito di un cMOOC sono riutilizzabili da altri, anche per essere combinate in nuovi contenuti. Inoltre, quelle risorse devono restare disponibili anche dopo che il corso è finito: «Our course was so different from the current MOOCs because we strove to drive the students away from our central platform whereas edX, Coursera, Udacity bring you to their own spaces and determine when you can no longer access the materials» (Siemens 2014).

Se i numeri di Siemens erano alti, quelli di Stanford quasi li centuplicano, forse per questo erano stati pensati per un'interazione più limitata secondo il modello che verrà detto xMOOC: erogati da istituzioni di alto livello, non prevedono costi di iscrizione e possono essere fruiti senza alcun piano di studi o propedeuticità con altri corsi. Queste caratteristiche coincidono con quelle accolte in Italia dalla CRUI (Progetto MOOCs Italia - CRUI 2017, p. 2).

Anche se i suoi stessi creatori lo definirono il «nuovo MOOC, diretto dal docente» (Ng – Widom 2014, p. 34), non si deve necessariamente pensare che si trattasse della consapevole ricerca di un modello alternativo al cMOOC: le due esperienze si svolsero quasi contemporaneamente e senza incontrarsi. Ng e colleghi si trovarono

<sup>5</sup> Le OER (Open Educational Resources) sono definite dall'Unesco come «teaching, learning or research materials that are in the public domain or released with intellectual property licenses that facilitate the free use, adaptation and distribution of resources»,  
<<https://en.unesco.org/themes/building-knowledge-societies/oer>>



a dover fronteggiare un numero inaspettato di iscritti, le piattaforme inizialmente non ressero, così rifletterono sulla necessità di progettare secondo principi di scalabilità da un punto di vista sia tecnico che didattico (Ng – Widom 2014, p. 39). Secondo Widom e Ng se si può pensare che in un corso con 100 iscritti i forum vengano letti abbastanza agevolmente da tutti, numeri più grandi richiedono perlomeno l'affinamento di strumenti come l'autocorrezione per i forum per eliminare le ridondanze, che rendono sostenibile lo scambio di idee e diventano molto utili prassi come la valutazione tra pari. Thrun non prevede alcun forum, tanto che i corsisti crearono da loro stessi degli spazi esterni di discussione.

Probabilmente la soluzione è un modello ibrido che combini attività formali e non formali con l'apprendimento informale che, aumentando il coinvolgimento dei partecipanti, motivi a completare il percorso (Fidalgo-Blanco – Sein-Echaluce – García-Peñalvo 2016), come è stato nel caso di EMMA che vedremo oltre.

Bisogna poi dire due parole a proposito della lingua: nati in area angloamericana, su piattaforme alle quali aderiscono le più grandi università di quell'area, certamente anche qui, come altrove, assistiamo ad una preponderanza dell'inglese. Ad oggi, secondo ClassCentral<sup>6</sup>, abbiamo 8061 corsi in inglese contro 896 in spagnolo, 747 in francese, 324 in russo e nessun'altra lingua sopra questo valore. A questo proposito segnaliamo la rilevanza che ha la disponibilità di trascrizioni per aumentare la frequenza dei non madrelingua (Miró – Baquero-Arnal – Civera – Turró – Juan 2018). Per quanto riguarda l'italiano, vedremo che gli esempi non mancano.

## Frequentare un MOOC

Il MOOC può essere immaginato come un contenitore – o meglio un'aula virtuale – dove il discente incontra i suoi pari, docenti, tutor, oltre che materiali e servizi dei quali può fruire in modalità asincrona, organizzandosi autonomamente.

Normalmente la breve presentazione dei corsi in piattaforma descrive contenuti, durata, date di svolgimento, impegno settimanale stimato, lingua ed eventuali prerequisiti. L'iscrizione è aperta a tutti, senza alcuna distinzione, e può prevedere il pagamento di una piccola quota solo nel caso si desideri un attestato finale verificato.

Una volta iscritti di norma si può vedere un Syllabus con una scaletta delle attività di solito a cadenza settimanale e una chiara indicazione delle attività obbligatorie per il completamento del percorso, delle quali talvolta è possibile tener traccia nello spazio personale.

Le lezioni settimanali prevedono nella gran parte dei casi una o due videolezioni per un totale di 40-60 minuti, seguite da un test delle conoscenze basato sui con-

<sup>6</sup> Classcentral (<<https://www.classcentral.com/>>) è un motore di ricerca per i MOOCs con periodici report sulla situazione mondiale e sul loro uso. Le dichiarazioni del sito dicono che vive dei profitti che riceve quando chi si iscrive ai MOOC per loro tramite acquista un certificato verificato.

tenuti. Vengono suggeriti materiali per l'approfondimento che non sono obbligatori ai fini della frequenza, la scelta se andare fino in fondo o fermarsi più in superficie dipende dagli obiettivi didattici personali e dal tempo.

Oltre ai test possono venir richiesti contributi creativi, come post sul forum del corso o la scrittura di testi liberi. Dal momento che un MOOC può avere centinaia di iscritti spesso si ricorre alla valutazione tra pari, che risponde a più obiettivi: dare un ritmo al lavoro del partecipante, fornire un riscontro sulla sua adeguatezza rispetto agli obiettivi del corso, favorire l'interazione tra pari che costituisce un momento importante di relazione. La prassi presenta ovviamente delle criticità dal momento che livelli e aspettative dei corsisti possono essere disomogenei, per questo ad ognuno vengono assegnati più lavori da valutare, ed i valutatori vengono a loro volta valutati, in modo da bilanciare i giudizi troppo soggettivi.

Mediante i MOOCs durano tra 3 settimane e 2 mesi, alcuni hanno precise date di inizio e di fine e un numero limitato di edizioni, mentre altri sono ricorsivi e ci si può iscrivere in qualsiasi momento scegliendo il proprio ruolino di marcia: ovviamente il secondo tipo è meno interattivo. Ad esempio, il corso *Library Advocacy Unshushed*<sup>7</sup> dell'Università di Toronto ha avuto purtroppo solo due edizioni, la pagina tutt'ora disponibile invita gli interessati ad iscriversi per ricevere notizie di una eventuale edizione futura; al contrario, il corso *English Grammar and Style* della University of Queensland<sup>8</sup> viene continuamente riproposto e può essere iniziato in qualunque momento.

A frequentare i MOOC sono spesso, come scoprì subito lo stesso Thrun, persone che già lavorano e desiderano imparare qualcosa di nuovo per motivi personali o professionali, o studenti universitari che desiderano conoscere argomenti non trattati nel loro corso di studi, dai corsi di lingua alla gestione delle relazioni interpersonali, dall'uso del digitale alle nuove normative come quella sulla privacy, dalla storia contemporanea all'intelligenza artificiale, solo per fare qualche esempio.

La frequenza di un MOOC quindi presuppone sostanzialmente la disponibilità di dispositivi e collegamenti di rete adeguati e la capacità di organizzare il proprio tempo. Le diverse piattaforme tengono conto delle raccomandazioni del W3C riguardo all'accessibilità e offrono soluzioni alternative, come la trascrizione dei video e dei materiali, per favorire la partecipazione con collegamenti instabili.

## Le piattaforme

### Le esperienze internazionali: edX, Coursera, Udacity

edX, Coursera e Udacity sono le tre piattaforme per l'erogazione di MOOC più grandi e importanti a livello mondiale: possono contare su grandi investimenti,

<sup>7</sup> <https://www.edx.org/course/library-advocacy-unshushed-university-torontox-la101x>.

<sup>8</sup> <https://www.classcentral.com/course/edx-english-grammar-and-style-2022>.

hanno fra i loro soci le università più prestigiose e insieme servono 50 milioni di studenti in tutto il mondo. Nonostante i diversi modelli di business, hanno alcuni elementi comuni – che di fatto stabiliscono uno standard qualitativo molto alto per tutta l’offerta di MOOC:

- solitamente il corso è offerto gratuitamente, con attestato di frequenza. Si paga una tariffa, che di solito non supera i 100 \$ a corso, solo nel caso si desideri una certificazione validata;
- la certificazione non ha di solito valore di tipo universitario, ma costituisce una motivazione a completare il corso e può essere inserita in piattaforme come LinkedIn;
- Sono disponibili forum di discussione che costituiscono l’aspetto più interessante del corso, perché favoriscono l’interazione con gli altri iscritti al corso e permettono un minimo di comunicazione a due vie; purtroppo spesso non c’è una particolare attenzione da parte dei tutor che si limitano a dare indicazioni di servizio, quindi non tutti i corsi offrono esperienze altrettanto felici;
- la durata dei corsi è solitamente espressa in settimane ed è compresa fra 4 e 6; spesso è consentito completare il corso in minor tempo;
- la maggior parte dei corsi sono in lingua inglese: è molto utile la trascrizione, anche perché non tutti i docenti sono madrelingua o ugualmente comprensibili;
- le piattaforme sono ottimizzate per dispositivi sia fissi che mobili;
- come per i forum, anche l’assistenza è spesso carente se non del tutto assente;
- i contenuti non sono necessariamente aggiornati: utile verificare la data della prima edizione del corso per determinati argomenti, ad esempio in campo tecnologico, dove tre anni possono essere molti;
- hanno una sezione a pagamento dedicata alle aziende, dove si offrono come fornitrici di formazione interna.

Vediamole ora in dettaglio:

### **edX** ([www.edx.org](http://www.edx.org))

Organizzazione senza scopo di lucro, si definisce l’unica piattaforma completamente governata da sole istituzioni accademiche. Fondata nel 2012 dall’Università di Harvard e dal MIT, offre circa 2.000 corsi organizzati dalle migliori università al mondo: oltre alle due già nominate troviamo Berkeley, Delft, il politecnico di Hong Kong, la Sorbona.

Le lezioni video sono accessibili via YouTube e spesso scaricabili per essere usate anche in modalità offline. Molti corsi sono accessibili gratuitamente e comunque sono organizzati per rispondere ad esigenze diverse, ad esempio è possibile frequentare dei MOOC per capire se si hanno competenze di base sufficienti per accedere al campus reale, oppure ottenere crediti accademici presso le università degli Stati Uniti. I corsi “archiviati” sono accessibili ma senza tutoraggio e non offro-

no certificati o esami: i corsi più popolari: “Introduzione alla Computer Science”, “Giustizia” e “La scienza della Felicità”.

Sostiene il progetto Open edX, la piattaforma open-source disponibile liberamente, e si definisce «the only leading MOOC provider that is both nonprofit and open source»<sup>9</sup>.

#### **Coursera** ([www.coursera.org](http://www.coursera.org))

Si tratta della più nota piattaforma profit, quasi un sinonimo di MOOC. Fondata nel 2012 da Daphne Koller e Andrew Ng di Stanford in partnership con Princeton e Yale, comprende la University of Illinois at Urbana-Champaign, Johns Hopkins, la University of Michigan, lo HEC di Parigi, la University of London: sono oltre 150 i partner tra le università. I dati del sito parlano di 35 milioni di studenti e più di 2.700 corsi, con video tradotti in più di 30 lingue. Ha un settore, ovviamente a pagamento, dedicato alle aziende, agli enti governativi ed alle associazioni per la formazione del loro personale.

Offre MOOCs, veri e propri corsi di laurea certificati, raggruppamenti di corsi chiamati “specializzazioni” che sono pagati con una tariffa mensile. È previsto un periodo di prova di sette giorni, per cui chi è molto veloce e motivato può riuscire a completare il corso anche nel periodo di prova. La modalità di valutazione è fra pari: normalmente per poter avere la valutazione del proprio elaborato è necessario avere valutato tre o più elaborati degli altri studenti, e questo a volte può portare piccoli rallentamenti. I corsi più popolari di Coursera sono “Machine Learning”, il consigliabile “Learning How to Learn” e “Introduction to Mathematical Thinking”.

#### **Udacity** (<http://eu.udacity.com>)

Udacity nasce quando Sebastian Thrun - che aveva vaticinato nel 2012 la fine delle Università tradizionali - e Peter Norvig aprono il loro corso “Introduction to Artificial Intelligence”.

Eroga corsi di IT e discipline professionalizzanti, dalle reti neurali all’intelligenza artificiale, dal digital marketing alla linguistica computazionale, che sono in alcuni casi progettati con imprese come Google, Amazon, At&T o IBM che li utilizzano come strumento di reclutamento.

La piattaforma offre corsi gratuiti e serie di corsi a pagamento detti “Nanodegrees”; può essere una buona idea iscriversi ad un corso gratuito per comprenderne il funzionamento - ad esempio, programmazione Java come prova per il più lungo “Android Nanodegree”. I programmi “Nanodegree” di solito prevedono un’attività di tipo pratico e offrono supporto individuale, che ovviamente si riverbera sui costi, compresi tra \$ 500 e \$ 1.000 circa. Tra i corsi più popolari di

<sup>9</sup> <https://www.edx.org/about-us>.

Udacity troviamo “Intro to Computer Science”, “Intro to HTML and CSS” e “Intro to Artificial Intelligence”.

### Le esperienze europee

**European MOOC Consortium** (<https://emc.eadtu.eu/>)

Le principali piattaforme europee (OpenUPed, FutureLearning, Miriadax, Fun e Eduopen) hanno costituito un consorzio nel 2016 per rappresentare complessivamente la loro offerta che raggruppa circa 250 Università ed offre più di 1000 corsi. Il consorzio ha pubblicato un *position paper* (EMC 2018) sull’integrazione dei MOOC nel processo di Bologna in cui si chiede all’Europa di costituire un gruppo di lavoro per integrare i MOOC nei processi educativi formali.

**OpenupEd** (<https://www.openuped.eu>)

OpenupEd è stato lanciato nell’aprile 2013 da EADTU (European Association Distance Education) con il supporto dell’Unione Europea. Sono 11 i partner, con sede in otto paesi dell’Unione (Francia, Italia, Lituania, Paesi Bassi, Portogallo, Slovacchia, Spagna e Regno Unito), e in tre paesi extra-UE (Russia, Turchia e Israele). Si pone come aggregatore di corsi e non offre una piattaforma propria; partito con 40 corsi oggi fornisce accesso ad oltre 200 MOOCs nelle 13 lingue dei partner, oltre all’arabo. Le istituzioni possono associarsi alla piattaforma rispettando certi criteri con una quota pari a 2.500 euro. Per l’Italia è presente l’Università Telematica UniNettuno.

**FutureLearn** (<https://www.futurelearn.com/>)

FutureLearn è la piattaforma inglese per MOOCs fondata nel 2012 da Open University, ente che festeggia oggi i suoi 50 anni nella formazione a distanza e online. Attiva dal 2013, conta nove milioni di utenti ed ha partnership con più di 200 enti nel mondo, tra cui prestigiose università e istituzioni con un ricco archivio di materiale culturale ed educativo, come il British Council, la British Library, il British Museum, la National Film and Television School ed enti privati come Accenture.

I corsi hanno una durata media tra le 6 e le 10 settimane, ma ve ne sono anche di più brevi, di 3 o 5 settimane. Le singole unità contengono video-lezioni, articoli, quiz e specifiche sezioni di discussioni che consentono l’interazione tra gli studenti. L’apprendimento è visto, dunque, come un’esperienza sociale nella quale, tramite specifici spazi e strumenti, i partecipanti entrano in contatto tra loro arricchendo il proprio bagaglio di conoscenze.

A seguito del completamento di una determinata percentuale di lezioni di un percorso, per la maggior parte c’è la possibilità di acquistare un certificato di partecipazione. FutureLearn, inoltre, consente ai propri studenti di sostenere un esame

nei centri dedicati per ottenere un certificato di superamento del corso. Come alcune piattaforme americane anche FutureLearn offre un abbonamento con tariffa *flat* a vari corsi.

Dal 2016 ha iniziato ad offrire crediti universitari e consentire il conseguimento di una laurea online.

In una recente intervista (Bailey – Skelton 2019) Kathryn Skelton, *Chief Strategy Officer* di FutureLearn si focalizza su alcuni punti della strategia della piattaforma, legati al contesto internazionale e all'uso dei MOOC per la formazione continua:

- la formazione continua è sempre più importante, l'università deve offrire opportunità a chi ha già un lavoro: Open University lo fa da sempre, il digitale facilita il compito;
- l'accesso ai corsi online avviene sempre più da dispositivi mobili, fatto che apre nuove modalità di fruizione e richiede adattamenti anche nella progettazione;
- la collaborazione a livello internazionale fra università è cruciale per condividere buone prassi e sperimentare l'impatto di nuovi pubblici;
- una recente ricerca a cura di FutureLearn (Somerville 2018) rivela che i datori di lavoro tengono in grande considerazione la formazione online, per questo motivo la progettazione di nuovi corsi viene fatta anche con le aziende;
- la progettazione è centrata su chi apprende. L'intelligenza artificiale viene utilizzata per automatizzare alcune funzioni della piattaforma e consentirne quindi la scalabilità a costi ridotti, mentre la parte umana viene demandata al rapporto fra pari, al cosiddetto "social learning" che ha una forte correlazione con il tasso di completamento dei corsi.

#### **MiriadaX** (<https://www.miriadax.net>)

MiriadaX, nata all'inizio del 2013 a cura della rete di collaborazione accademica di Spagna e America Latina Universia, è sostenuta dal Banco Santander e dalla divisione Learning Services di Telefónica. È stata la prima piattaforma MOOC per gli studenti di lingua spagnola, e agli oltre quattro milioni di iscritti offre 690 corsi di oltre 100 università, coinvolgendo più di 2.500 insegnanti<sup>10</sup>.

La presenza maggiore è della Spagna, con 39 partner tra i quali l'Universidad Politécnica de Madrid con 41 corsi e l'Universidad de Navarra con 20 corsi. Segue la Colombia con 8 partner, e il Brasile che con 5 partner offre corsi in portoghese. Il corso più popolare del 2016 è stato "Sviluppo di HTML5, CSS e Javascript Web, app mobili", con 292.000 visite, seguito da "Introduzione alla Business Intelligence e ai Big Data", che ha avuto 286.000 visite<sup>11</sup>.

La maggior parte dei corsi sono in spagnolo o in inglese, divisi per moduli, con-

<sup>10</sup> Dati tratti dalla piattaforma MiriadaX alla pagina <<https://miriadax.net/nuestros-numeros>>.

<sup>11</sup> Dati tratti da «Getting to Know MiriadaX, Latin America's Largest MOOC Platform — Class Central» 2017.

tengono video-lezioni e materiale supplementare, hanno durata variabile con un impegno orario settimanale comunicato nel video introduttivo del corso. Anche questa piattaforma offre la certificazione a pagamento dei corsi e propone collaborazioni ad imprese.

#### **France Université Numérique - FUN** (<https://www.fun-mooc.fr/>)

FUN (France Université Numérique) è un ambiente lanciato dal Ministero dell'Istruzione Superiore e della ricerca francese nel 2013 con il contributo di tre partner pubblici, INRIA, CINES e RENATER e di soggetti specializzati del sistema universitario francese. L'intento del progetto è quello di riunire in un'unica piattaforma i MOOCs delle università e delle scuole francesi garantendo libero accesso ai contenuti dei corsi. Sono circa 1 milione gli iscritti ai 150 corsi offerti dalle 50 università che aderiscono alla piattaforma.

Il livello di collaborazione e interazione dell'esperienza di apprendimento per gli studenti di FUN dipende dal livello individuale di conoscenza e abilità disciplinare. È necessaria l'iscrizione per usufruire della vasta gamma di argomenti e corsi che soddisfano le diverse esigenze di apprendimento; i corsi hanno cadenza settimanale e sono suddivisi in ulteriori unità.

La piattaforma FUN-MOOC offre MOOCs aperti a tutti, mentre, FUN Campus<sup>12</sup> e FUN Corporate<sup>13</sup> sono rivolti a pubblici specifici: FUN Corporate è orientata alle aziende che erogano formazione a dipendenti, clienti e partner; FUN Campus è la versione riservata agli studenti universitari o delle scuole superiori, dove solo gli insegnanti responsabili delle lezioni possono autorizzare l'accesso agli studenti che utilizzano un loro identificativo personale, così i loro dati restano a disposizione dei docenti e i punteggi ottenuti possono essere riconosciuti dalle istituzioni di provenienza ai fini della valutazione. Entrambe, oltre ai MOOCs, distribuiscono SPOC (*Small Private Online Course*), che supportano la trasformazione della didattica adottando metodologie come la classe capovolta o lo studio per gruppi di apprendimento collaborativo.

#### **EduOpen** (<https://learn.eduopen.org>)

EduOpen è un progetto finanziato dal Ministero dell'Istruzione, dell'Università e della Ricerca, finalizzato alla realizzazione di una piattaforma per l'erogazione di MOOC da parte di un network di atenei Italiani e di un insieme di partner selezionati. Ad oggi conta 20 istituzioni, 259 docenti, 68 Tutor, 42 editor di contenuti, più di 50 mila utenti per 245 corsi offerti e 30 pathway (un insieme coordinato di MOOC finalizzato a certificare Master o Corsi di specializzazione). La rete costituita è aperta all'adesione di altre università ed enti/associazioni/reti di particolare

<sup>12</sup> <https://www.fun-campus.fr/>.

<sup>13</sup> [www.fun-corporate.fr](http://www.fun-corporate.fr).

rilievo scientifico e culturale.

**Open education Italia** (<https://educazioneaperta.eu/wp/>)

Si tratta di una rete di educatori che, senza scopo di lucro, lavorano insieme per sviluppare le risorse educative aperte – le OER – e l’inclusione nella pratica educativa. Interessante il “Corso OERuP!”<sup>14</sup>, gratuito, che spiega come reperire e come creare queste risorse. La rete invita all’adesione chiunque sia interessato.

**iversity** ([www.iversity.org](http://www.iversity.org))

Si tratta di una piattaforma educativa che offre corsi online in collaborazione con università, aziende e organizzazioni non governative (ONG). L’azienda, originariamente fondata nel 2011, sul proprio sito dichiara di essere stata la prima piattaforma europea ad offrire i Massive Open Online Courses (MOOC) e la prima a livello mondiale ad offrire crediti formativi universitari a livello europeo (ECTS). Nel 2013, grazie ad un finanziamento della fondazione tedesca *Stifterverband für die Deutsche Wissenschaft*, iversity aveva messo a disposizione un finanziamento per offrire ai docenti interessati la produzione di MOOCs. Fu una bella operazione di marketing perché le proposte furono poi votate dal web. Ci furono molte candidature italiane e fra i dieci vincitori due erano progetti italiani.

Altre piattaforme, che qui abbiamo solo lo spazio di nominare, sono:

**Udemy** ([www.udemy.com](http://www.udemy.com)), un grande mercato della formazione online dove chiunque può offrire un proprio corso;

**Cognitive Class.ai** (<https://cognitiveclass.ai/>), un’iniziativa IBM per diffondere la conoscenza dei big data;

**Alison** (<https://alison.com>) si definisce come un’impresa che crede nell’educazione aperta, offre circa 1000 corsi gratuiti e si finanzia con la pubblicità e con alcuni servizi a pagamento;

**Eliademy** (<https://eliademy.com>) *open source*, finlandese, prevede sia servizi gratuiti che a pagamento.

**La piattaforma EMMA: uno sguardo dall’interno**  
**([www.europeanmoocs.eu](http://www.europeanmoocs.eu))**

EMMA è l’acronimo di “European Multiple MOOC Aggregator”, un progetto co-finanziato dall’Unione Europea svoltosi tra il 2014 e il 2016, coordinato

<sup>14</sup> <https://educazioneaperta.eu/wp/corso-oerup/>.



dall'Università di Napoli Federico II e che coinvolgeva altre quattro università (University of Leicester nel Regno Unito, l'Université de Bourgogne in Francia, l'Università di Tallin, in Estonia, e l'Universitat Politècnica di València, in Spagna), tre "open university" che, a differenza sia dagli istituti tradizionali sia dalle università telematiche, si basano sul concetto di apprendimento continuo (l'Universitat Oberta De Catalunya in Spagna, la Open Universiteit Nederland nei Paesi Bassi e l'Universidade Abierta in Portogallo), alcuni enti di ricerca privati e anche piccole e medie imprese (PMI). Inoltre, l'opportunità di un'offerta gratuita aperta ad enti esterni per una valutazione di prova aveva permesso a più di altre venti entità, europee e non, di valutare la piattaforma.

EMMA fornisce una piattaforma per ospitare corsi online gratuiti e aperti a tutti, in diverse lingue europee, con lo scopo di conservare e valorizzare il ricco patrimonio culturale, educativo e linguistico dell'Europa promuovendo un apprendimento transculturale e multilingue. I MOOC presenti su EMMA hanno una durata variabile tra le 6 e le 8 settimane e sono costituiti da lezioni divise per unità. Per favorire un approccio transculturale e plurilinguistico i corsi sono offerti secondo una logica di "cluster", su temi contigui e, grazie ad un sistema di traduzione automatica, sono disponibili in due o tre lingue a partire da quella del paese di origine. Gli studenti di EMMA sono accompagnati nella creazione di un progetto di apprendimento personalizzato tramite lo strumento della "toolbox", che permette di costruire percorsi autonomi di apprendimento utilizzando porzioni di corsi diversi e contigui. L'interazione tra pari e tra studenti e docenti tramite blog e post era un elemento caratterizzante di EMMA volto ad arricchire l'esperienza formativa: l'obiettivo finale era creare una vetrina per le eccellenze e le innovazioni nei metodi di insegnamento e negli approcci all'apprendimento, realizzando un'esperienza pilota su larga scala. Le funzionalità di trascrizione automatica dal video al testo, con traduzione automatica, hanno costituito l'ambito di ricerca più promettente del progetto, grazie al contributo del Politecnico di Valencia: permettendo una riduzione del tempo di revisione dei testi di circa un decimo nella trascrizione, e del 50% nella traduzione, rappresentava un indubbio supporto ad una visione europea della formazione superiore (Miró – Baquero-Arnal – Civera – Turró – Juan 2018).

Nel progetto EMMA una specifica attenzione era posta sull'analisi dei dati. Da un lato veniva misurata la percezione soggettiva dei partecipanti attraverso questionari pre- e post-corso; dall'altro un software di tracking sviluppato ad hoc poteva misurare le interazioni dello studente in piattaforma con diverse modalità. Era possibile di monitorare con diverse modalità, ad esempio, gli abbandoni, i livelli di interazione, l'utilizzo degli strumenti e delle attività. Diverse profilazioni socio-demografiche erano ricavabili sulla base delle informazioni ottenute in fase di registrazione alla piattaforma, e si disponeva di una serie completa di dati sulle aspettative e le esigenze raccolte tramite un questionario. L'idea era quella di confrontare dati soggettivi e oggettivi sulla fruizione dei contenuti della piattaforma.

Il progetto EMMA era stato finanziato in parte nell'ambito del Programma quadro CIP (Competitività e Innovazione) della Commissione Europea<sup>15</sup>, mentre il resto di quanto necessario a rendere operativa la piattaforma proveniva dai partner stessi.

Alla fine del progetto, nel 2016, la piattaforma aveva avuto 150.000 visitatori unici per quasi 4 milioni di pagine visitate, 36 MOOC attivi realizzati da 174 docenti e tutor, circa 30 mila utenti registrati; 21 MOOC avevano avuto fra i 150 e i 500 iscritti per un totale di circa 7.000, 5 MOOC contavano fra i 500 e 1.000 iscritti, 2 MOOC avevano fra 1.000 e 5.000 iscritti ed 1 MOOC superava i 5.000. Le lingue erano, oltre all'inglese, francese, spagnolo, italiano, olandese, estone e portoghese. Il 76% degli utenti erano donne, il 46% aveva un'età compresa fra i 46 ed i 57 anni e gli insegnanti erano il 47%. Due terzi dell'utenza non aveva esperienza pregressa e il 41% di questi erano docenti di area STEM. In particolare, l'82% degli iscritti al corso di *coding* erano donne, cioè 4.000 persone.

L'analisi dei dati relativamente alle modalità d'uso della piattaforma era stata fatta definendo 4 profili: *Iscritti*, coloro che si erano collegati alla piattaforma almeno 5 volte; *Osservatori*, coloro che si erano collegati più di 5 volte ma senza interagire con altri iscritti; *Contributori*, coloro che avevano consegnato un compito o scritto almeno un post; *Attivi*, coloro che avevano consegnato compiti e scritto post più di una volta.

Questa segmentazione era stata applicata ad alcuni corsi per valutare quali caratteristiche di MOOC ne condizionavano la frequenza, in quanto la piattaforma era progettata per fornire supporti ai diversi approcci pedagogici (xMOOC e cMOOC). Ad esempio, il progetto MOOC "Progettare corsi online con framework 7Cs" durava 6 settimane ed era basato principalmente su una funzionalità: la conversazione. Gli studenti avevano una varietà di opzioni per utilizzare lo strumento per le discussioni, presentare compiti, fare domande, confrontarsi con i colleghi e così via: in questo MOOC ad esempio la percentuale del profilo *Iscritti*, cioè a bassa interazione, era del 71% mentre sommando i *Contributori* e gli *Attivi* si arrivava ad un 15%. Il corso sul *coding*, che durava 13 settimane e per l'intensa interazione si affidava ai social media esterni, aveva il 27% di *Osservatori*, il 32% di *Contributori* e il 7% di *Attivi*.

In generale tutti gli strumenti innovativi forniti dalla piattaforma venivano accolti in modo positivo dagli utenti: il multilinguismo, la "toolbox" ma soprattutto i "cruscotti", che permettevano di personalizzare l'esperienza di apprendimento, molto apprezzati da studenti e insegnanti.

EMMA ha dato l'opportunità a chi lo desiderava di testare gratuitamente una nuova forma di insegnamento e aprire nuovi orizzonti, come è stato per il corso "Digital Library" tenuto da Anna Maria Tammaro, che dice: «Secondo me, il ruolo dell'inse-

<sup>15</sup> Il programma quadro per la competitività e l'innovazione (CIP) della Commissione europea è archiviato al sito <[http://ec.europa.eu/cip/index\\_it.htm](http://ec.europa.eu/cip/index_it.htm)>.

gnante e il modo in cui può essere supportato dalla piattaforma è di primaria importanza per il MOOC. L'insegnante MOOC è spesso necessario per svolgere molteplici ruoli: insegnante, designer, mentore, marketing istituzionale ecc. in un ambiente completamente trasparente, e quindi con un rischio alto. La pedagogia del MOOC non è solo qualcosa che ha incorporato la piattaforma online, insegnare in un MOOC è diverso dall'insegnare in una classe. Avrei potuto assegnare un ruolo più rilevante ai partecipanti nei criteri di valutazione del MOOC, mi sarebbe piaciuto stabilirli in modo partecipativo fin dalla prima lezione» (Tammaro 2016).

Le molteplici caratteristiche innovative di EMMA costituivano al tempo stesso un'opportunità ma anche una grande sfida, in un momento in cui si stavano rafforzando alleanze e progetti nazionali di grande portata come il già ricordato FUN<sup>16</sup> francese. EMMA avrebbe potuto essere una grande opportunità per l'Europa, ma soprattutto per l'Italia, visti i risultati eccezionali nella formazione per gli insegnanti, infatti EMMA ha permesso a migliaia di loro di approcciarsi al fenomeno MOOC, ad oggi ancora poco diffuso fra studenti e professionisti. La piattaforma EMMA è ancora attiva tre anni dopo la chiusura del progetto, aperta a sperimentazioni con altri soggetti pubblici e privati, anche se sicuramente le questioni legate alla proprietà intellettuale della piattaforma e dei contenuti, oltre che l'assenza di un sostegno pubblico in termini di politiche e di risorse, ne condizionano fortemente la sopravvivenza.

### Le Università Italiane

Nel 2015 la Fondazione CRUI pubblicò un rapporto sulla situazione dei MOOCs nelle Università italiane (CRUI 2015) che fotografava una situazione in forte movimento: l'offerta era raddoppiata nel biennio 2013-14 (da 18 a 39 MOOC) con circa un centinaio di corsi attivati dagli Atenei italiani a fine 2015. L'indagine CRUI svolta a fine 2014 aveva ricevuto 55 risposte dal 70% dei 79 atenei invitati a compilare il questionario online, che rilevava anche le piattaforme utilizzate: *Moodle* per Ferrara, Foggia, Roma Tor Vergata; *Coursera* per Roma La Sapienza – la prima in Italia ad entrarvi - Milano Bocconi e Politecnico; *Iversity* per Padova e Foggia; il Politecnico di Milano utilizzava anche *OpenEdx*, mentre Catania e Urbino usavano *YouTube*; *Future Learn* utilizzata dall'Università per Stranieri di Siena e di nuovo Padova; la Federico II di Napoli aveva personalizzato la propria offerta per offrire il servizio Federica.eu ed il progetto europeo EMMA.

Il rapporto evidenziava punti di forza e di debolezza dell'uso dei MOOC per la didattica universitaria e tratteggiava i possibili destinatari di tali servizi: lavoratori che potrebbero conciliare i propri tempi di studio e lavoro; studenti indecisi sul percorso di studio; studenti stranieri che potevano così avvicinarsi all'offerta di

<sup>16</sup> V. sopra tra le esperienze europee.

istruzione superiore italiana. L'analisi prendeva in considerazione anche dati relativi ai costi, alle figure professionali impiegate e agli studenti, e si chiudeva con la proposta di arrivare alla definizione di linee guida per gli Atenei italiani finalizzate a coordinarne l'offerta sul territorio nazionale.

A seguito del rapporto del 2015 la Fondazione CRUI accoglieva la proposta di creare strumenti per il coordinamento dell'offerta delle università italiane e avviava il progetto MOOCs Italia, coordinato da Flavio Corradini, attraverso la costituzione di tre tavoli di lavoro che nel 2017 pubblicano tre documenti: le "Linee guida nazionali per la predisposizione di MOOCs di qualità erogati dalle Università italiane"<sup>17</sup>, che definisce i criteri ma apre anche ad alcuni aspetti problematici, l'"Accordo Quadro per il Coordinamento tra Atenei Italiani aderenti al Progetto MOOCs Italia"<sup>18</sup>, che si basa sulla costituzione dell'Osservatorio CRUI sui MOOC, di cui pubblica il "Regolamento dell'Osservatorio CRUI e relativa Bozza di Statuto"<sup>19</sup>. Purtroppo, si deve constatare che il suddetto Osservatorio non risulta ancora effettivamente attivato.

Nell'ambito de "I magnifici incontri CRUI" nel 2018 la Fondazione CRUI e il MIUR hanno riunito più di 250 rappresentanti che hanno lavorato ad «8 tavoli tematici volti a porre al centro dell'alta formazione le tecnologie digitali come nuovo linguaggio per l'apprendimento continuo»<sup>20</sup> e definire il Piano Università Digitale.

Tutti i tavoli si sono occupati a vario titolo delle tecnologie digitali a supporto dell'apprendimento: il tema dei MOOC è stato citato a più tavoli, non solo a quello che vi era dedicato, ed è stata delineata un'azione specifica per il rafforzamento dell'offerta, che si vede sia come opportunità per avvicinare la cittadinanza all'offerta universitaria tradizionale, sulla scia della Terza Missione, che come strumento di orientamento nella scelta del percorso di studi<sup>21</sup>.

Avere un proprio corso ospitato su una piattaforma internazionale è una scelta valida per un ateneo: essere su Coursera significa aprire una vetrina mondiale per l'offerta di formazione online. I costi vivi non sono sicuramente contenuti, tuttavia l'investimento può valere la spesa, considerando che di solito un ateneo non ha la totalità dei propri corsi su piattaforme di tale portata, ma solo un numero limitato.

<sup>17</sup> Progetto MOOCs Italia - CRUI 2017.

<sup>18</sup> Progetto MOOCs Italia - CRUI 2017.

<sup>19</sup> Progetto MOOCs Italia - CRUI 2017.

<sup>20</sup> CRUI, *I Magnifici Incontri CRUI 2018 - Piano Nazionale Università Digitale. Atti*, p. 2, <[https://www2.cruai.it/cruai/magnifici\\_incontri\\_cruai\\_2018/ATTI\\_I\\_MAGNIFICI\\_INCONTRI\\_CRUI2018.pdf](https://www2.cruai.it/cruai/magnifici_incontri_cruai_2018/ATTI_I_MAGNIFICI_INCONTRI_CRUI2018.pdf)>.

<sup>21</sup> CRUI, *I Magnifici Incontri CRUI 2018*, cit., p. 22. Si noti che i MOOCs come strumenti di orientamento possono essere utili anche agli stessi Atenei per decidere quali corsi attivare, si veda Pozzi - Conole 2014.

## MOOC AIB come esperienza di formazione professionale informale

A fine 2018 l'Associazione Italiana Biblioteche (AIB) si trovava nella necessità di formare delle figure in grado di assistere gli associati nel rinnovo dell'attestazione - per il quale andavano inviate le domande entro il 15 gennaio 2019 - e sul modello di formazione sviluppato dopo la L. 4/2013.

Facciamo un breve passo indietro per spiegare di cosa si tratta: a seguito della Legge 4/2013, l'AIB si era iscritta presso il Ministero dello Sviluppo Economico all'elenco delle associazioni professionali che rilasciano l'attestato di qualità e di qualificazione professionale dei servizi prestati dai propri associati. La prima tornata di attestazione si era svolta nel 2014 attraverso una selezione basata essenzialmente sui titoli, ma la Direttiva approvata dal Comitato Esecutivo Nazionale (CEN) nella seduta del 29 dicembre 2015 aveva avviato un processo che, a inizio 2019, avrebbe portato per la prima volta ad una procedura basata sulle competenze, che prevedeva tra gli strumenti la possibilità di compilare un portfolio<sup>22</sup>. L'AIB è un'associazione di volontari, presente su tutto il territorio nazionale suddivisa in sezioni regionali; gli organi di governo erano stati rinnovati nel 2017 con l'elezione di molti nuovi consiglieri che non erano necessariamente informati sul processo. C'era perciò il bisogno di formare, in poco tempo e in tutte le sezioni, delle persone che potessero assistere gli associati nella procedura.

Matilde Fontanin, componente esperta dell'Osservatorio Formazione, preparò per il CEN il progetto di un corso basato sul modello MOOC, adattato alla situazione. Si prevedeva di utilizzare la piattaforma Moodle AIBformazione.it, già vivacemente popolata, di sfruttare le competenze di alcuni membri dell'Osservatorio Formazione - Manuela De Noia e la stessa Matilde Fontanin - in qualità di docenti, e di affidare tutoraggio e organizzazione ad associati qualificati, che avevano acquisito le loro competenze in seno all'Associazione nei percorsi intrapresi dopo il 2014.

Il corso ricalcava il modello MOOC: videolezioni settimanali, test delle conoscenze, materiali di approfondimento e attività individuale e di gruppo in piattaforma. Si discostava dalle definizioni viste sopra in quanto prevedeva un numero massimo di partecipanti, dei prerequisiti e una quota di iscrizione, per quanto piccola. Queste limitazioni erano dettate dall'esiguità di risorse finanziarie, non paragonabili a quelle, ad esempio, di un Ateneo.

Il corso, svoltosi tra il 5 novembre e il 17 dicembre, prevedeva 100 ore tra lezioni, studio individuale e lavori di gruppo, scandite in 6 settimane. Nelle prime tre si presentavano i contenuti in videolezioni seguite da test delle conoscenze e dibat-

<sup>22</sup> La *Direttiva sulla Formazione continua* alla quale si fa riferimento è quella approvata dal CEN nella seduta del 29 dicembre 2015, ora sostituita da quella approvata il 16 novembre 2018, che affianca al portfolio altre forme meno riflessive di autocertificazione. Il documento è disponibile in rete a: <http://www.aib.it/chi-siamo/statuto-e-regolamenti/direttiva-formazione-continua>.

titi nei forum, nelle ultime tre il forum di gruppo e lo strumento *compito*<sup>23</sup> venivano usati per guidare i partecipanti nella compilazione del loro portfolio individuale. L'organizzazione fu un lavoro di gruppo: le docenti predisposero le videolezioni, i materiali, la scaletta delle attività didattiche; la webmaster, Maria Accarino, pensò all'apertura della piattaforma e all'iscrizione dei partecipanti, compresa la stesura di un suo report finale; l'organizzatrice del corso, Paola Monno, si occupò di raccogliere le iscrizioni, rispondere alle domande degli interessati, curare i rapporti sia con la Segreteria di AIB Nazionale che con la webmaster; il caricamento dei materiali in piattaforma in genere è a cura delle tutor, ma, vista l'esperienza delle docenti stesse, se ne occupò tutto il team, comprese le 3 tutor coinvolte a fronte dei 60 iscritti: Martina Contessi, Immacolata Murano e la stessa Paola Monno.



Figura 1. La pagina del corso AIB su [www.aibformazione.it](http://www.aibformazione.it)

L'attività fu organizzata secondo le Linee Guida AIB (AIB 2016) che recepiscono gli standard formativi e professionali e delineano gli obiettivi formativi sulla base delle competenze. Tali obiettivi consistevano in:

CONOSCENZE	<p>A1) Individuare e consultare i provvedimenti di legge e le direttive del CEN AIB che hanno portato alla nascita del corrente modello di formazione;</p> <p>A2) Analizzare e applicare i requisiti per l'attestazione professionale;</p> <p>A3) Trasmettere le modalità di preparazione di autocertificazione, CV e portfolio.</p>
CAPACITÀ	<p>B1) Essere in grado di rispondere o reindirizzare le domande degli associati della propria Sezione in merito agli strumenti per l'attestazione (incluso il portfolio);</p> <p>B2) Essere in grado di confrontarsi con gli associati in merito a domande sul processo di attestazione (nascita, sviluppo, stato attuale);</p> <p>B3) Essere in grado di collaborare alla pianificazione di percorsi di formazione nella propria Sezione utilizzando correttamente i modelli AIB</p>

<sup>23</sup> Il "compito" è lo strumento di Moodle che consente di predisporre la consegna e la valutazione di elaborati.

Il percorso fu effettivamente seguito da 55 partecipanti e portato a termine da 49. Alla chiusura fu svolta, come da modello formativo, un'indagine sul gradimento del corso, con domande sulla percezione dei corsisti rispetto ai contenuti e all'organizzazione del corso, alle relazioni tra pari e con i docenti e tutor, oltre ad uno spazio aperto ai commenti. Il giudizio veniva espresso con un punteggio da 1 a 5, si riportano di seguito le medie per le 44 risposte anonime:

organizzazione del corso:	4
quantità di ore dedicate al corso:	4
strutturazione del corso:	4
scelta degli argomenti:	4,7
materiali distribuiti (bibliografia):	4,6
ambiente di apprendimento:	4
video - lezioni:	4,2
interazione tutor - gruppo:	4,8
interazione con i compagni di corso:	3,4
interazione con i compagni di gruppo:	3,7
valutazione del corso rispetto alle esigenze del corsista:	3,7
valutazione del corso rispetto alle aspettative del corsista:	3,6
valutazione del corso rispetto al costo:	4
utilità rispetto all'ambiente lavorativo:	3,2

Manca lo spazio qui per compiere delle riflessioni sulla base dei risultati, ma si può affermare che il giudizio in generale è positivo e che gli aspetti meno graditi sembrano essere le interazioni tra pari, mentre il ruolo delle tutor ha avuto il gradimento maggiore. Quanto ai commenti aperti, molti sono di persone che ritengono l'esperienza significativa per la loro crescita professionale, ma alcuni evidenziano delle criticità, come la fatica a rispettare le scadenze, la quantità di lavoro richiesta, la difficoltà a seguire i numerosi messaggi dei forum. Riprendendo il commento di Widom e Ng riportato sopra<sup>24</sup>, la nostra esperienza ci mostra, come già in altre occasioni, che anche un forum di gruppo composto da 20 persone, delle quali non tutte ugualmente attive, genera un numero di post che è difficile seguire, perlomeno nel contesto della formazione professionale: è un aspetto sul quale si può lavorare, ma che fa parte del gioco.

Riguardo ai tempi di svolgimento e la quantità di lavoro, parlando di formazione professionale il dilemma è il conflitto tra la necessità di svolgerla e il poco tempo a disposizione. La formazione in presenza assicura che chi partecipa sia materialmente presente, ma non sempre se ne monitora la ricaduta sulle competenze indi-

<sup>24</sup> Ci si riferisce alla considerazione che con 100 partecipanti si può pensare che i messaggi vengano letti abbastanza agevolmente da tutti (Ng - Widom 2014).

viduali e sulle prassi lavorative. I corsi online apparentemente sono più facili da frequentare, ma richiedono in aggiunta uno sforzo organizzativo che, se certo deve essere facilitato dal datore di lavoro, resta personale – il corsista deve impegnarsi a riorganizzare il proprio impegno lavorativo<sup>25</sup>.

In questa esperienza, tuttavia, i risultati finali, il numero di portfolio<sup>26</sup> presentati e la qualità dei lavori dicono che l'esperienza è stata positiva, ha raggiunto molte persone in tutta Italia, le ha fatte incontrare. Questo si è rivelato fondamentale anche alla luce della revisione della direttiva sulla formazione approvata dal CEN<sup>27</sup> a corso già iniziato che si è potuta affrontare come una comunità di lavoro, con uno spazio per domande e risposte.

L'organizzazione di un tale corso è complessa e il presente, breve riepilogo non pretende di riuscire a darne conto, ma fornisce un esempio di come, con poche risorse, adattando il modello MOOC si possa rispondere ad un bisogno di formazione sul territorio nazionale per un numero significativo di persone.

### **Project Mu.Sa: per i musei**

Sempre nel settore dei beni culturali, ma passando ai musei, segnaliamo un'iniziativa di respiro internazionale in corso, il Project Mu.Sa che si propone di chiudere il divario creatosi tra la formazione formale e il mondo del lavoro «due to the quickening pace of the adoption of ICT in the museum sector»<sup>28</sup>. Anche qui vediamo un modello promettente, finanziato da Erasmus+, corrente: potrebbe essere uno spunto da cogliere.

## **Conclusioni**

Il saggio si proponeva di descrivere i MOOCs a chi li conoscesse solo superficialmente, con particolare riferimento all'Italia di oggi e con uno sguardo sulla formazione professionale. Come abbiamo visto il panorama è ampio, ci sarebbe molto da dire sugli aspetti pedagogici, andragogici, organizzativi, finanziari, ma questo esulerebbe dagli intenti delle autrici oltre che dallo spazio offerto dalla rivista. Se, come ci auguriamo, siamo riuscite nel nostro scopo, vorremmo chiudere con delle osservazioni.

L'apprendimento online, o e-learning, viene talvolta confuso con la magia: si pensa che accedere ad un corso online equivalga ad apprendere. Esattamente come per le esperienze di formazione in presenza, l'atto di apprendere implica un cambiamento da parte del discente. Inoltre, strutturare un corso online richiede un ri-

<sup>25</sup> Sulle criticità e le barriere culturali nell'uso dell'e-learning nella formazione professionale si riflette da anni, si veda ad esempio Fontanin 2011.

<sup>26</sup> Per approfondire la tipologia di portfolio delle competenze atteso si veda AIB 2017.

<sup>27</sup> AIB 2018.

<sup>28</sup> <http://www.project-musa.eu/musa-in-180-words/>.



pensamento da parte del docente rispetto alle proprie prassi consolidate, e non tutti sono disposti a farlo. Per questo assistiamo a casi dove il semplice fatto di caricare in rete dei materiali viene definito e-learning, e il semplice fatto di averli scaricati viene considerato apprendimento, mentre dobbiamo partire dal principio che la forma digitale in sé non cambia la natura dell'esperienza. Chi organizza in questo modo i corsi online crea un danno tanto più grave perché rafforza la convinzione che l'e-learning non funziona, mentre neanche i corsi in presenza funzionano di fronte a un atteggiamento passivo, perché apprendere significa venire coinvolti e cambiare, e questo costa sempre tempo e fatica, che si affrontano con la giusta motivazione.

Secondo il report del Pew Research Centre 2016, il 74% degli americani adulti ha partecipato a qualche tipo di formazione personale volontaria (Li 2019), il che poco sorprende in un mondo del lavoro e della conoscenza che è in continua evoluzione. Con questo atteggiamento i MOOCs possono costituire una preziosa opportunità per l'apprendimento continuo e per la formazione professionale, e potrebbero risultare risolutivi laddove vi sia il desiderio o la necessità di formarsi relativamente a temi specifici ed esistano barriere fisiche o di organizzazione del tempo – di chi lavora a tempo pieno come di chi lavora a chiamata, di chi è iscritto un corso di studio con obbligo di frequenza o di chi lavora anche a tempo parziale ma è impossibilitato ad assentarsi dal lavoro.

Perché questa opportunità si realizzi, tuttavia, occorre che vi sia chiarezza sui requisiti e sulle necessità, che gli enti che li erogano siano per primi consci che essi richiedono un diverso sforzo organizzativo, e che ne venga riconosciuto il valore. D'altro canto, molte delle difficoltà si possono risolvere consorziandosi e lavorando in rete per preparare un'offerta che risponda alle esigenze di informazione e formazione su questioni specifiche, anche appoggiandosi a quei consorzi e piattaforme già esistenti che abbiamo nominato sopra, e questo è il nostro auspicio.

*MOOCs – Massive Open Online Courses – are online courses born under the house of OERs – Open Educational Resources – and characterized by large number of participants from diverse nationalities, education, cultural background. Born within the US academic context on the aftermath of e-learning and the digital revolution, from the early beginnings to their launch, from the first enthusiasms to the following challenges, their growth saw different models and participatory approaches and a variety of platforms. They were offered by many institutions, from large universities to institutes for research and education. Nowadays they are still provided in large number, and they offer a precious opportunity for personal and professional development in a society which is becoming increasingly aware of the need for lifelong learning.*

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

## RIFERIMENTI BIBLIOGRAFICI

- AIB 2016 Associazione italiana biblioteche. *Linee guida per la formazione continua AIB. AIB-WEB*, 16 marzo 2016.  
<<http://www.aib.it/struttura/osservatorio-formazione/2016/55066-linee-guida-la-formazione-continua-aib/>>.
- AIB 2017 Associazione italiana biblioteche. *Il portfolio delle competenze: un nuovo strumento per il professionista dell'informazione*, a cura dell'Osservatorio formazione, coordinatore P. Lùperi, contributi di M. De Noia, M. Fontanin, P. Lùperi. Roma: Associazione italiana biblioteche, 2017.
- AIB 2018 Associazione italiana biblioteche. *Direttiva sulla Formazione continua: documento approvato dal CER nella seduta del 16 novembre 2018*. In: AIB-WEB:  
<<https://www.aib.it/chi-siamo/statuto-e-regolamenti/direttiva-sulla-formazione-continua/>>.
- Bailey – Skelton 2019 Sophie Bailey. *Education 4.0: The FutureLearn Perspective - Interview with Kathryn Skelton*. In: *The Edtech Podcast* (blog). 3 aprile 2019.  
<<https://theedtechpodcast.com/education-4-0-the-futurelearn-perspective/>>.
- CRUI 2015 Conferenza dei Rettori delle Università italiane. *MOOCs - Massive Open On-Line Courses: Prospettive e Opportunità per l'Università Italiana*. 2a ed. Roma: Fondazione CRUI, 2015.  
<<https://www.cruai.it/atenei-e-rettori-cruai/atenei-per-regione/item/353-moocs-massive-open-on-line-courses.html>>.
- Deng – Benckendorff – Gannaway 2019 Ruiqi Deng — Pierre Benckendorff — Deanne Gannaway. *Progress and new directions for teaching and learning in MOOCs*. «Computers & Education», 129 (2019), p. 48–60.  
<<https://doi.org/10.1016/j.compedu.2018.10.019>>.
- EMC 2018 European MOOC Consortium. *Lifelong Learning and the EHEA Agenda: Contribution of the European MOOC Consortium*. 2018.  
<<https://tinyurl.com/yxuutktv>>.
- Fidalgo-Blanco – Sein-Echaluze – García-Peñalvo 2016 Ángel Fidalgo-Blanco — María Luisa Sein-Echaluze — Francisco José García-Peñalvo. *From massive access to cooperation: lessons learned and proven results of a hybrid xMOOC/cMOOC pedagogical approach to MOOCs*. «International Journal of Educational Technology in Higher Education», 13 (2016), n. 1, p. 24.  
<<https://doi.org/10.1186/s41239-016-0024-z>>.
- Fontanin 2011 Matilde Fontanin. *Inglese professionale in modalità blended learning: principi pedagogici e vantaggi di Moodle*. In: *MoodleMoot*. Trieste, 23-24 giugno 2011.  
<<https://tinyurl.com/y52u2bym>>.

- Gaggi 2013 Massimo Gaggi. *Dietrofront, l'università online non funziona*. «Corriere della Sera», 15 dicembre 2013. <[http://www.corriere.it/scuola/13\\_dicembre\\_15/universita-online-non-funziona-mooc-sebastian-thrun-50b2d04a-6579-11e3-95f1-73e6b5fcc151.html](http://www.corriere.it/scuola/13_dicembre_15/universita-online-non-funziona-mooc-sebastian-thrun-50b2d04a-6579-11e3-95f1-73e6b5fcc151.html)>.
- FutureLearn 2018 *Online learning bolsters chances of promotion, reveals new research*. «FutureLearn», 27 settembre 2018. <<https://about.futurelearn.com/press-releases/online-learning-bolsters-chances-of-promotion-reveals-new-research>>.
- Hollands – Tirthali 2014. Fiona M. Hollands — Devayani Tirthali. *MOOCs: Expectations and Reality. Full Report*. New York: Center for Benefit-Cost Studies of Education, Teachers College, Columbia University, 2014. <<https://tinyurl.com/y58runhr>>.
- Kenny – Gunter 2014 Robert Kenny — Glenda Gunter. *MOOCs: The New EDU Imperative*. In: *Proceedings of SITE 2014 - Society for Information Technology & Teacher Education International Conference*, a cura di M. Searson, M. Ochoa. Jacksonville: Association for the Advancement of Computing in Education (AACE), 2014, p. 1462–1463, <<https://www.learntechlib.org/primary/p/131234/>>.
- Li 2019 Kun Li. *MOOC Learners' Demographics, Self-Regulated Learning Strategy, Perceived Learning and Satisfaction: A Structural Equation Modeling Approach*. «Computers & Education», 132 (aprile 2019), p. 16–30. <<https://doi.org/10.1016/j.compedu.2019.01.003>>.
- Merrill 2002 David Merrill. *First Principles of Instruction*. «Educational Technology Research and Development», 50 (2002) n. 3, p. 43–59. <<https://doi.org/10.1007/BF02505024>>.
- Miró – Baquero-Arnal – Civera – Turró – Juan 2018 Juan Daniel Valor Miró — Pau Baquero-Arnal — Jorge Civera — Carlos Turró — Alfons Juan. *Multilingual Videos for MOOCs and OER*. «Journal of Educational Technology & Society», 21 (2018) n. 2, p. 1. <<https://www.jstor.org/stable/26388375>>.
- Ng – Widom 2014 Andrew Ng — Jennifer Widom. *Origins of the Modern MOOC (xMOOC)*. In: *MOOCs: expectations and reality. Full report*. [s.l.]: ERIC Clearinghouse, 2014, p. 34–47. <<https://tinyurl.com/y5skv5b5>>.
- O'Donnell 2012 James J. O'Donnell. *The Future Is Now, and Has Been for Years*. «The Chronicle of Higher Education», 3 settembre 2012. <<https://www.chronicle.com/article/The-Future-Is-NowHas/134070>>. Pappano 2012 Laura Pappano. *The Year of the MOOC*. «The New York Times», 4 novembre 2012, <<https://tinyurl.com/y2hvha8k>>.

- Pozzi – Conole 2014 Francesca Pozzi – Grainne Conole. *Quale futuro per i MOOC in Italia?* «Italian Journal of Educational Technology», 22 (2014) n. 3, p. 173–82. <<https://doi.org/10.17471/2499-4324/187>>.
- Progetto MOOCs Italia – CRUI 2017 Progetto MOOCs Italia – CRUI. Conferenza dei Rettori delle Università italiane 2017. *Linee guida nazionali per la predisposizione di MOOCs di qualità erogati dalle Università italiane: documento approvato dalla CRUI*, aprile 2017. Fondazione CRUI. <[https://www.cru.it/images/1-\\_LineeGuidaMOOCsItalia\\_aprile2017.pdf](https://www.cru.it/images/1-_LineeGuidaMOOCsItalia_aprile2017.pdf)>.
- Shah 2018 Dhawal Shah. *By The Numbers: MOOCs in 2017*. «Class Central's MOOC Report», 19 gennaio 2018. <<https://www.classcentral.com/report/mooc-stats-2017/>>.
- Siemens 2012 George Siemens. *Talks about MOOCs and their benefits for the teacher and the learner in the context of open education and OER. Dr Siemens also gives advice for anyone wishing to run a MOOC (Massive Open Online Course)*. <<https://www.youtube.com/watch?v=-a2cEzsMEMY>>; <<https://tinyurl.com/y2zqjnmr>>.
- Siemens 2014 George Siemens. *Origins of the cMOOC*. In: Fiona M. Hollands – Devayani Tirthali. *MOOCs: Expectations and Reality*. New York: Center for Benefit-Cost Studies of Education, Teachers College, Columbia University, 2014.
- Tammaro 2016 Anna Maria Tammaro. *Putting the users into the digital library: reflections on a MOOC teaching experience*. «DLIS», 19 luglio 2016. <<https://dlis.hypotheses.org/642>>.
- Trejo 2016 Orlando Trejo. *Getting to Know MiríadaX, Latin America's Largest MOOC Platform*. «Class Central's MOOC Report», 21 novembre 2017. <<https://www.classcentral.com/report/getting-know-miriadax/>>.
- Wetterstrom 2014 Luke Wetterstrom. *The Year After the Year of the MOOC*. «The Gate», 28 gennaio 2014. <<http://uchicagogate.com/articles/2014/1/28/years-after-mooc/>>.



## Progetti



# L'attuazione in Italia del Progetto GoogleBooks

**Andrea De Pasquale**

*Direttore della Biblioteca Nazionale Centrale di Roma (BNCR)<sup>1</sup>*

*Con la sottoscrizione dell'Accordo di Cooperazione tra il MiBACT e Google Ireland Limited (9 marzo 2010) e di quello tra il Dipartimento per la digitalizzazione della pubblica amministrazione della Presidenza del Consiglio dei Ministri e il MiBACT (2 febbraio 2012), si è dato avvio in Italia al progetto GoogleBooks. Esso si è svolto in quattro fasi dal dicembre 2012 al dicembre 2018 e ha visto la partecipazione di undici biblioteche statali sotto il coordinamento della Biblioteca Nazionale Centrale di Roma, portando alla digitalizzazione di circa 630.800 volumi. Se ne dettagliano le operazioni svolte, i vantaggi ottenuti e si effettua una valutazione dei costi/benefici in riferimento alle digitalizzazioni fatte alla Biblioteca Nazionale Centrale di Roma.*

## Le fasi del progetto

**C**on la sottoscrizione dell'Accordo di Cooperazione tra il MiBACT e Google Ireland Limited per la digitalizzazione e messa in rete del patrimonio librario delle Biblioteche Nazionali Centrali di Firenze e di Roma, avvenuta in data 9 marzo 2010 e il successivo Addendum relativo alla Biblioteca Nazionale di Napoli del 17 febbraio 2011, si è avviata la più importante e longeva collaborazione internazionale della Biblioteca Nazionale Centrale di Roma (BNCR), che ha portato all'attuazione in Italia del Progetto GoogleBooks, per il quale la Biblioteca ha coordinato, su incarico del Ministero, tutte le attività tecniche-informatiche, organizzative e amministrative (stazione appaltante per tutte le fasi con stipula di tutti i contratti per servizi di movimentazione e altre prestazioni, forniture e lavori di restauro). L'Italia si allineava così alle scelte di altri importanti Paesi, come Stati Uniti, Inghilterra, Germania, Francia, Austria, Spagna, Olanda, Repubblica Ceca, Giappone e di biblioteche nazionali di rilevanza mondiale, come la Bayerische Staatsbibliothek di Monaco di Baviera, la British Library di Londra, l'Österreichische Nationalbibliothek di Vienna, la Nationale Bibliotheek van Nederland de L'Aja, che avevano stipulato accordi con Google per la digitalizzazione del loro patrimonio.

<sup>1</sup> Ringrazio Monica Sperabene, responsabile dell'Ufficio Innovazione, Polo BVE, Sistemi informativi della Biblioteca Nazionale Centrale di Roma, per la collaborazione nella redazione di questo contributo.



Il progetto si è svolto per quasi un decennio in più fasi contraddistinte dalla partecipazione di numerosi soggetti, da lotti di quantità diversa di libri e da finanziamenti differenti.

I lavori presero avvio dopo la stipula dell'accordo di cooperazione del 2 febbraio 2012 tra il Dipartimento per la digitalizzazione della pubblica amministrazione della Presidenza del Consiglio dei Ministri e il MiBACT per la realizzazione del progetto Catalogazione e creazione metadati a supporto del Progetto Google. Sono poi seguite una nota del 24 novembre 2014, prot. 21863 della Direzione Generale Biblioteche e istituti culturali (DGBIC), con la quale si comunicava l'intenzione del Ministero di proseguire insieme ad AgID l'intervento iniziato con il citato accordo del 2012, un nuovo Accordo di cooperazione stipulato il 6 marzo 2015 tra l'Agenzia per l'Italia digitale della Presidenza del Consiglio dei Ministri e la DGBIC del MiBACT per la realizzazione del progetto Interventi per la valorizzazione e promozione del patrimonio, conseguente al progetto Catalogazione e creazione metadati a supporto del Progetto Google. Nell'ambito dei citati Accordi, venivano messi a disposizione i finanziamenti necessari, erogati da DGBIC, AgID e ARCUS spa.

Le fasi di svolgimento sono state le seguenti:

### **GoogleBooks I (dicembre 2012 – agosto 2014)**

Biblioteca Nazionale Centrale di Roma (95.249 volumi di cui 60.000 catalogati ex novo in SBN);

Biblioteca Nazionale Centrale di Firenze (87.238 volumi di cui 73.000 catalogati ex novo in SBN);

Biblioteca Nazionale di Napoli (72.004 volumi di cui 39.500 catalogati ex novo in SBN).

In questa fase ha partecipato anche l'ICCU – Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche - per la parte informatica relativa ad SBNWeb.

Totale volumi digitalizzati: circa 254.491.

Finanziamento: € 2.299.826,76 così suddivisi: MiBACT: € 300.000,00 e AgID/ARCUS € 1.999.826,76

### **GoogleBooks II (autunno 2015-2016)**

Biblioteca Nazionale Centrale di Roma (22.983 volumi),

Biblioteca Nazionale Centrale di Firenze (28.723 volumi);

Biblioteca Nazionale di Napoli (33.652 volumi);

Biblioteca Angelica (7.687 volumi);

Biblioteca universitaria Alessandrina (24.840 volumi);

Biblioteca Casanatense (30.872 volumi);

Biblioteca di Storia moderna e contemporanea (33.150 volumi);

Biblioteca universitaria di Torino (50.567 volumi);

Totale volumi digitalizzati: circa 232.474.

Finanziamento: € 281.164,38 così suddivisi AgID: € 232.050,00; ARCUS € 49.114,38.

### **GoogleBooks III (gennaio 2017-marzo 2018, con la conclusione effettiva nel mese di giugno)**

Tale fase ha previsto la catalogazione in SBN dei volumi e la digitalizzazione delle collezioni di periodici della BNCR realizzate sia tramite propri autonomi finanziamenti di bilancio sia tramite i contributi AgID e ARCUS per l'invio di n. 60.000 fascicoli di periodici della BNCR.

Biblioteca Nazionale Centrale di Roma (50.000 volumi);

Biblioteca Nazionale Centrale di Firenze (12.663 volumi);

Totale volumi digitalizzati: circa 62.663.

Finanziamento: € 140.971,02 così suddivisi AgID: € 87.243,48;

ARCUS € 53.727,54; bilancio BNCR: €. 58.191,30

### **GoogleBooks IV (gennaio 2018 – dicembre 2018)**

La fase si avviava a seguito di richiesta di finanziamenti da parte di BNCR alla DGBIC che con nota del Serv. 1 n. 19907 del 08/11/2017 autorizzava il contributo per l'attuazione del Progetto Google – Fase 4 nell'anno 2018, con l'invio di circa 85.000 volumi di tre biblioteche partner: Biblioteca statale di Cremona, Biblioteca statale Isontina di Gorizia, Biblioteca annessa al monumento nazionale di Santa Scolastica di Subiaco.

La Biblioteca Nazionale Centrale di Roma ha coordinato le attività per le seguenti biblioteche partner:

Biblioteca statale di Cremona (circa 55.000 volumi);

Biblioteca statale Isontina di Gorizia (circa 15.000 volumi);

Biblioteca statale annessa al Monumento Nazionale di S. Scolastica (circa 15.000 volumi).

Totale volumi digitalizzati: circa 85.000.

Finanziamento previsto: € 130.000,00 a carico della Direzione Generale Biblioteche e istituti culturali

Attualmente, il totale dei volumi digitalizzati nelle quattro fasi operative del progetto che si sono succedute nel corso degli anni, tra il mese di dicembre 2012 e il dicembre 2018, al quale hanno partecipato in totale 11 biblioteche statali del Ministero, è stato di circa 630.800 così suddivisi:

Biblioteca Nazionale Centrale di Roma (169.400 volumi);

Biblioteca Nazionale Centrale di Firenze (128.624 volumi);

Biblioteca Nazionale Centrale di Napoli (105.656 volumi);

Biblioteca Angelica (7.687 volumi);  
Biblioteca universitaria Alessandrina (24.840 volumi);  
Biblioteca Casanatense (30.872 volumi);  
Biblioteca di Storia moderna e contemporanea (33.150 volumi);  
Biblioteca universitaria di Torino (50.567 volumi).  
Biblioteche statale di Cremona (circa 55.000 volumi)  
Biblioteca statale Isontina di Gorizia (circa 15.000 volumi)  
Biblioteca statale annessa al Monumento nazionale di Santa Scolastica di Subiaco (circa 15.000 volumi).  
Considerato che il progetto prevedeva la digitalizzazione di almeno 500.000 unità e fino a raggiungere un numero massimo di 1.000.000 di volumi, il MiBACT sta ora valutando se proseguire con la digitalizzazione ulteriore di circa 180.000 opere, nel prossimo biennio. Le biblioteche candidate sono: la Biblioteca Nazionale Centrale di Roma (anche con il fondo ISIAO) per circa 15.000 volumi, la Biblioteca Nazionale di Napoli per 100.000 volumi, le Biblioteche Universitarie di Pavia e di Padova rispettivamente per 55.000 e 10.000 volumi.

### **Il workflow delle operazioni**

Alcune attività tecniche sono state indispensabili per l'avvio del progetto nelle singole biblioteche. Innanzitutto, condizione imprescindibile è stata la catalogazione (generalmente in SBNWeb) di tutti i volumi da inviare allo Scan Center di Google; quindi, la biblioteca interessata – eventualmente anche con il supporto del Polo – ha provveduto allo scarico in formato UNIMARC delle notizie bibliografiche complessive relative al materiale scelto, facendo attenzione ad includere solo i volumi effettivamente posseduti dalla biblioteca e, tra questi, soltanto i volumi la cui data di edizione fosse compresa nell'intervallo temporale idoneo per la scansione, come stabilito in accordo con Google nel rispetto del diritto d'autore (monografie editate fino al 1877 e periodici fino al 1947). È stato indispensabile inoltre che i volumi considerati fossero in buono stato di conservazione.

Le biblioteche partecipanti hanno quindi dovuto preparare i volumi ritenuti idonei e renderli disponibili negli appositi carrelli da inviare per la scansione, usufruendo delle tecnologie e dei software appositamente creati per la movimentazione ragionata delle opere; quindi, a valle, assicurare o affidare al coordinatore la conservazione dei file digitali restituiti da Google dopo la lavorazione.

Ciò ha avuto ovviamente un considerevole impatto sull'impegno organizzativo delle risorse umane e delle attività, oltre che sul flusso dei procedimenti amministrativi; nello specifico le voci di spesa ammesse negli stanziamenti sono servite per l'acquisto e spedizione di attrezzature informatiche e materiali di facile consumo, per l'affidamento di servizi tecnico-specialistici come le attività informatiche,

amministrative e contabili, la catalogazione e l'inventariazione, il restauro e la movimentazione libraria; infine sono state effettuate spese di missione per i sopralluoghi indispensabili all'avvio delle attività nelle biblioteche.

Tutto il processo di digitalizzazione è avvenuto presso lo Scan Center appositamente creato da Google in Italia. Google ha provveduto anche al trasporto e all'assicurazione dei volumi verso qualsiasi forma di rischio, per la fase del trasporto da e verso la biblioteca e durante la lavorazione allo Scan Center.

La BNCR ha mantenuto, nel suo ruolo di coordinamento, un costante contatto sia con la sede USA sia con i responsabili dello Scan Center, oltre che con le biblioteche partner del progetto.

Per quanto riguarda le attività informatiche connesse alla movimentazione dei libri, la Biblioteca Nazionale Centrale di Roma e la Biblioteca Nazionale Centrale di Firenze hanno reso disponibile *GoogleBooks Light*, un software prezioso per rendere possibile il processo e monitorare con assoluta certezza i volumi dalla fase della selezione, all'invio allo Scan Center, alla restituzione finale alla biblioteca.

Quindi l'Ufficio Innovazione, Polo BVE, Sistemi informativi della BNCR, un volta ricevuto il file UNIMARC complessivo dalla biblioteca, ha provveduto alla verifica dell'idoneità dei dati ai fini del loro caricamento nella banca dati *GoogleBooks Light*; in tal senso, si è proceduto innanzitutto con l'analisi delle collocazioni dei volumi, per appurare che non ci fossero condizioni di ambiguità nella descrizione delle collocazioni e nella digitazione degli inventari (questa fase ha perciò richiesto talvolta il coinvolgimento dei bibliotecari); si è verificato quindi il formato del file UNIMARC e si è proceduto alla conversione in un file XML Marc21 che è stato caricato nella banca dati *GoogleBooks Light* della biblioteca interessata, nel frattempo allestita dalla BNCR su un pc dedicato.

L'esito positivo di tutte le fasi suddette ha quindi consolidato l'idoneità della biblioteca candidata a partecipare al Progetto Google usando *GoogleBooks Light*. Solo allora il pc preparato dalla BNCR, completo di software e banca dati personalizzata, è stato inviato alla biblioteca dove un tecnico informatico ha provveduto sia alla verifica del perfetto funzionamento di tutti i processi, inclusa l'accessibilità per l'assistenza da remoto garantita da BNCR, sia all'addestramento del personale all'uso delle procedure *GoogleBooks Light*.

Sono stati anche verificati i parametri logistici necessari negli ambienti di lavoro presso la biblioteca candidata ad entrare nel progetto. A tal fine sono stati fissati con netto anticipo gli incontri con i rappresentanti di Google e dello Scan Center presso le biblioteche, per discutere, anche con la direzione della BNCR, tutte le questioni organizzative.

Per un agevole svolgimento del lavoro sono occorsi infatti spazi per la movimentazione dei volumi, per lo stoccaggio e il riempimento dei carrelli, dovendo assicurare l'adeguata distanza tra il locale destinato alla preparazione dei carrelli e i magazzini

librari, garantendo che il percorso fosse libero da ostacoli sia per il trasporto all'interno della biblioteca sia da questa al camion di Google, che stazionava al piano stradale per il carico della spedizione e la restituzione nelle date pattuite. Una volta completato il carico, il camion è stato chiuso in sicurezza anche con l'applicazione di sigilli di garanzia, aperti solo una volta che il camion era giunto a destinazione.

Google ha organizzato il piano delle spedizioni con estrema cura, in modo che lo Scan Center fosse costantemente in attività, garantendo efficienza e puntualità, con un alto livello di controllo dei processi, creando un circuito costante. Dato il coinvolgimento di un elevato numero di volumi tutti di pregio, inviati dalle diverse biblioteche, il presupposto strategico basilare per ogni fase di lavorazione, è stato quello di garantire corrispondenze univoche e rigorose, che rendessero praticamente impossibile qualsiasi genere di errore, o consentissero di ravvisarlo immediatamente.

A tal fine, su ambedue i lati di ogni carrello, la biblioteca ha affisso un foglio contenente l'identificativo del carrello stesso, composto dalla sigla della biblioteca e da un numero progressivo composto da tre cifre; il foglio ha identificato il carrello per tutta la durata del progetto, sia per Google che per la biblioteca ed è stato quindi fissato in modo tale da non poter essere staccato.

Al pari, tra la coperta e il frontespizio di ogni volume, via via che veniva caricato il carrello, è stato inserito un modulo con un codice a barre costituito dal codice della biblioteca e dal numero di inventario del volume stesso; il modulo è stato lasciato appositamente sporgente rispetto al taglio del volume, tanto quanto necessario al riconoscimento ottico immediato, che è stato effettuato con un apposito lettore, collegato al pc della banca dati *GoogleBooks Light*.

Salvo che per le Biblioteche di Napoli e Firenze che hanno provveduto autonomamente a tutte le procedure, la conservazione delle risorse digitalizzate è effettuata dall'Ufficio Innovazione, Polo BVE, Sistemi informativi della BNCR per un totale di circa 5,8 TB; mentre è stato possibile solo per la prima parte del Progetto Google inviare le copie master ai Magazzini digitali per la conservazione a lungo termine, per un totale di circa 100.000 documenti tra l'autunno 2014 e la primavera 2015. Sono previsti prossimi invii per il successivo materiale.

### Vantaggi

Il progetto ha avuto conseguenze positive importanti sugli assi portanti della *mission* istituzionale delle biblioteche e della DGBIC del MiBACT: la conservazione e la tutela, ma insieme anche l'accessibilità, la valorizzazione, la divulgazione delle raccolte. Il progetto infatti ha costituito una grande opportunità per la salvaguardia dei testi rari e di pregio, patrimonio inestimabile delle biblioteche, coniugando le esigenze di tutela con una spinta cospicua a favore della fruizione, grazie al digitale<sup>2</sup>.

<sup>2</sup> Marina Battaglini, *Digitalizzazione e tutela: la Biblioteca Centrale Nazionale di Roma e il Progetto Google*, in *I Beni Bibliografici nelle strategie dei fondi europei*. Siracusa, ISISC, 3-4 dicembre 2015.

Possibilità di ricerca delle parole nei testi tramite OCR, accesso libero al materiale, ricostruzione virtuale di edizioni incomplete, sono solo alcuni dei vantaggi anche economici che ha comportato il progetto, considerando che la lettura delle copie digitali è l'alternativa alla consultazione di quelle originali, e consente di proteggere il materiale librario dai danni derivanti dall'uso nel tempo. Dunque grazie all'alta qualità dei processi standardizzati e ai corposi numeri in termini di opere riprodotte, il progetto GoogleBooks ha permesso di attuare una politica per l'accesso, la conoscenza, la conservazione e la tutela del patrimonio tanto massiva quanto incisiva, portando vantaggi quantificabili anche in termini economici, a medio e lungo termine.

I file digitali dei volumi possono essere liberamente e gratuitamente visualizzabili in full text e scaricabili poiché si tratta di opere di pubblico dominio che non rientrano nelle limitazioni imposte dalla legge sul diritto d'autore, e sono reperibili oltre che sul sito di GoogleBooks, su alcuni siti web delle biblioteche e potranno essere fruiti in futuro attraverso portali nazionali o internazionali. Occorre anche tener conto che, non essendo stata concessa alcuna esclusiva a Google, ogni biblioteca è stata libera nel frattempo di stipulare contratti di digitalizzazione con terze parti, senza alcuna limitazione, dunque anche dando seguito alle attività con progetti successivi o paralleli.

### **Costi e benefici per la Biblioteca Nazionale Centrale di Roma**

Precedentemente all'avvio del progetto GoogleBooks, la BNCR aveva digitalizzato quanto di seguito indicato:

- manoscritti (inclusi manoscritti musicali e manoscritti miniati) e documenti autografi: circa 6.770, per 77.200 pagine
- volumi rari: 558 per un totale di circa 100.000 pagine
- periodici storici: 441 per circa 200.000 pagine
- quotidiani e settimanali: 104.900 fascicoli per 1.325.000
- fotografie e cartoline: 2.318 per 4.979 immagini digitali
- partiture musicali: 531 per 120.000 pagine
- dischi in vinile: 5.000 per 10.000 mp3.

Una parte considerevole dei file non era corredata da metadati, che furono forniti in seguito, nell'ambito del progetto europeo Athena Plus. Inoltre la BNCR non disponeva fino al 2015 di scanner planetari per la ripresa digitale del materiale raro e di pregio, indispensabile per la gestione interna del servizio di digitalizzazione.

*Atti del Convegno*, a cura di A. Campagnolo, L. Catalano, R. C. Giordano, G. Lo Piccolo, Palermo: Assessorato dei Beni Culturali e dell'Identità siciliana. Dipartimento B.C. e I.S., 2015, p. 140-146.

Ora con il Progetto Google la BNCR ha ottenuto la digitalizzazione di 169.400 volumi tra monografie e periodici, per un totale di 64.160.220 pagine che sono attualmente accessibili, sia in sede che da remoto.

I costi sostenuti, pur approssimativi, per tali digitalizzazioni sono stati i seguenti: Anno 2012: €. 48.862,22; Anno 2013: €. 695.554,58; Anno 2014: €. 228.376,71; Anno 2015: €. 265.221,70; Anno 2016: €. 48.360,70; Anno 2017: €. 56.729,94; Anno 2018: €. 70.296,84; per un totale di €. 1.413.392,69.

Il costo medio per pagina digitalizzata e metadatata è stato pari a: €. 1.413.392,69 / 64.160.220 pagine = €. 0,02, a fronte di un costo medio di mercato per le suddette operazioni di € 0,40.

Circa 90.000 monografie della Biblioteca Nazionale digitalizzate nel Progetto Google sono accessibili sia in sede dal Sistema integrato di servizi Ermes sia da remoto tramite l'OPAC SBNWeb, in attesa di riunirle nella nuova Teca digitale, che mira ad essere il portale deputato alla consultazione di tutte le risorse digitalizzate negli anni passati, presenti e prossimi della Biblioteca<sup>3</sup>.

Il progetto ha anche dato lo spunto per avviare un'importante operazione culturale, da sempre auspicata, che riprende l'antico ruolo della Biblioteca di sede dell'Emeroteca nazionale<sup>4</sup>.

Nella seconda fase del progetto GoogleBooks sono stati digitalizzati circa 10.000 volumi di periodici per circa 1.000 testate e nella fase terza, iniziata nel gennaio 2017, circa 60.000 volumi per un totale di 6.000 testate, complessivamente per un ammontare di 20 milioni di pagine.

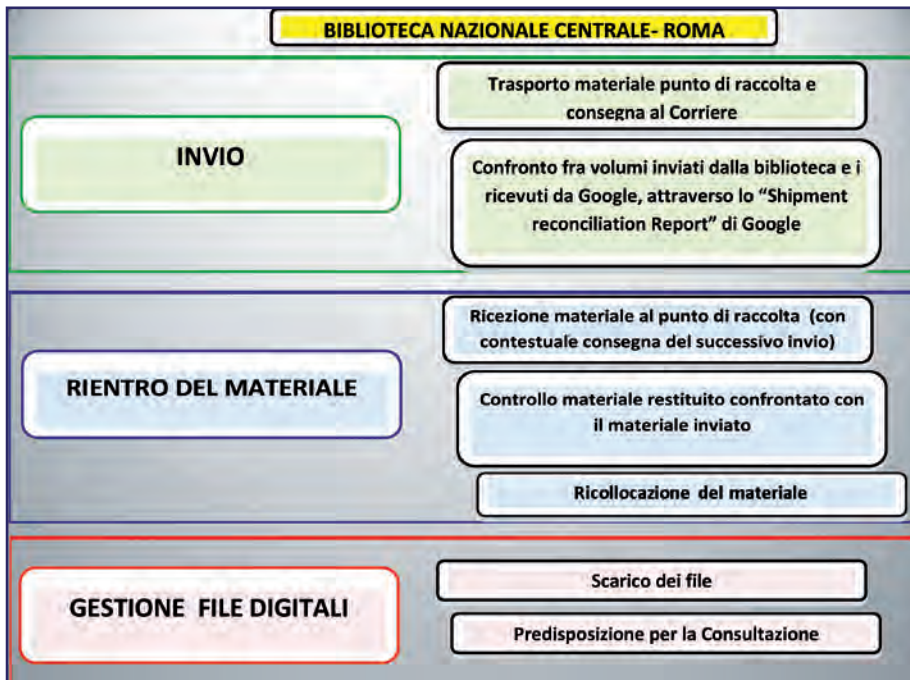
Contestualmente la Biblioteca, dotata di apposite strumentazioni, ha avviato la digitalizzazione delle sue oltre 50.000 bobine in microfilm di testate periodiche (1.025 titoli italiani e stranieri), di cui circa 6.000 sono state già digitalizzate.

In più il 18 luglio 2017 è stata sottoscritta con il Senato della Repubblica una convenzione di durata quinquennale, che ha come obiettivo la realizzazione di un'Emeroteca nazionale contenente una collezione di periodici italiani digitalizzati mediante la ricostruzione completa di annate periodiche possedute dalle due biblioteche. Il Senato ha messo a disposizione 60 testate storiche italiane già digitalizzate e disponibili sul loro portale GiSID oltre alla collezione completa dell'*Avanti*, 42 testate italiane e straniere microfilmate e circa 800 testate italiane e straniere in formato cartaceo.

È stato quindi realizzato un portale detto Emeroteca Digitale, ospitato all'interno della Teca digitale della BNCR, che contiene attualmente 2.117 testate periodiche provenienti non solo dal Progetto Google ma anche dal trascorso progetto europeo Athena Plus, a cui la Biblioteca aveva partecipato, e dalla

<sup>3</sup> <http://digitale.bnc.roma.sbn.it/tecadigitale/>.

<sup>4</sup> Andrea De Pasquale, *Per un'emeroteca nazionale digitale*, in *Bibliographic journals between past and future. International Conference Proceedings (Bologna, Biblioteca Universitaria, 22-23 February 2018)*, a cura di R. Cesana e F. Sabba, «Bibliothecae.it», 7/2 (2018), p. 348-370, <<https://bibliothecae.unibo.it/article/view/8951>>.





260 nominal books a carrello (1 nominal book= 330 pagine)  
Cm 114 x 71 prof. x 177 h  
Peso 250 kg. ca pieno



**T-Cart Single Sided**  
con dimensioni cm 88 (lunghezza) x cm 38 (profondità) x cm 137 (h), 3 piani cm 30,5 profondi, distanza fra i piani cm. 35, peso pieno ca. kg 102. I metri lineari disponibili per i libri sono m. 2,28 a carrello.  
Calcolando come misura standard il *nominal book* (330 p., spessore cm 2,54, peso Kg. 0,695), su questo carrello si possono caricare ca. 90 volumi.



**90 VOLUMI**

 B1000153320	ColiBa 20 00002 1985 v.1 1985, v. 1
Google Book Search   Biblioteca Isontina	
Autore Titolo: <i>British books in print ... 1874-1987 ...</i> Data: 21/10/2017 09:51 © Google Books 13225 (T000009232) (Periodico) Barcode: B1000153320 Data reimagio: 19/01/2018	
<div style="border: 1px solid red; padding: 2px; display: inline-block;"><b>Da inserire nel libro</b></div>	
 ColiBa 20 00002 1985 v.1	
3 di 6	

ColiBa 20 00002 1985 v.1 1985, v. 1

Google Book Search   Biblioteca Isontina
Autore Titolo: <i>British books in print ... 1874-1987 ...</i> Data: 21/10/2017 09:51 © Google Books 13225 (T000009232) (Periodico) Barcode: B1000153320 Data reimagio: 19/01/2018
<div style="border: 1px solid red; padding: 2px; display: inline-block;"><b>Segnaposto nello scaffale</b></div>
Motivo della negativi
<input type="checkbox"/> Mancante <input type="checkbox"/> Non idoneo per Google <input type="checkbox"/> Non disponibili al momento <input type="checkbox"/> In cattivo stato
3 di 6

digitalizzazione dei microfilm della Biblioteca, sviluppando una collezione che conta attualmente oltre 17 milioni di immagini.

All'interno dell'Emeroteca digitale è possibile effettuare sia una ricerca semplice che avanzata o scorrere semplicemente la lista alfabetica dei titoli presenti e delle annate disponibili. Attualmente la Teca non supporta la tecnologia OCR (la funzionalità è in fase di progettazione), ma per i circa 60.000 volumi digitalizzati con il



progetto GoogleBooks è possibile la ricerca a testo intero sul portale del progetto. L'idea è quella di potenziare il portale nell'immediato futuro con altre risorse di periodici digitali provenienti da altre campagne di digitalizzazione italiane, al fine di ricostruire una bibliografia completa che integri le collezioni stesse, colmando le lacune e restituendo le testate nella loro completezza.

*With the signing of the Cooperation Agreement between the MiBACT and Google Ireland Limited (9 March 2010) and the one between the Department for the digitalization of the public administration of the Presidency of the Council of Ministers and the MiBACT (2 February 2012), it was given launch in Italy to the GoogleBooks project. It took place in four phases from December 2012 to December 2018 and saw the participation of eleven state libraries under the coordination of the Central National Library of Rome, leading to the digitization of approximately 630.800 volumes. The operations performed are detailed, the advantages obtained and an evaluation of the costs/benefits is carried out with reference to the digitalizations made at the Central National Library of Rome.*

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

# L'emeroteca digitale dei giornali locali del Piemonte

**Dimitri Brunetti**

*Regione Piemonte*

*I giornali locali sono un patrimonio prezioso per la storia del territorio e costituiscono un bene di grande interesse dal punto di vista sociale, economico, culturale e storiografico. Purtroppo trovare notizie su di un giornale di carta non è affatto semplice: bisogna avere a disposizione le testate e i fascicoli e bisogna cercare con grande attenzione avendo fortuna.*

*La Regione Piemonte offre a tutti la possibilità di consultare i giornali locali già digitalizzati pubblicati in Piemonte dal 1846 ad oggi. Il servizio è libero e gratuito e non necessita di registrazione.*

*Il portale [www.giornalidelpiemonte.it](http://www.giornalidelpiemonte.it) rende oggi disponibili oltre tre milioni di pagine, ma è in continua crescita grazie agli accordi con gli editori e al lavoro delle biblioteche civiche e di altri soggetti culturali.*

I giornali locali sono un patrimonio prezioso per la storia del territorio e costituiscono un bene di grande interesse dal punto di vista sociale, economico, culturale e storiografico. Le informazioni quotidiane sono un elemento di straordinaria importanza per conoscere cosa succede nel nostro presente, ma sapere anche cosa è accaduto in passato è altrettanto significativo per ricostruire i frammenti della nostra identità.

I giornali di prossimità sono l'unica fonte in grado di raccontare giorno per giorno le vicende dei nostri territori, delle comunità e delle persone. In molti casi sono i soli testimoni di storie e situazioni che nessun altro documento riesce a testimoniare e trasmettere.

Il Piemonte vanta una storia eccezionale nell'editoria locale, sia per numero di testate, sia per continuità nel tempo. Purtroppo si tratta di documenti fragilissimi e, in più, trovare notizie su di un giornale di carta non è affatto semplice: bisogna avere a disposizione le testate e i fascicoli e bisogna cercare con grande attenzione avendo fortuna.

Per valorizzare l'informazione giornalistica locale e per facilitare la ricerca su di essa, la Regione ha commissionato al CSI-Piemonte la realizzazione del portale «Giornali del Piemonte». Il progetto è stato definito dalla Giunta Regionale e dal Consiglio Regionale, con la collaborazione della Federazione italiana piccoli editori

giornali (Fipeg), degli editori che hanno condiviso questa idea, degli istituti culturali, dei sistemi bibliotecari e delle biblioteche pubbliche.

Il portale offre ai cittadini, ai ricercatori, agli insegnanti, agli studenti e ai giornalisti la possibilità di consultare online i giornali locali già digitalizzati, pubblicati in Piemonte dal 1846 ad oggi. Il servizio è gratuito e non necessita di registrazione. Però, in accordo con gli editori, non viene consentita la lettura delle ultime due edizioni (quella in edicola e quella immediatamente precedente), che comunque progressivamente entrano a far parte del patrimonio accessibile.

Il progetto sui giornali del Piemonte probabilmente è stato il primo e il più ampio intervento pubblico a favore dell'accesso libero e gratuito alle pagine dei giornali locali. Ancora oggi il portale [www.giornalidelpiemonte.it](http://www.giornalidelpiemonte.it) è al centro di un'azione sostenuta in modo significativo dalla Regione, che da anni si impegna per garantirne lo sviluppo in termini di pagine disponibili e servizi offerti<sup>1</sup>.

### 1. La digitalizzazione dei giornali

Da una ventina d'anni la Regione, i sistemi bibliotecari, le biblioteche pubbliche e gli istituti culturali del territorio investono numerose energie nella digitalizzazione dei giornali, inizialmente come elemento qualificante dello sviluppo della Biblioteca Digitale Piemontese e ora in funzione del nuovo portale, con lo scopo di offrire alla consultazione sul web una parte consistente del patrimonio culturale conservato. Nei primi tempi sono state riprodotte alcune testate di particolare rilevanza, come l'«Avanti! Quotidiano del Partito socialista italiano» in varie edizioni dal 1896 al 1993, «Il Popolo Nuovo» dal 1948 al 1960, la «Rivoluzione Liberale» dal 1922 al 1925, una trentina di giornali di fabbrica pubblicati dal 1949 al 1991 e l'edizione piemontese de «L'Unità» dal 28 aprile 1945 al 31 luglio 1957 (quando viene soppressa per esigenze di bilancio) che costituisce una delle quattro edizioni del quotidiano nazionale insieme a quelle di Milano, Genova e Roma. Si provvede inoltre alla riproduzione di alcune riviste quali «Notizie Olivetti» dal 1952 al 1995, «Comunità» dal 1946 al 1992, la «Riforma sociale», confluita poi in «Rivista di storia economica» dal 1894 al 1943, nonché numerosi periodici del cinema muto pubblicati a Torino dal 1910 al 1930.

Nei primi anni Duemila viene avviato dal comune di Fossano e dalla Regione un vasto intervento per la digitalizzazione di tutti i giornali della provincia di Cuneo, che porta alla riproduzione (integrale o limitata al secondo dopoguerra) di 67 testate e 418.000 pagine. Intanto altre biblioteche realizzano analoghi interventi, ad esempio nella zona del comune di Chieri e nell'alessandrino.

<sup>1</sup> Il presente contributo prosegue e aggiorna l'articolo già pubblicato nella sezione Progetti di questa stessa rivista dell'ICCU: Dimitri Brunetti, *Testate giornalistiche d'informazione locale: interventi di digitalizzazione e valorizzazione in Piemonte*, «DigItalia. Rivista del digitale nei beni culturali», 9 (2014), n. 1, p. 89-101.

Il 25 giugno 2008 viene emanata la legge regionale n. 18 “Interventi a sostegno dell’editoria piemontese e dell’informazione locale” in favore dell’editoria locale e nel novembre dello stesso anno, in occasione della sottoscrizione del protocollo d’intesa e di collaborazione fra la Regione e la Fipeg, viene presentato il portale *MenteLocale* che si propone come vetrina della stampa locale piemontese. *MenteLocale* nasce dopo una sperimentazione iniziata nel 2006, quando viene realizzata una piattaforma tecnologica per poter consultare le fonti giornalistiche del territorio in un’unica interfaccia in modalità digitale e per realizzare rassegne stampa. Ogni giorno nella piattaforma entravano oltre mille documenti corrispondenti ad altrettante pagine di giornale, circa cinquanta comunicati stampa e molte immagini fotografiche. Al 2014 aderiscono a *MenteLocale* 76 testate distribuite su tutto il territorio regionale con periodicità trisettimanale, bisettimanale, settimanale, quindicinale o mensile.

Il 6 giugno 2012 viene presentato in anteprima al Presidente della Repubblica Giorgio Napolitano il portale web con cui si accede all’intera raccolta digitalizzata del quotidiano «La Stampa» dalla sua fondazione del 9 febbraio 1867 come «Gazzetta Piemontese» al 2005. Si è trattato del primo archivio di un quotidiano nazionale ad essere completamente digitalizzato e ad aver messo a disposizione dei lettori in maniera libera e gratuita circa 1.800.000 pagine che contengono i 12 milioni di articoli pubblicati in 138 anni: all’indirizzo [www.archiviola stampa.it](http://www.archiviola stampa.it) sono consultabili le pagine di tutte le 47.243 edizioni de «La Stampa» e le 18.314 edizioni di «Stampa Sera», incluse le edizioni locali e gli inserti TorinoSette, TuttoLibri, TuttoScienze, MondoScuola ecc. Attualmente il portale è in fase di aggiornamento e si prevede di ampliarne l’offerta con l’aggiunta delle annate più recenti.

Nel 2013 il Consiglio regionale del Piemonte organizza il convegno “Tra la carta e il web. Le nuove sfide dell’informazione locale” che consolida la collaborazione con gli editori e porta all’ampliamento delle testate aderenti a *MenteLocale*.

Nel 2014 con fondi comunitari prende avvio un nuovo progetto che prevede di assegnare contributi alle biblioteche comunali centro rete dei sistemi bibliotecari per avviare o completare la digitalizzazione di alcune testate giornalistiche locali scelte per rilevanza e sulla base della rappresentatività nelle provincie piemontesi. Il bando assegna un contributo fino a un massimo dell’80%. Al bando rispondono i sistemi bibliotecari di Biella, Casale Monferrato, Fossano (in associazione con i sistemi Alba, Cuneo e Mondovì), Novara, Novi Ligure, Pinerolo, Tortona e Verbania, che in accordo con gli editori di riferimento propongono undici progetti di digitalizzazione di altrettanti giornali locali. Nel corso degli anni successivi, alle prime testate se ne aggiungeranno altre, grazie ai risparmi e al reperimento di ulteriori risorse, e il progetto porterà alla riproduzione digitale e alla metadattazione di oltre novecentomila pagine di giornali in attività distribuiti su quasi tutto il territorio regionale, nonché al riconoscimento dei testi mediante OCR grazie all’applicazione delle specifiche *Linee guida* di intervento definite dal CSI-Piemonte.

Il 9 maggio 2014, in occasione del Salone internazionale del libro di Torino, viene organizzato l'incontro dal titolo "Testate giornalistiche d'informazione locale piemontese. Progetti di digitalizzazione" cui partecipano l'assessore regionale, il presidente dell'Ordine dei giornalisti del Piemonte, il presidente della Fipeg e alcuni editori. Nell'occasione viene presentato il nuovo bando di finanziamento inquadrando nell'ampio contesto di azioni per la valorizzazione dei giornali locali.

Da settembre 2015 ad aprile 2016 l'Ordine dei giornalisti del Piemonte e la Fipeg organizzano otto incontri di formazione riservati ai giornalisti dal titolo "Carte, digitale e territorio. Informatizzazione locale e comunicazione pubblica a servizio del cittadino-lettore-navigatore" che si svolgono in altrettanti luoghi (Alessandria, Asti, Baveno, Biella, Ciriè, Fossano, Novara e Vercelli) e registrano una vasta partecipazione. Gli incontri, pur non interamente dedicati a raccontare il lavoro che si stava facendo sulla digitalizzazione dei giornali, hanno avuto il merito di coinvolgere i giornalisti che, dal canto loro, hanno mostrato il proprio apprezzamento per quanto realizzato con numerosi articoli pubblicati su diverse testate e distribuiti in un lungo periodo.

## 2. [www.giornalidelpiemonte.it](http://www.giornalidelpiemonte.it)

Mercoledì 2 marzo 2016 l'assessore regionale alla cultura Antonella Parigi, il presidente dell'Ordine dei giornalisti del Piemonte Alberto Sinigaglia e il presidente nazionale della Fipeg Pietro Policante hanno presentato, nella sala stampa della sede di piazza Castello della presidenza della Regione, il nuovo portale *I Giornali del Piemonte* che permette l'accesso libero e gratuito alle raccolte digitalizzate dei giornali locali piemontesi a partire dal 1846 fino ad oggi. Nella stessa giornata viene pubblicato il portale all'indirizzo [www.giornalidelpiemonte.it](http://www.giornalidelpiemonte.it).

La prima versione offriva al consultatore 1.359.697 pagine di giornale riferite a 105 testate diverse, le funzionalità di ricerca semplice e avanzata, la possibilità di sfogliare la singola edizione e di scaricarne il contenuto in versione immagine o testo. Da quel momento al portale viene riservata una grande attenzione, così che oltre all'incremento numerico delle pagine e delle testate, già nel corso del primo anno e di quello successivo il CSI, su incarico della Regione, provvede ad ampliarne le funzioni e ad aggiungere servizi all'utente.

Sul finire del 2017 il sistema rileva alcune criticità nelle ricerche, soprattutto in quelle avviate sulla base di termini generici o di larghissimo utilizzo. Il problema viene identificato nell'alto numero delle risorse digitali disponibili: già nella tarda primavera del 2016 le pagine disponibili sono diventate oltre un milione e settecentomila, a fine anno vengono superati i due milioni di pagine, il 5 giugno 2017 sono 2.636.245, il 29 maggio 2018 3.062.557, il 27 giugno dello stesso anno sono 3.115.604. Si decide così di intervenire sul portale e nel settembre del 2018 viene pubblicata la sua seconda versione che amplia il servizio con nuove funzio-



Figura 1. L'home page della prima versione del portale I Giornali del Piemonte (2016-2018)

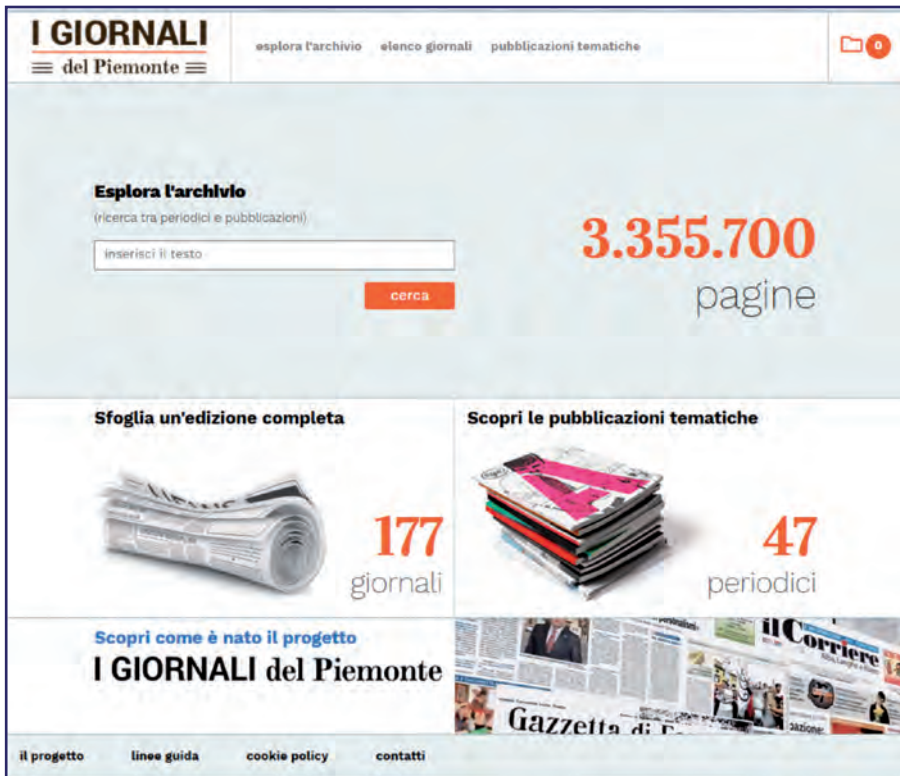


Figura 2. L'home page dell'attuale versione del portale I Giornali del Piemonte

nalità di ricerca e propone una comunicazione più efficace anche grazie all'interfaccia rinnovata.

La nuova versione del sistema si avvale di un motore di ricerca semantico full text e propone una completa revisione grafica del front-end progettato con un approccio *mobile first*. Il *layout* è completamente *responsive* in relazione al nuovo motore di ricerca semantico, per migliorare l'operatività da parte dell'utente e garantirne la fruizione ottimale sui diversi sistemi *device* e in particolare sui dispositivi *mobile*.

Il portale *I Giornali del Piemonte* espone ad ogni accesso il numero aggiornato delle pagine disponibili e delle testate dei giornali. Propone una modalità di ricerca semplice sull'intera base dati per parola o insieme di parole, ma anche una modalità avanzata di esplorazione dell'archivio che permette all'utente di selezionare i termini (tutte le parole, qualunque parola o frase esatta), di limitare l'interrogazione alle prime pagine o di ampliarla all'intero giornale, di scegliere l'arco temporale di interesse, infine di selezionare la provincia o un insieme territoriale in cui hanno sede gli editori delle risorse sia sulla base di un elenco, sia di una mappa dove sono geolocalizzate le testate. Gli elenchi proposti dal sistema possono essere organizzati per rilevanza, testata, provincia o data, crescente o decrescente. Una volta identificato il giornale che interessa, il sistema ne propone la pagina con il testo ricercato, con la possibilità di ingrandire o rimpicciolire la schermata, di stampare, di sfogliare l'intera edizione, di scaricare il testo o di scaricare il Pdf. Le pagine e gli articoli selezionati formano un fascicolo che l'utente può salvare alla fine della sessione e che include l'elenco dettagliato delle testate, edizioni e pagine scaricate e riprodotte nel medesimo Pdf di fascicolo.

Lo sviluppo più interessante della nuova piattaforma è l'ampliamento dell'offerta dai giornali locali "da edicola" ai periodici di altro genere o con periodicità più lunga. A fianco della sezione «Giornali» è stata inserita la sezione «Periodici» che rende disponibili altre testate riferite a temi e ambiti professionali specifici. La sezione Periodici permette di esporre i materiali già presenti sulla Teca Digitale Piemontese che non era stato possibile far confluire nel nuovo portale dedicato ai giornali, ma che andavano pubblicati per assicurarne la consultazione. Inoltre la predisposizione di questa sezione rende concreta la possibilità di raccogliere altri materiali o di avviare nuove campagne di digitalizzazione riservate a periodici di ambito economico, culturale, sociale ecc. I canali tematici attivi sono «Pubblica amministrazione», «Cinema» e «Lavoro», mentre altri sono in fase di predisposizione.

Infine, l'home page propone anche un testo di presentazione del progetto e il video promozionale del marzo 2016. Una sezione del portale è dedicata a coloro che gestiscono i progetti di digitalizzazione mettendo a disposizione le *Linee guida per la digitalizzazione dei periodici piemontesi*, definite dal CSI-Piemonte, che descrivono nel dettaglio le modalità di digitalizzazione e i conseguenti flussi di lavorazione.



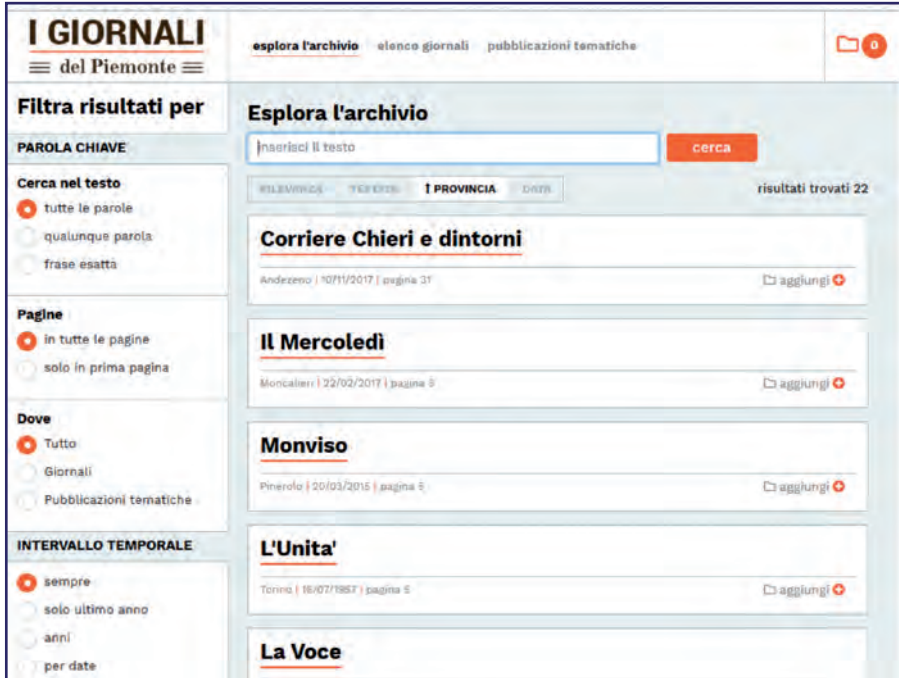


Figura 3. La pagina di ricerca



Figura 4. La visualizzazione di una pagina e il box per scaricarne il testo

Oggi, 11 luglio 2019, il portale offre al pubblico 224 testate (177 giornali e 47 periodici) per un totale di 207.302 edizioni e 3.355.700 pagine.

	<i>testate</i>	<i>edizioni</i>	<i>% testate</i>	<i>% edizioni</i>
Alessandria	36	39.577	16,1%	19,1%
Asti	3	1.921	1,3%	0,9%
Biella	11	28.822	4,9%	13,9%
Cuneo	53	79.455	23,7%	38,3%
Novara	4	9.066	1,8%	4,4%
Torino	95	28.108	42,4%	13,6%
Verbania	16	15.931	7,1%	7,7%
Vercelli	6	4.422	2,7%	2,1%
<b>totale</b>	<b>224</b>	<b>207.302</b>		

Figura 5. Il contenuto complessivo del Portale

	<i>testate</i>	<i>edizioni</i>	<i>% testate</i>	<i>% edizioni</i>
Cinema	20	394	42,6%	26,2%
Lavoro	26	770	55,3%	51,2%
Pubblica amministrazione	1	340	2,1%	22,6%
<b>totale</b>	<b>47</b>	<b>1504</b>		

Figura 6. Il contenuto dei Canali tematici

### 3. L'alimentazione continua

Inizialmente l'offerta del portale *I Giornali del Piemonte* è cresciuta grazie alla migrazione delle testate digitalizzate negli anni precedenti alla sua realizzazione e ancora residenti nella Teca Digitale Piemontese ormai dismessa. Già alla fine del 2017 in Teca non rimaneva nessun giornale, e nel corso dell'anno successivo sono stati trasferiti anche i periodici che sono andati ad alimentare i canali tematici pubblicati a partire da settembre del 2018. Intanto fra il 2016 e l'inizio del 2018 sono state consegnate le digitalizzazioni realizzate grazie al bando regionale del 2014. Queste due condizioni hanno permesso il raggiungimento di traguardi numerici importanti rispetto alle pagine disponibili.

Nell'estate del 2017 il portale del Consiglio Regionale *MenteLocale* è stato chiuso

e gli archivi in esso contenuti sono stati riversati in [www.giornalidelpiemonte.it](http://www.giornalidelpiemonte.it), che da quel momento si avvantaggia anche dello scarico automatico garantito dagli editori delle nuove edizioni prodotte, con un significativo incremento annuale di pagine.

Nel corso del 2018 è stato inserito nel protocollo d'intesa triennale con la Conferenza Episcopale Piemontese l'obiettivo di procedere alla digitalizzazione dei giornali cattolici. Allo stato attuale, delle diciotto testate piemontesi aderenti alla Federazione Italiana Settimanali Cattolici (Fisc), otto sono state interamente digitalizzate, due sono in lavorazione e otto in programma.

Da qualche anno, poi, gli accordi di finanziamento delle attività dei sistemi bibliotecari e dei poli d'area bibliotecari del Piemonte prevedono contributi vantaggiosi per la digitalizzazione dei giornali da loro conservati, pur sempre nel rispetto degli standard tecnici definiti dal progetto complessivo. In questo modo si è potuto intervenire su molte testate minori, sulle raccolte di giornali cessati, oppure completare interventi realizzati in passato, ma non conclusi o mancanti del riconoscimento testuale.

Non vanno dimenticati i progetti speciali, su istanza regionale o dei comuni, delle Università o di altri soggetti, che stanno portando alla riproduzione digitale di giornali e periodici di particolare interesse.

Infine, talvolta capita che qualche editore si proponga spontaneamente e conferisca le proprie raccolte digitali per favorire un migliore accesso a quanto ha pubblicato.

#### 4. Le linee guida per la digitalizzazione

Con l'avvio del bando del 2014 per la digitalizzazione dei giornali piemontesi, riservato alle biblioteche centro rete dei sistemi bibliotecari, il CSI-Piemonte ha definito specifiche linee guida derivandole dall'analogo documento di carattere più generale indirizzato agli interventi di riproduzione digitale dei beni documentali, la cui prima versione risale alla fine degli anni Novanta. La stesura iniziale delle *Linee guida per la digitalizzazione dei periodici locali piemontesi*, del gennaio 2015, è stata successivamente rivista e consolidata nel novembre 2017 alla luce delle prime applicazioni. Oggi tale documento definisce le regole che devono essere rispettate da tutti i soggetti, pubblici e privati, che intendono produrre oggetti digitali compatibili con il portale *I Giornali del Piemonte*.

Il documento propone istruzioni per le molteplici forme in cui il periodico potrebbe essere conservato o disponibile: in copia originale su carta, già riprodotto su microfilm, in formato immagine, in formato Pdf di tipo immagine o in formato Pdf con testo ricercabile.

Se il giornale è su carta o su microfilm si deve procedere con la creazione per ciascuna pagina di un file master, finalizzato alla conservazione a lungo termine.

Riguardo all'utilizzo dei microfilm occorre verificare la qualità della pellicola e tenere presente che in corso di confezionamento delle bobine potrebbero essere stati stampati singoli scatti di cattiva qualità, con sovrapposizioni, con parti scure o altri elementi che rendono illeggibile tutto il testo o una sua parte richiedendo, quindi, di disporre di una copia cartacea del giornale per digitalizzare nuovamente la pagina. In fase di scansione le pagine devono essere rifilate in maniera da evitare l'acquisizione di zone bianche estese al di fuori della pagina, pur garantendo di contenere per intero la pagina stessa (non la sola area di testo) con un margine tutt'intorno. Il file master deve avere una risoluzione di almeno 300 dpi, una scala di colore di 8 bit per pixel per ottenere 256 tonalità di grigio; le immagini devono essere salvate in formato Tiff non compresso. Durante la digitalizzazione viene richiesto di compilare i tag contenuti nell'*header* del file con i metadati tecnici (sezione IMG dei MAG). I file e le cartelle del *file system* devono essere organizzati secondo uno specifico modello definito per la corretta alimentazione del portale.

Dalla copia master di ogni pagina viene generato il formato Pdf con testo ricercabile per la consultazione online, accompagnato da un file Txt con il contenuto testuale integrale delle notizie per la ricerca, generato tramite un procedimento di OCR con un livello minimo di precisione del 97%.

Se l'editore, o altro soggetto, mette a disposizione il solo formato Pdf per ogni pagina di ogni singola edizione non si avrà a disposizione il file per la conservazione. In questo caso le prescrizioni delle *Linee guida* vengono seguite per la nomenclatura dei file Pdf e delle cartelle (limitandone il contenuto a 50mila file), nonché per la creazione del file testuale. Il documento regionale fornisce anche alcuni esempi di immagini nei diversi formati.

Dopo aver realizzato le prime fasi di intervento, l'azienda o il soggetto che ha svolto la digitalizzazione del periodico consegna al CSI, per ciascuna testata, un campione significativo del prodotto che viene verificato e validato. In un secondo tempo, dopo aver eventualmente provveduto alla correzione degli elementi non conformi, avviene la consegna finale e completa della testata su supporto informatico a cura dell'ente o dell'azienda e il successivo caricamento a cura del CSI sulla piattaforma Archivio storico periodici piemontesi (ASPP) rendendo consultabile le edizioni sul portale *I Giornali del Piemonte*.

## 5. L'emeroteca digitale

I giornali e i periodici sono sempre stati identificati come una fonte e una risorsa culturale imprescindibile per la conoscenza del territorio, per la storia e per la ricostruzione delle vicende minute che ci coinvolgono ogni giorno. Per questo motivo l'obiettivo che ci si è dati è quello di digitalizzare e rendere accessibili gratuitamente tutti i giornali piemontesi e, per quanto possibile, salvaguardare gli ori-

ginali cartacei sottratti alla consultazione<sup>2</sup>. Fino ad oggi è stato fatto molto: è stato digitalizzato il quotidiano «La Stampa» insieme ad altre 177 testate, quasi tutte dal primo all'ultimo numero, rendendo complessivamente già disponibili circa 5.150.000 pagine. Numerosi altri interventi sono in corso e altri ancora programmati<sup>3</sup>.

La sfida più complessa, ma necessaria è quella di digitalizzare «La Gazzetta del Popolo», forse il primo quotidiano d'Italia, stampato a Torino dal 16 giugno 1848 al 31 dicembre 1983. Centotrentasei anni difficili da consultare a causa delle poche collezioni esistenti, peraltro non complete e in genere in pessime condizioni. Un decennio è stato digitalizzato dal comune di Chieri e dall'Università di Torino, i primi trent'anni sono stati inseriti nella programmazione di un progetto transfrontaliero Italia-Francia e per gli altri si stanno cercando le risorse.

In tutto questo percorso l'intenzione è quella di consolidare sempre più i rapporti con gli editori e i giornalisti, con le biblioteche civiche che conservano le collezioni dei giornali, con gli altri enti pubblici e i soggetti privati interessati e impegnati negli interventi di riproduzione. Nel futuro vogliamo integrare il portale *I Giornali del Piemonte* nell'Ecosistema culturale piemontese e in *Mèmora*<sup>4</sup> e, naturalmente, contribuire alla creazione di un'emeroteca digitale nazionale<sup>5</sup>.

Tuttavia, in questi ultimi tempi sempre più spesso ci si trova ad affrontare le richieste di soggetti che, appellandosi al diritto all'oblio, chiedono la cancellazione di articoli a loro riferiti o di intere pagine, sostenendo che per offrire la possibilità di fare ricerche per parola i portali come quello piemontese indicizzano i termini e quindi raccolgono e gestiscono illecitamente i dati personali. Senza entrare nel merito di questo tema, che richiede una riflessione specifica anche in relazione ad un quadro normativo nazionale ed europeo non chiaro, la volontà è quella di preservare l'integrità delle fonti (peraltro sempre disponibili su carta nelle biblioteche) provvedendo, quando le richieste risultano legittime, alla sola deindicizzazione dei nomi propri, tentando anche, fin quanto possibile, di non riservare l'accesso al portale ai soli utenti registrati.

<sup>2</sup> Riguardo all'attenzione posta alla conservazione dei periodici si veda: Mariagrazia Chiazza, *Progetti e problemi per la conservazione dei periodici in Piemonte*, in *Conservare il Novecento: la stampa periodica. 2. Convegno nazionale, Ferrara, Salone internazionale dell'arte del restauro e della conservazione dei beni culturali e ambientali 29-30 marzo 2001. Atti*, a cura di M. Messina e G. Zagra, Roma: Associazione italiana biblioteche, 2002, p. 116-120.

<sup>3</sup> Per l'individuazione delle testate cessate o in attività su cui intervenire si fa riferimento, oltre che alle molte ricerche locali disponibili, all'*Atlante della stampa periodica del Piemonte e della Valle d'Aosta (1789-1989)*, a cura di R. Allio, Torino: Centro studi piemontesi, 1996, recentemente aggiornato.

<sup>4</sup> *Mèmora. La nuova piattaforma digitale per i beni culturali piemontesi*, «DigItalia. Rivista del digitale nei beni culturali» 13 (2018), n. 1, p. 109-124.

<sup>5</sup> Andrea De Pasquale, *Per un'emeroteca nazionale digitale*, «Bibliothecae.it», 2018, n. 2, p. 348-370.

*Local newspapers are a precious heritage for the history of the territory and have great interest from a social, economic, cultural and historiographical point of view. Unfortunately, finding news on a paper journal is not easy at all: you need to have the newspapers and files available and you need to look very carefully being luck.*

*The Regione Piemonte offers everyone the opportunity to consult local digitalized newspapers published in Piedmont from 1846 to today. The service is free and does not require registration.*

*The [www.giornalidelpiemonte.it](http://www.giornalidelpiemonte.it) portal makes over three million pages available today, but it is constantly growing thanks to agreements with publishers and to the work of public libraries and other cultural subjects.*

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

# Urania digitale: il patrimonio storico scientifico degli osservatori astronomici italiani in Polvere di stelle e Internet Culturale

**Antonella Gasperini** - INAF Osservatorio Astrofisico di Arcetri

**Emilia Olostro Cirella** - INAF Osservatorio Astronomico di Capodimonte

*L'Istituto Nazionale di Astrofisica (INAF), il principale ente di ricerca italiano per lo studio dell'Universo, promuove, realizza e coordina attività di ricerca nei campi dell'astronomia e dell'astrofisica. Inoltre progetta e sviluppa tecnologie innovative e strumentazione d'avanguardia per lo studio e l'esplorazione del cosmo. È costituito da 19 strutture distribuite sul territorio nazionale, alcune delle quali di antica fondazione. Gli osservatori astronomici possiedono infatti un cospicuo patrimonio storico, bibliografico e archivistico che costituisce un unicum di grande interesse culturale nell'ambito degli enti di ricerca italiani ed europei. Da qualche anno è in corso un progetto complessivo di tutela e valorizzazione che ha condotto anche alla scelta di digitalizzare i volumi e le carte di archivio più significativi sia da un punto di vista culturale sia da un punto di vista iconografico. Le digitalizzazioni compiute sono a disposizione sul portale dei beni culturali dell'INAF "Polvere di stelle" dove è possibile sfogliare integralmente i volumi e consultare una scelta di documenti provenienti dagli archivi storici, in particolar modo osservazioni solari, stellari e meteorologiche fatte nel corso del XIX secolo. Inoltre dello stesso materiale bibliografico e archivistico sono state realizzate scansioni ad alta risoluzione conservate presso il repository nazionale di INAF (Italian data-center for astronomical archives) utilizzando gli stessi criteri di storage e sharing riservato ai dati astronomici. Grazie alla collaborazione di ICCU, dal 2019 è possibile consultare tali collezioni digitali con Internet Culturale ed Europea.*

**L'** Istituto Nazionale di Astrofisica (INAF), il principale ente pubblico di ricerca per lo studio delle scienze dell'Universo in Italia, è costituito da 17 strutture distribuite sul territorio nazionale, alcune delle quali di antica fondazione. Oltre a svolgere attività di ricerca nei campi dell'astronomia e dell'astrofisica, a sviluppare tecnologie innovative e a progettare strumentazione d'avanguardia, INAF conserva e promuove un importantissimo patrimonio bibliografico, archivistico e strumentale ereditato principalmente dagli Osservatori astronomici che vi afferiscono.

Tale cospicuo patrimonio rappresenta, per la sua peculiarità, un unicum di grande interesse storico-scientifico e culturale nell'ambito degli enti di ricerca italiani ed europei. Strumenti, libri e carte d'archivio di INAF costituiscono una vera e propria "collezione diffusa" su tutto il territorio italiano, che testimonia lo sviluppo della scienza astronomica e l'evoluzione della tecnica strumentale in Italia, dalle osservazioni pre-galileiane ad oggi. Custodite per lo più negli edifici che sin dall'origine ospitarono i singoli osservatori astronomici, le diverse raccolte bibliografiche, archivistiche e storico-strumentali hanno un profondo legame con l'architettura monumentale in cui sono inserite, sono fortemente radicate nella storia di ciascuna città che le ospita e sono di fondamentale rilevanza per la storia del nostro Paese, dell'Europa e del mondo.

Per quanto riguarda la dotazione libraria antica, il patrimonio delle biblioteche dell'INAF si compone oggi di circa 7000 volumi che trattano, per la maggior parte, di astronomia, fisica, matematica e di tutte quelle discipline che in passato erano campo di studio degli astronomi, alcune con importanti ricadute pratiche, quali la geodesia, la topografia e la meteorologia. Questi volumi consentono di compiere un viaggio che tocca le principali tappe del percorso di conoscenza scientifica, rappresentando delle vere pietre miliari della cultura, come ad esempio le opere di Galileo, Copernico, Tolomeo, Keplero e Newton, possedute nelle prime edizioni.

Del tutto ragguardevole e di grande pregio scientifico è il materiale documentario custodito da INAF. Tutti gli archivi storici degli Osservatori conservano carte di diversa tipologia: oltre ai diari di osservazioni astronomiche, schizzi e disegni di nebulose, comete, eclissi e pianeti, vi si trovano anche annotazioni e carteggi che rivelano il percorso compiuto dagli astronomi dei secoli passati per giungere a nuove scoperte scientifiche. Vi sono, inoltre, osservazioni meteorologiche che si protraggono ininterrotte per secoli, resoconti di viaggi in terre sconosciute, effemeridi della Luna richieste dai governanti per la programmazione dell'illuminazione pubblica in una certa città o ancora consulenze ai tribunali. Non è neppure trascurabile la presenza di documenti (lettere e carte familiari, diari, conti) relativi agli aspetti della vita privata di chi lavorava in una specola. Grazie al ruolo che gli astronomi hanno ricoperto nei secoli passati, la documentazione archivistica conservata negli Osservatori rappresenta, dunque, una fonte insostituibile non solo per la storia della scienza, ma anche per l'indagine in una gran varietà di settori con una spiccata ricaduta nella vita civile.

Nell'ambito dei fondi archivistici fondamentale importanza è rivestita anche dal patrimonio fotografico, in particolare quello inerente la fotografia astronomica, dal momento che negli Osservatori dell'INAF sono conservate le lastre fotografiche realizzate durante le eclissi o i transiti planetari, oppure intere collezioni frutto di osservazioni sistematiche effettuate ai telescopi. Vi sono anche autentiche rarità nell'ambito della fotografia scientifica, come il primo Atlante fotografico lunare della storia, che il gesuita Angelo Secchi realizzò tra il 1857 e il 1858, e che lui



stesso rese noto con il titolo *Mappe fotografiche delle principali fasi lunari fatte all'Osservatorio del Collegio Romano l'anno 1858*.

Dal 2013 INAF si è impegnato ad attuare un progetto sistematico, un vero e proprio MAB astronomico, che ha portato alla realizzazione di *Polvere di Stelle*, il *Portale dei beni culturali astronomici italiani*<sup>1</sup>, il sito web che raccoglie i database archivistici, bibliografici e strumentali di tutti i beni culturali INAF (Fig.1). Il portale costituisce uno strumento informatico che consente agli studiosi ricerche simultanee su tutti i database delle differenti tipologie di materiale storico ed è arricchito sia dalla presenza di una Teca digitale, che permette la consultazione di un primo nucleo di volumi rari e di pregio e di particolari fondi archivistici, sia dal database delle biografie degli astronomi italiani.



Figura 1. Il portale dell'INAF Polvere di stelle. I beni culturali dell'astronomia italiana

All'interno della Teca Digitale è possibile sfogliare integralmente i volumi e consultare una scelta di documenti provenienti dagli archivi storici, in particolar modo osservazioni solari, stellari e meteorologiche fatte nel corso del XIX secolo. Inoltre dello stesso materiale bibliografico e archivistico sono state realizzate scansioni ad alta risoluzione conservate presso il repository nazionale di INAF IA2 (Italian Center for Astronomical Archives) utilizzando gli stessi criteri di storage e sharing riservato ai dati astronomici.

<sup>1</sup> [www.beniculturali.inaf.it](http://www.beniculturali.inaf.it).

Per quanto riguarda le digitalizzazioni dei volumi antichi, esse sono state realizzate nel rispetto degli standard definiti dall'ICCU, producendo un primo set master di immagini TIFF ad alta risoluzione (400 ppi) non compresse, destinate alla conservazione, e due ulteriori set di immagini JPEG compresse, a risoluzione minore (300 e 150 ppi). Le immagini sono state elaborate per operare i processi di riconoscimento ottico (OCR) e generare i relativi file PDF con testo incapsulato. Infine sono stati generati i file XML in formato MAG versione 2.0.1, con le restrizioni previste per la MagTeca dell'ICCU<sup>2</sup>.

Il criterio seguito nell'individuazione dei volumi da digitalizzare è stato di tipo cronologico. Partendo dalle testimonianze più antiche e più preziose possedute da INAF, la scelta è caduta innanzitutto su un codice manoscritto della seconda metà del XIV secolo conservato presso la Biblioteca antica dell'Osservatorio di Roma a Monte Mario: una raccolta di 17 testi "classici" di astronomia molto diffusi in epoca tardo medievale che formano una sorta di compendio di astronomia ad uso probabilmente universitario (Fig.2). Sono stati poi selezionati testi pubblicati nel Cinquecento e nel Seicento che non fossero già disponibili online. È stata fatta qualche eccezione nel caso in cui gli esemplari posseduti avessero un particolare valore e caratteristiche di unicità da un punto di vista culturale o iconografico. Come, ad esempio, per l'edizione del 1543 posseduta dall'Osservatorio di



Figura 2. Tractatus De Spera di Sacrobosco all'interno della Teca Digitale di Polvere di stelle (Osservatorio astronomico di Roma)

<sup>2</sup> <http://www.internetculturale.it/opencms/opencms/it/main/partner/servizi/standard/>.

Capodimonte del *De revolutionibus orbium coelestium* di Copernico, che contiene le note di censura della Santa Inquisizione, oppure come per l'*Uranometria* del Bayer nell'edizione del 1661 posseduta dall'Osservatorio di Padova, che contiene splendide tavole raffiguranti le costellazioni. Una collezione, piccola di numero ma di particolare rilievo, è costituita dagli 11 incunaboli dell'Osservatorio astronomico di Brera che comprendono sia edizioni in latino di testi di astronomi arabi (Albumasar e Alchabitius) sia opere di Sacrobosco, Boezio e Regiomontano. Sul fronte archivistico, alla data attuale, il portale contiene le digitalizzazioni di osservazioni solari (macchie e protuberanze) fatte da Pietro Tacchini e Angelo Secchi nel periodo 1865-1880 e conservate rispettivamente presso gli Osservatori di Catania e di Roma (Fig.3).

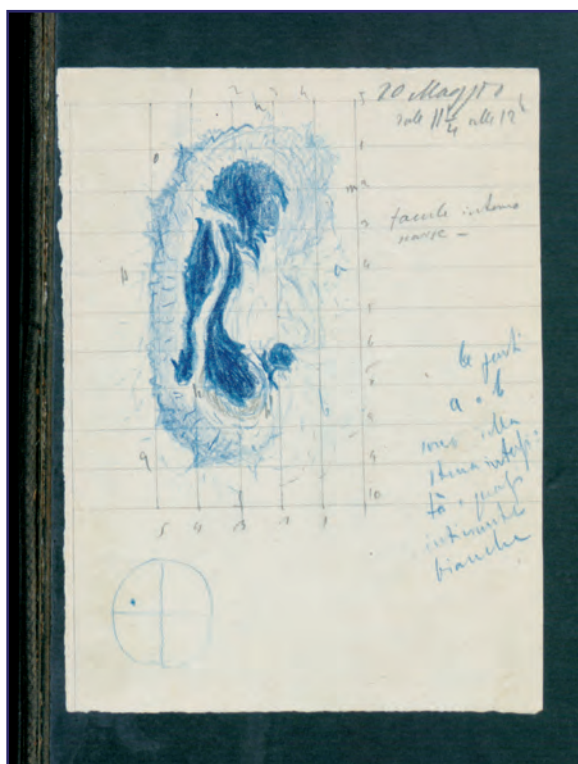


Figura 3. Osservazioni solari di Pietro Tacchini, 20 maggio 1865 (Osservatorio astrofisico di Catania)

Inoltre è presente all'interno del portale l'importante serie delle osservazioni meteorologiche eseguite a Padova per mano di Giovanni Poleni, che coprono un periodo che va dal 1725 al 1764.

Grazie alla collaborazione con ICCU, e in particolar modo grazie alla competenza e disponibilità della dott.ssa Laura Ciancio, è possibile consultare tali collezioni digitali anche all'interno di Internet Culturale e di Europeana. Ma il progetto si sta ampliando coinvolgendo anche la collezione più propriamente museale.

Il patrimonio strumentale storico dell'INAF infatti è costituito da oltre un migliaio di pezzi, retaggio

della storia scientifica degli Osservatori astronomici di più antica fondazione. Si tratta di una collezione completa che documenta in modo esaustivo l'evoluzione della strumentazione scientifica e in particolare astronomica dal XVI al XX secolo: astrolabi, quadranti, sestanti, cannocchiali, telescopi, teodoliti, cronometri, globi, strumenti matematici, strumenti geodetici e strumentazione meteorologica varia

accantonata nel tempo man mano che diveniva obsoleta per la quotidiana attività di ricerca. Questi strumenti rappresentano anche una testimonianza concreta della perizia dei meccanici e dei tecnici che li hanno costruiti, o degli astronomi che li hanno modificati per renderli più funzionali alle mutevoli esigenze osservative e di studio che si susseguivano con il trascorrere del tempo. Per questa ragione ogni singolo oggetto costituisce un pezzo unico nella storia dell'astronomia e della tecnica costruttiva degli strumenti scientifici. All'interno del portale "Polvere di stelle" la strumentazione antica è in gran parte descritta, ma è in corso un progetto per completare e uniformare le descrizioni dei pezzi museali ed arricchire l'apparato iconografico. Una volta completato tale lavoro, la strumentazione astronomica antica potrà essere disponibile, oltretutto sul sito istituzionale dell'INAF dedicato al proprio patrimonio storico, anche sulle piattaforme Internet Culturale ed Europeana.

*The National Institute for Astrophysics (INAF), the main Italian research center for the study of Universe, promotes, carries out and coordinates research activities in the fields of astronomy and astrophysics. It also projects and develops innovative technologies and state-of-the-art tool for the study and exploration of the cosmos. INAF is made up of seventeen facilities spread all over the national country, some of which were founded in ancient times. The astronomical observatories, in fact, hold a remarkable bibliographic, archival and historical instrumental heritage that represents a unique entity of great cultural interest in the context of the Italian and European research institutes. To promote its historical collection, INAF has been making an overall project of safeguarding and conserving for a few years, aimed at digitizing the most important books and archival documents both culturally and iconographically. The digitizations are available on the web portal of the INAF cultural heritage "Polvere di stelle" which allows the users to fully browse the volumes and consult a selection of documents from historical archives, in particular solar, stellar and meteorological observations made during the nineteenth century. INAF has also made high resolution digital images of the same material stored into INAF national repository (Italian data-center for astronomical archives) using the same storage and sharing methods reserved for astronomical data.*

*Thanks to collaboration of ICCU, from 2019 it is possible to consult these digital collections on the web portals of Internet Culturale and Europeana.*

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

# Linked Books: un indice citazionale per la storia di Venezia

**Giovanni Colavizza** - University of Amsterdam (NL)

**Matteo Romanello** - Digital Humanities Laboratory, EPFL (CH)

**Andrea Giuliano** - Biblioteca Nazionale Centrale di Roma (BNCR)

**Maria Cristina Mataloni** - Istituto Centrale per il Catalogo Unico (ICCU)

**Daniela Grandin** - Università Ca' Foscari Venezia

*Presentiamo i risultati del progetto Linked Books, che ha portato alla creazione di un prototipo per un indice citazionale capace di collegare il catalogo bibliotecario nazionale italiano (Opac SBN) con il sistema informativo dell'Archivio di Stato di Venezia e con i portali d'autorità e "metamotori" di ricerca internazionali (VIAF.org, Europeana). Il prototipo include 3.850.581 citazioni estratte da un corpus di 2.475 volumi, di cui 1.905 monografie e 552 numeri di rivista, da cui 5.496 articoli, riguardanti la storia di Venezia. Il progetto Linked Books ha permesso di esplorare la fattibilità e l'opportunità della creazione di un indice citazionale per discipline umanistiche, e di affrontare e risolvere ostacoli tecnici quali la creazione di un corpus rappresentativo a partire da risorse e competenze bibliografiche, la digitalizzazione dei materiali nel rispetto dei diritti d'autore, l'estrazione automatica di citazioni e lo sviluppo di interfacce di ricerca per il pubblico.*

## Introduzione

I materiali bibliografici e le fonti per lo studio nelle discipline umanistiche sono allo stesso tempo abbondanti e dipendenti dal contesto in cui si trovano. Una collezione bibliotecaria o un fondo archivistico richiedono, di conseguenza, una cura specifica per la loro preservazione, studio e messa a disposizione degli studiosi interessati. In questo modo, e specialmente in paesi come l'Italia, istituti culturali come archivi e biblioteche si sono specializzati nel valorizzare specifiche collezioni. Anche servizi digitali inter-istituzionali, come il Catalogo del Servizio Bibliotecario Nazionale<sup>1</sup> ed il Sistema Archivistico Nazionale<sup>2</sup> non comunicano tra loro, pur portando organicità e permettendo l'accesso centralizzato a (meta)dati all'interno di uno stesso ambito. Tuttavia, nella letteratura scientifica collegamenti espliciti a materiali bibliografici e fonti primarie di varia natura sono presenti sotto forma di *citazioni*, reperibili in riferimenti a piè di pagina o in liste apposite. Un *in-*

<sup>1</sup> <https://opac.sbn.it>.

<sup>2</sup> <http://san.beniculturali.it>.

dice citazionale per le discipline umanistiche che consideri questi collegamenti può (e deve) mettere in relazione ambiti documentari per ora separati, e fornire agli utenti un punto di accesso e uno strumento di navigazione finora insperato.

Il progetto *Linked Books* è nato con l'obiettivo di creare un prototipo per un indice citazionale capace di collegare il catalogo bibliotecario nazionale italiano (Opac SBN) con i sistemi informativi archivistici e con i portali d'autorità oppure con i "metamotori" di ricerca internazionali (VIAF.org, Europeana)<sup>3</sup>. Il progetto si è focalizzato su un caso di studio specifico: la storiografia su Venezia e il collegamento tra l'Opac SBN e il sistema informativo dell'Archivio di Stato di Venezia (SiasVe)<sup>4</sup>. L'interesse scientifico del progetto risiede nella volontà di studiare la fattibilità di un indice citazionale per le discipline umanistiche. Il caso di studio proposto trova invece la sua motivazione negli scopi specifici del progetto *Venice Time Machine*<sup>5</sup>. Il progetto *Venice Time Machine*, iniziato nel 2012 da una collaborazione tra l'École polytechnique fédérale de Lausanne (EPFL), l'Archivio di Stato di Venezia e l'Università Ca' Foscari di Venezia, ha come scopo la creazione di un sistema informativo digitale riguardante ogni aspetto di Venezia e della sua storia, dall'urbanistica alla geografia, dall'arte ai materiali bibliografici ed archivistici. Questi ultimi – le collezioni dell'Archivio di Stato – hanno costituito fin dal principio il cantiere di lavoro principale. La necessità, dapprima pratica poi vieppiù ragionata, di reperire dati sull'uso delle collezioni dell'Archivio che la comunità di ricerca ha fatto e continua a fare, è stata all'origine del progetto *Linked Books*: un indice citazionale per la storia di Venezia. Il progetto, iniziato nel 2014, è stato ufficializzato nel 2015 e sviluppato fino al 2018. Al progetto hanno collaborato l'EPFL, l'Archivio di Stato di Venezia, l'Università Ca' Foscari tramite il Sistema Bibliotecario di Ateneo e la Biblioteca di Area Umanistica (BAUM), la Biblioteca Nazionale Marciana, l'Istituto Veneto di Scienze, Lettere ed Arti, la Biblioteca Europea di Informazione e Cultura (BEIC) e l'Istituto Centrale per il Catalogo Unico (ICCU)<sup>6</sup>.

Nel corso di questi anni, il progetto *Linked Books* ha realizzato la digitalizzazione di un corpus di circa 2000 monografie e oltre 500 numeri di rivista, da cui sono stati estratti circa 4 milioni di riferimenti a fonti primarie e secondarie. Due strumenti di ricerca digitali sono stati inoltre sviluppati e pubblicati in linea: una biblioteca digitale (*Venice Scholar – Library*)<sup>7</sup> e l'indice citazionale (*Venice Scholar*)<sup>8</sup>. Questi due strumenti comunicano tra di loro e con tutti i cataloghi o sistemi informativi menzionati precedentemente. Il progetto *Linked Books* ha dimostrato come

<sup>3</sup> <https://viaf.org>, <https://www.europeana.eu>.

<sup>4</sup> <http://www.archiviostatovenezia.it/siasve>.

<sup>5</sup> <https://vtm.epfl.ch>.

<sup>6</sup> Il progetto è stato sostenuto dal Fondo di ricerca nazionale svizzero con due finanziamenti, progetti 205121\_159961 e P1ELP2\_168489, da un *research grant* di Europeana e dal supporto continuo di EPFL e della Fondazione Lombard Odier di Ginevra.

<sup>7</sup> <http://venicescholar.dhlab.epfl.ch/library>.

<sup>8</sup> <https://venicescholar.dhlab.epfl.ch>.

un indice citazionale possa agire come aggregatore ed integratore di sistemi informativi provenienti da ambiti culturali differenti, come archivi e biblioteche. Ha inoltre offerto lo spunto per un approccio alla creazione di indici citazionali collaborativi e distribuiti, sul modello dei cataloghi bibliotecari nazionali.

La sezione che segue offre una breve digressione sulla storia degli indici citazionali e la loro copertura delle discipline umanistiche. In seguito, presentiamo il nostro approccio alla selezione dei materiali bibliografici da cui estrarre citazioni, il contesto legale per quanto riguarda il diritto d'autore e l'acquisizione dei dati e metadati. Dettagliamo, quindi, la metodologia di estrazione delle citazioni con procedure automatizzate, nonché lo sviluppo delle interfacce di ricerca per gli utilizzatori. Il contributo si conclude con una discussione e degli spunti di lavoro per il futuro.

### Indici citazionali e discipline umanistiche

Gli indici citazionali nascono dalla volontà di superare i limiti della catalogazione per soggetto, in presenza di una crescente massa di materiale pubblicato, e offrire uno strumento non mediato per il reperimento della letteratura d'interesse. Il primo indice citazionale, *Science Citation Index (SCI)*, ebbe per antecedenti empirici gli strumenti di indicizzazione della giurisprudenza (secolo diciannovesimo) e gli esperimenti in bibliografia statistica risalenti ai decenni subito precedenti. Al primo volume dello SCI, pubblicato nel 1963, ha fatto seguito un'espansione degli studi quantitativi sulla scienza (*scientometrics*). Il legame simbiotico tra lo SCI e lo sviluppo del campo "scientometrico" è stato paragonato a quello fra il sincrotrone e la fisica delle particelle moderna<sup>9</sup>. Nonostante il lancio già nel 1978 di un indice specializzato per le discipline umanistiche, l'*Arts & Humanities Citation Index (A&HCI)*, il suo impatto per gli studiosi e per gli stessi "scientometristi" è stato quantomeno marginale<sup>10</sup>.

Diversi indici citazionali commerciali sono attualmente a disposizione della comunità scientifica, e l'offerta è in continua espansione. Gli indici a pagamento includono *Web of Science* (l'attuale evoluzione dello SCI) e *Scopus*, mentre *Google Scholar*, *Dimensions* e *Microsoft Academic*, per limitarsi a quelli con maggiore copertura, sono a libero accesso. Esistono oltretutto indici specializzati, come *Meta* per la letteratura biomedica o *Semantic Scholar* per quella in informatica. La copertura dei dati di questi indici è oggetto di costanti analisi da parte della comunità di ricerca in bibliometria, tuttavia la maggior parte degli studi concordano nel sottolineare che i dati sono in costante miglioramento, sia in quantità che qualità, e specialmente per le pubblicazioni più recenti<sup>11</sup>. Cenerentola restano le discipline

<sup>9</sup> Nicola De Bellis, *Bibliometrics and citation analysis: from the Science Citation Index to cybermetrics*, Lanham (Md): Scarecrow Press, 2009, p. 38.

<sup>10</sup> Jordy Ardanuy, *Sixty years of citation analysis studies in the humanities (1951-2010)*, «Journal of the American Society for Information Science and Technology», 64 (2013), n. 8, p. 1751-1755.

<sup>11</sup> Anne-Will Harzing – Satu Alakangas, *Google Scholar, Scopus and the Web of Science: A longitudinal and cross-disciplinary comparison*, «Scientometrics», 106 (2016), n. 2, p. 787-804.

umanistiche e le arti (*Arts and Humanities*) e, in parte, le scienze sociali, nonostante segnali di miglioramento anche in questo settore<sup>12</sup>.

Come hanno sottolineato precedenti studi, la copertura delle discipline umanistiche negli indici citazionali non è sempre del tutto esaustiva per almeno due gruppi di fattori<sup>13</sup>:

- Ragioni intrinseche che dipendono da alcune caratteristiche della letteratura umanistica. Per esempio, l'uso di varie tipologie di pubblicazioni, specialmente monografie; la varietà di lingue; la diffusa pratica di pubblicare prevalentemente in cartaceo; l'ampia varietà di fonti utilizzate e del modo di citare<sup>14</sup>.
- Ragioni estrinseche, legate invece agli ecosistemi informativi in cui l'estrazione delle citazioni avviene. Facciamo riferimento alla possibilità non soddisfacente di accedere in maniera aperta e centralizzata ai metadati necessari per identificare una fonte citata, a partire da identificativi permanenti (*PID: Permanent Identifiers*)<sup>15</sup>.

Se gli ostacoli intrinseci derivano perlopiù dalla cultura scientifica umanistica, e vanno considerati come dati di fatto da risolvere con accorgimenti appositi, gli ostacoli estrinseci ricadono invece nell'area di responsabilità delle istituzioni che si occupano di curare questi servizi.

La conseguenza è evidente: se da un lato un indice citazionale per le discipline umanistiche può offrire accesso sistematico a collezioni documentarie e bibliotecarie in maniera integrata, dall'altro la sua messa in opera offre uno stimolo alle istituzioni culturali del settore per la messa a disposizione digitale, ed in forma programmaticamente interrogabile, dei propri metadati. Punto di partenza per il raggiungimento di questo obiettivo è la collaborazione su progetti specifici ed esplorativi, come *Linked Books*.

### Creazione del corpus: selezione e digitalizzazione

Il punto di partenza del progetto è stata la domanda: in cosa consiste la storiografia su Venezia? La necessità di reperire pubblicazioni sulla storia di Venezia ricade in teoria nell'ambito della ricerca per soggetto, ma richiede in pratica un approccio

<sup>12</sup> Björn Hammarfelt, *Beyond coverage: Toward a bibliometrics for the humanities*, in: *Research assessment in the humanities*, a cura di M. Ochsner, S. E. Hug, H.-D. Daniel, Cham: Springer International Publishing, 2016, p. 115-131.

<sup>13</sup> Giovanni Colavizza – Matteo Romanello, *Citation mining of humanities journals: The progress to date and the challenges ahead*, «Journal of European Periodical Studies», in corso di pubblicazione.

<sup>14</sup> Björn Hellqvist, *Referencing in the humanities and its implications for citation analysis*, «Journal of the American Society for Information Science and Technology», 61 (2009), n. 2, p. 310-318; Anton Nederhof, *Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A review*, «Scientometrics», 66 (2006), n. 1, p. 81-100; Chris Alen Sula – Matthew Miller, *Citations, contexts, and humanistic discourse: Toward automatic extraction and classification*, «Literary and Linguistic Computing», 29 (2014), n. 3, p. 452-464.

<sup>15</sup> Cfr., ad esempio, il progetto europeo *Freya*: <<https://www.project-freya.eu>>.



pragmatico e differenziato. In questo contesto, la BAUM ha supportato adeguatamente le finalità del progetto, considerato che al suo interno raccoglie una collezione significativa di materiale bibliografico (fondi speciali, pubblicazioni, volumi di pregio) sulla storia di Venezia. La biblioteca ha consentito la conservazione e l'accessibilità di questi materiali, facilitate dall'adozione di una collocazione del materiale bibliografico prevalentemente a scaffale aperto, secondo un sistema classificatorio per materia, basato sulla Classificazione Decimale Dewey. I criteri per la selezione del corpus per il progetto sono stati: 1) la selezione di materiali principalmente dedicati a Venezia, ad esclusione della letteratura, peraltro molto abbondante, che tocca Venezia in maniera meno centrale; 2) la selezione di materiali tipicamente sottorappresentati negli indici commerciali, quali ad esempio monografie e lingue vernacolari (italiano, francese, ecc.); 3) la selezione di un volume di opere congruo ma non sovrabbondante rispetto alle risorse del progetto, anche in vista dell'espansione del corpus in fasi successive; ed infine, 4) la selezione di materiali perlopiù recenti (dal 1970 in poi).

Rispetto ai diritti d'autore, la soluzione trovata con l'accordo delle varie parti è stata quella di affidare ad EPFL la digitalizzazione delle opere su mandato delle istituzioni proprietarie delle singole copie utilizzate. L'istituzione proprietaria ha diritto a creare una copia digitale, per scopo di preservazione e di messa a disposizione dell'utenza all'interno dei propri locali fisici. Dalla copia digitale, EPFL ha creato una copia temporanea finalizzata al trattamento automatico dettagliato nel seguito, il cui obiettivo è l'estrazione di riferimenti bibliografici. Tali copie temporanee, distrutte al termine del trattamento, hanno permesso l'estrazione dei riferimenti bibliografici escludendo ogni altro contenuto sotto diritto d'autore.

Un primo gruppo di pubblicazioni individuato è stato prevalentemente di tipo monografico, in lingue nazionali che non fossero antecedenti al 1970. L'approccio euristico assunto nella progressiva costruzione del corpus si è affidato, nella fase iniziale, alle conoscenze dirette dei *subject librarians* in merito alla disponibilità di particolari fondi librari di studiosi della storia di Venezia le cui biblioteche personali erano state donate nel corso degli anni<sup>16</sup>. La conoscenza di questi fondi restituiva la certezza che le pubblicazioni estratte dal catalogo fossero coerenti con il tema del progetto. In questo modo si è potuta avviare la digitalizzazione per verificarne contestualmente gli esiti.

Successive estrazioni dal catalogo di Ateneo si sono focalizzate su una ricerca mirata per argomento, con lo scopo di recuperare risorse ordinate per classe, utilizzando la notazione numerica 945.31 (Storia di Venezia) in base al sistema di Classificazione Dewey. Il sistema classificatorio viene utilizzato in biblioteca anche per la collocazione del materiale bibliografico, un aspetto che ha consentito di

<sup>16</sup> In particolare, si fa riferimento ai fondi di Frederic C. Lane e di Gaetano Cozzi per un totale di 2.300 volumi. Le schede bio-bibliografiche sono accessibili ai seguenti indirizzi: <[https://www.unive.it/pag/fileadmin/user\\_upload/SBA/documenti/BAUM/012\\_Fondo\\_Frederic\\_Lane.pdf](https://www.unive.it/pag/fileadmin/user_upload/SBA/documenti/BAUM/012_Fondo_Frederic_Lane.pdf)> e <[https://www.unive.it/pag/fileadmin/user\\_upload/SBA/documenti/BAUM/007\\_Fondo\\_Gaetano\\_Cozzi.pdf](https://www.unive.it/pag/fileadmin/user_upload/SBA/documenti/BAUM/007_Fondo_Gaetano_Cozzi.pdf)>.

operare un'ulteriore selezione direttamente a scaffale. Un approccio pragmatico che ha permesso di escludere dal corpus bibliografico, ad esempio, le pubblicazioni che presentavano nel testo un corredo di illustrazioni e immagini piuttosto significativo: libri coerenti con l'argomento ricercato, ma con una assente o ridotta lista di riferimenti.

Successive estrazioni si sono affidate ad una ricerca per soggetto utilizzando più descrittori; su tutti, "Venezia" e "Venezia-Storia". La ricerca per soggetto ha restituito un complesso di risorse piuttosto consistente, anche se più limitato rispetto alle opere disponibili nelle collezioni. Non tutte le notizie bibliografiche, infatti, sono corredate da accessi semantici e potevano quindi essere viziate dall'uso, in passato, di vocabolari non uniformi o controllati. D'altra parte, si deve considerare che il catalogo è frutto di un processo di progressiva stratificazione che nel tempo ha sovrapposto norme e persone e può quindi soffrire di una mancanza di uniformità nella qualità delle descrizioni bibliografiche. Nell'espansione del corpus bibliografico sono state inoltre comprese pubblicazioni periodiche focalizzate sulle tematiche della storia veneziana che hanno richiesto la collaborazione con altre istituzioni veneziane, in particolare l'Archivio di Stato di Venezia e la Biblioteca Nazionale Marciana, per la ricostruzione della serie completa delle annate<sup>17</sup>.

Le liste di titoli elaborate dai *subject librarians* nel corso del tempo hanno prodotto complessivamente più di 22.000 risultati che sono stati oggetto di una valutazione e di una deduplicazione, prima di essere sottoposti alla digitalizzazione per l'estrazione delle citazioni. La digitalizzazione delle opere è stata messa in atto utilizzando uno scanner fornito da EPFL e personale specializzato, avendo cura di associare ad ogni volume digitalizzato i metadati disponibili dal polo SBN di Venezia. La taglia finale del corpus, dopo una successiva campagna di espansione, è di 2.475 volumi, di cui 1.905 libri, perlopiù monografie, e 552 numeri di rivista, da cui 5.496 articoli sono stati estratti<sup>18</sup>.

### Creazione del corpus: citazioni e uso dei dati del catalogo SBN

La creazione di un corpus di citazioni bibliografiche è un processo che consiste di due fasi distinte: in primo luogo l'estrazione (*citation extraction*) e, in secondo luogo, la loro disambiguazione (*citation matching*). L'estrazione richiede di classificare ciascuna parola del testo di partenza come facente parte di una citazione oppure no e se sì, di determinare il ruolo che essa ricopre al suo interno (ad es. in-

<sup>17</sup> Si tratta in particolare di pubblicazioni periodiche: Archivio Veneto, Studi Veneziani, gli Atti dell'Istituto Veneto di Scienze, Lettere ed Arti e Ateneo Veneto.

<sup>18</sup> Per un approfondimento riguardo alla selezione del corpus e sua coerenza rispetto alle citazioni da esso estratte in seguito, rimandiamo il lettore a Giovanni Colavizza – Matteo Romanello – Frédéric Kaplan, *The references of references: A method to enrich humanities library catalogs with citation data*, «International Journal on Digital Libraries», 19 (2017), n. 2-3, p. 151-161.

dicazione dell'autore, titolo, anno di pubblicazione, ecc.). La disambiguazione, invece, consiste nello stabilire, per ogni citazione così estratta, un legame tra la citazione stessa ed un record bibliografico in modo da poter ricondurre due o più citazioni della stessa pubblicazione al medesimo oggetto bibliografico citato<sup>19</sup>.

Risorse esistenti di metadati bibliografici, come il Catalogo del Servizio Bibliotecario Nazionale (SBN), rivestono un ruolo fondamentale per la disambiguazione di citazioni in quanto contengono già delle descrizioni dettagliate delle pubblicazioni che troviamo poi citate nel nostro corpus sulla storia di Venezia. Tali descrizioni sono estremamente preziose per qualsiasi algoritmo che cerchi di stabilire una corrispondenza tra citazione e pubblicazione citata.

È in questo frangente che il progetto *Linked Books* ha visto la collaborazione tra EPFL e ICCU al fine di migliorare la fruibilità dei *record* bibliografici contenuti in SBN. Tale collaborazione ha permesso di risolvere alcuni ostacoli esistenti nell'utilizzo di SBN, dovuti principalmente alle modalità di accesso ai dati – soprattutto vista la loro mole imponente – e al formato dei dati stessi, permettendo di accedere facilmente ai metadati contenuti in ciascun *record*.

La prima opzione considerata è stata l'utilizzo delle API (*Application Programming Interface*) implementate per l'app "OPAC SBN"<sup>20</sup>, realizzata per l'ICCU da Inera Srl<sup>21</sup>. Fra le altre funzionalità messe a disposizione da queste API c'è l'estrazione di un record in formato JSON dato il suo BID, l'identificatore univoco che viene assegnato ad ogni notizia bibliografica in SBN. JSON è un formato di dati testuale aperto, molto diffuso in ambito web per lo scambio di dati tra applicativi. Il formato JSON è ben noto alla maggioranza degli sviluppatori, a differenza di UNIMARC che, pur estremamente ricco, è comunque poco diffuso al di fuori dell'ambito biblioteconomico. L'*output* in JSON sarebbe stato facilmente utilizzabile per gli scopi del progetto, ma questo approccio presenta evidenti criticità.

La prima riguarda le prestazioni: una prova di *download* di 10.000 BID di tipologia varia richiede circa 45 minuti. Pertanto, per scaricare gli oltre 16 milioni di record attualmente in SBN, sarebbero stati necessari almeno cinquanta giorni, senza interruzioni. Oltre ad essere incompatibile con le tempistiche del progetto, questo approccio avrebbe avuto un serio impatto sul sistema OPAC, rischiando di degradare le sue prestazioni. Una seconda criticità è la mancanza, nell'*output* prodotto dalle API, di alcune informazioni importanti come soggetti, classificazioni Dewey, altri titoli e relazioni gerarchiche.

Per risolvere queste criticità, è stato realizzato un programma che, partendo dai

<sup>19</sup> Usiamo il termine "citazione" per indicare sia la porzione di testo contenente un riferimento bibliografico, che il legame tra una pubblicazione (citante) ed una fonte (citata). Il contesto chiarisce l'accezione usata.

<sup>20</sup> <<https://play.google.com/store/apps/details?id=it.inera.opacmobile&hl=it>> e <<https://itunes.apple.com/it/app/opac-sbn/id786155255>>.

<sup>21</sup> <https://www.inera.it>.

*record* UNIMARC estratti periodicamente dall'Indice SBN e destinati all'aggiornamento del Catalogo online, producesse record in un formato JSON ispirato a quello prodotto dalle API per l'applicazione mobile, con l'aggiunta delle informazioni mancanti ed eventuali altre migliorie. Le prime prove, effettuate su un campione di 100.000 record, hanno dato risultati molto confortanti anche dal punto di vista delle prestazioni: eseguiti su un normale *personal computer*, hanno richiesto meno di dieci minuti. Questo approccio, basato su un'elaborazione offline, evita di sovraccaricare il Catalogo online da un carico di lavoro eccessivo<sup>22</sup>.

Dopo una prova effettuata sull'intero scarico UNIMARC di oltre 16 milioni di record, durato solo qualche ora, è cominciata la fase di perfezionamento dell'output per renderlo più adatto alle esigenze del sistema di estrazione delle citazioni. Rispetto all'output delle API mobile sono state aggiunte, ad esempio, le classificazioni Dewey, i soggetti e altri titoli. Per quanto riguarda le localizzazioni, invece, si è preferito esportare solo il codice ISIL<sup>23</sup>, omettendo le denominazioni per esteso delle biblioteche e rimandando agli *open data* disponibili sul sito dell'Anagrafe delle biblioteche italiane per la decodifica dei codici ISIL.

Lo scarico dei dati SBN in formato JSON viene poi indicizzato in un'istanza locale di *Elastic Search*, un sistema *open source* di indicizzazione e ricerca di dati testuali, liberamente scalabile ed altamente performante. Questa soluzione permette di cercare nei metadati dei record SBN, permettendo così di usarli per disambiguare i riferimenti bibliografici estratti. Tuttavia, il limite principale di questa soluzione, che d'altro canto è inevitabile vista l'attuale infrastruttura tecnica di SBN, è la perdita di sincronizzazione con i dati in SBN<sup>24</sup>. Tali dati, infatti, sono tutt'altro che statici: nuove acquisizioni producono nuovi record, ed altri scompaiono in seguito alla rimozione di duplicati. Una soluzione ottimale dovrebbe permettere di usare SBN e altri sistemi analoghi come una corrente continua di dati (dinamica e costantemente in divenire), piuttosto che come uno scarico statico e destinato a diventare obsoleto.

L'estrazione di citazioni, nel progetto *Linked Books*, è avvenuta in due fasi successive. Nella prima fase citazioni a fonti primarie e secondarie sono state estratte dal corpus di pubblicazioni tramite un modello statistico che fa uso di citazioni manualmente estratte ed annotate. Tale modello consente di determinare, con un certo livello di accuratezza, l'inizio e la fine di ciascun riferimento, così come di classificare le informazioni che compongono una citazione (tipicamente nome degli autori, titolo, anno di pubblicazione, editore, ecc.). Questo modello ha poi per-

<sup>22</sup> Il programma è stato realizzato in Java 8, utilizzando le librerie MARC4J, JSONSimple e GSON, tutte distribuite sotto licenze aperte (LGPL 2.1 e Apache 2).

<sup>23</sup> <https://anagrafe.iccu.sbn.it/it/informazioni/cosa-e-lisil>.

<sup>24</sup> A questo riguardo, si segnala l'interessante proposta di utilizzare il software aperto Wikibase per la pubblicazione linked open data di dati UNIMARC, avanzata da Giovanni Bergamin – Cristian Bacchi, *New ways of creating and sharing bibliographic information: an experiment of using the Wikibase Data Model for UNIMARC data*, «JLIS.it» 9 (2018), p. 35-74.

nesso di estrarre 3.850.581 citazioni dal corpus considerato. La seconda fase, invece, ha permesso di collegare una citazione ad una autorità bibliografica o archivistica, e più precisamente SiASVe per i riferimenti a fonti di archivio ed i dati di ICCU e del *Virtual International Authority File* (VIAF) per le fonti secondarie<sup>25</sup>. Le citazioni così estratte possono essere esplorate dagli utenti utilizzando due applicazioni di ricerca, discusse nella sezione che segue.

### **Design e sviluppo di interfacce di ricerca ed esplorazione: Venice Scholar e Venice Scholar – Library**

Due applicazioni per la gestione e ricerca dei dati sono state sviluppate nel corso del progetto *Linked Books*: la *Venice Scholar Library* (VSL)<sup>26</sup> e il *Venice Scholar Index* (VSI)<sup>27</sup>. Si tratta, nel primo caso, di un applicativo per la biblioteca digitale (*digital library software*) che permette agli utenti di gestire ed esplorare i materiali digitalizzati. Nel secondo caso invece, del vero e proprio indice citazionale che permette agli utenti di esplorare le pubblicazioni sulla storia di Venezia seguendo i collegamenti tra di esse stabiliti dalle citazioni bibliografiche e di accedere ai dati disponibili per ciascuna pubblicazione (ad esempio, numero di citazioni date e ricevute).

Il design e sviluppo di entrambe le interfacce si è svolto seguendo un processo iterativo. In una prima fase di concezione e design, il team del progetto ha lavorato a stretto contatto con un designer per far sì che le interfacce fossero il più possibile intuitive ed ergonomiche. Terminato il design si è passati all'implementazione delle interfacce sotto forma di prototipi, che sono stati poi sottoposti ad una fase di *beta testing* al fine di ottenere dei commenti che potessero migliorarne l'usabilità.

Entrambe le interfacce sono state sottoposte ad un periodo di *beta testing* da parte di due gruppi di utenti. Un primo gruppo di cinque bibliotecari della biblioteca BAUM dell'Università Ca' Foscari di Venezia, ed un secondo gruppo composto da storici, archivisti e bibliotecari provenienti dalle altre istituzioni parte del progetto. Questa fase di prova è risultata di grande utilità per migliorare l'usabilità delle interfacce in quanto ha permesso di individuare elementi che erano fonte di confusione per l'utente (ad esempio, denominazioni dei pulsanti) o funzionalità che si sono rivelate insoddisfacenti, e pertanto sono state migliorate (come la precisione della ricerca testuale).

Passiamo ora ad esaminare le funzionalità principali di ciascuna interfaccia. La VSL offre una serie di funzionalità che sono tipiche di qualsiasi *digital library*, affiancate

<sup>25</sup> Per maggiori dettagli sui processi automatici di estrazione e disambiguazione e sulla loro accuratezza si veda Giovanni Colavizza – Matteo Romanello – Frédéric Kaplan, *The references of references: A method to enrich humanities library catalogs with citation data*, «International Journal on Digital Libraries», 19 (2017), n. 2-3, p. 156-158.

<sup>26</sup> <https://venicescholar.dhlab.epfl.ch/library>.

<sup>27</sup> <https://venicescholar.dhlab.epfl.ch>.

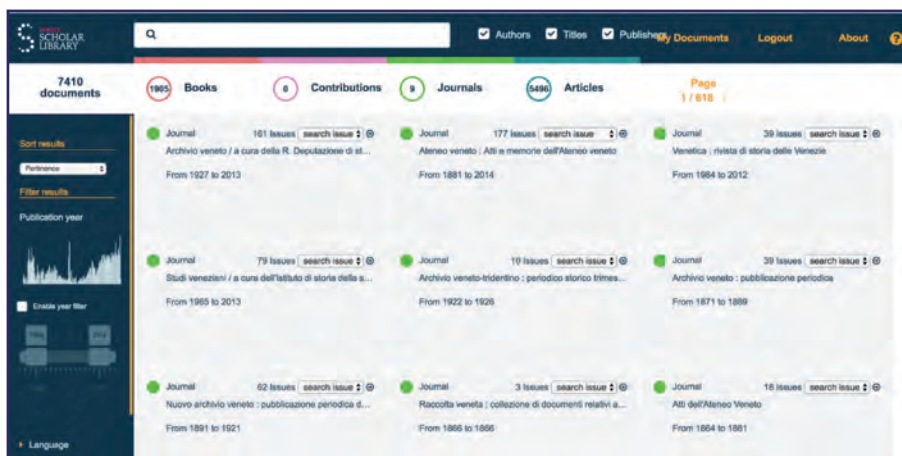


Figura 1. Venice Scholar Library: visualizzazione dei risultati di ricerca sui metadati.



Figura 2. Venice Scholar Library: modalità di lettura del materiale digitalizzato con immagine digitale ed OCR visualizzati una accanto all'altro.

da alcune funzionalità avanzate che permettono agli utenti amministratori di modificare manualmente le citazioni estratte automaticamente dalla VSL. Nella prima tipologia di funzionalità rientrano la visualizzazione e ricerca sui metadati del digitalizzato (Fig. 1), una modalità di lettura dove immagine digitale e testo prodotto dal riconoscimento automatico dei caratteri (OCR: *Optical Character Recognition*) sono visualizzati una accanto all'altro (Fig. 2) e la possibilità di scorrere rapidamente attraverso le miniature delle pagine digitalizzate. Per quanto riguarda le funzionalità avanzate, invece, la VSL offre agli utenti una modalità di lettura del testo dove le citazioni bibliografiche estratte automaticamente possono essere visualizzate ed eventualmente modificate (nel caso di utenti amministratori) (Fig. 3).

Sebbene nel corso del progetto *Linked Books* non ci sia stata alcuna attività di correzione manuale delle citazioni estratte, avere uno strumento che consente di

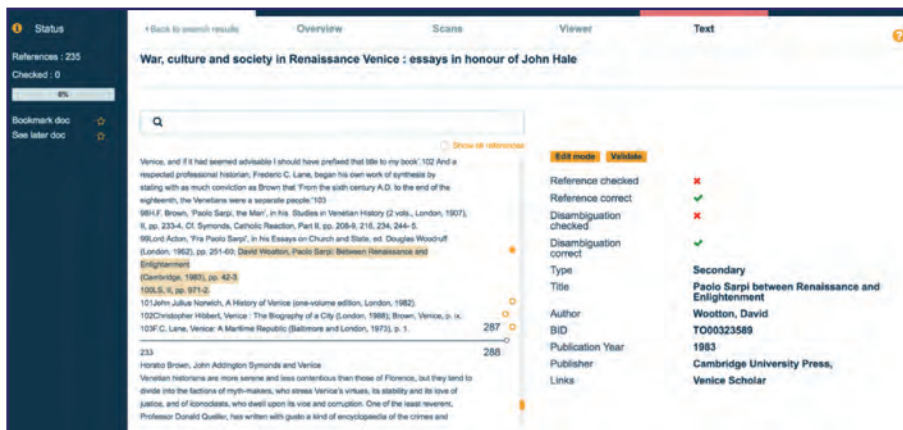


Figura 3. Venice Scholar Library: modalità di lettura per utenti amministratori con funzionalità di correzione dei riferimenti bibliografici.

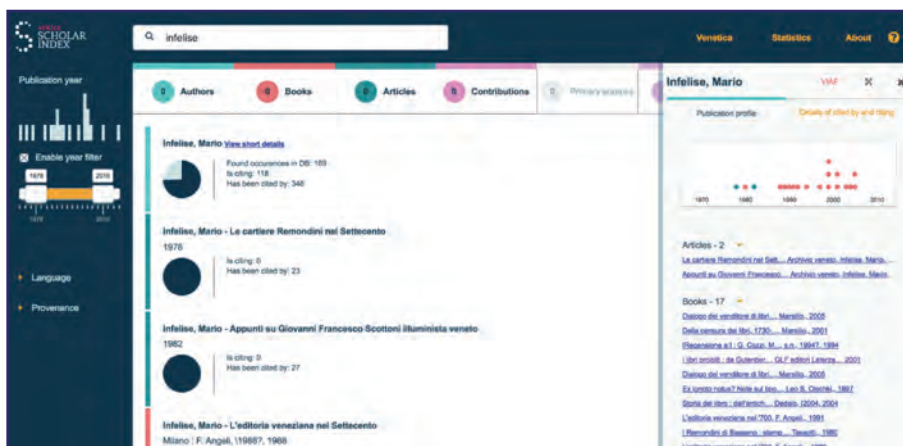


Figura 4. Venice Scholar Index: pagina dei risultati di ricerca.

farlo apre la strada ad un possibile coinvolgimento attivo di bibliotecari e archivisti non solo nell'attività di digitalizzazione ma anche in quella di miglioramento e correzione dei dati citazionali ricavati dalle pubblicazioni digitalizzate<sup>28</sup>. Tale correzione manuale è necessaria ed auspicabile al fine di garantire la maggiore accuratezza possibile delle analisi quantitative che un indice citazionale rende possibili.

L'indice di citazioni vero e proprio può essere esplorato attraverso una seconda interfaccia, il Venice Scholar Index. A differenza della VSL, l'accesso al VSI è intera-

<sup>28</sup> Il coinvolgimento dei bibliotecari nel miglioramento dei dati citazionali è al centro di un altro progetto di ricerca, la *Linked Open Citation Database*, sul quale si veda Anne Lauscher – Kai Eckert – Lukas Galke – Ansgar Scherp – Syed Tahseen Raza Rizvi – Sheraz Ahmed – Andreas Dengel – Philipp Zumstein – Annette Klein, *Linked Open Citation Database*, in: *Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries*, New York, ACM Press, 2018.



Figura 5. Venice Scholar Index: profilo citazionale dell'autrice Patricia Fortini Brown. Le cifre sul numero di citazioni fatte o ricevute sono cumulative rispetto a tutte le pubblicazioni dell'autrice che sono state indicizzate.

mente pubblico in quanto i dati che contiene e su cui si basa (soprattutto metadati bibliografici) non sono soggetti a diritto d'autore.

La prima pagina che accoglie l'utente presenta la consueta maschera di ricerca attraverso la quale è possibile effettuare una ricerca su a) nomi di autori, b) titoli di pubblicazioni citate (articoli o monografie), c) nomi di fonti d'archivio (fonti primarie) e d) sul testo stesso delle citazioni estratte. I risultati della ricerca vengono poi mostrati suddivisi per categoria e con la possibilità di filtrarli ulteriormente sulla base di criteri come anno di pubblicazione, lingua e biblioteca o archivio di provenienza del materiale digitalizzato (Fig. 4).

Ogni autore citato all'interno del corpus possiede un profilo delle pubblicazioni ed uno delle citazioni. Le pubblicazioni includono titoli che sono stati digitalizzati nel corso del progetto (e sono pertanto disponibili per la lettura attraverso la VSL), così come titoli non digitalizzati ma citati da altre pubblicazioni. Il profilo delle citazioni, disponibile sia per autori che singole pubblicazioni, presenta da un lato una visualizzazione della frequenza delle citazioni (sia in entrata che in uscita, cioè fatte e ricevute) e dall'altro offre la possibilità di esaminare la lista di pubblicazioni citate/citanti (Fig. 5).

Un aspetto che distingue il VSI da altri indici citazionali esistenti, come *Google Scholar*, è il fatto di indicizzare le citazioni a fonti primarie, e nello specifico i documenti archivistici. Il profilo citazionale di ogni autore contiene una sezione dedicata alle fonti primarie dove è possibile esplorare quali fondi, serie o sottoserie d'archivio sono stati citati da un certo autore nella totalità delle sue pubblicazioni (Fig. 6). Un altro punto di accesso alle fonti primarie nell'interfaccia del VSI è fornito dal profilo citazionale disponibile per ogni fonte primaria citata all'interno del corpus.



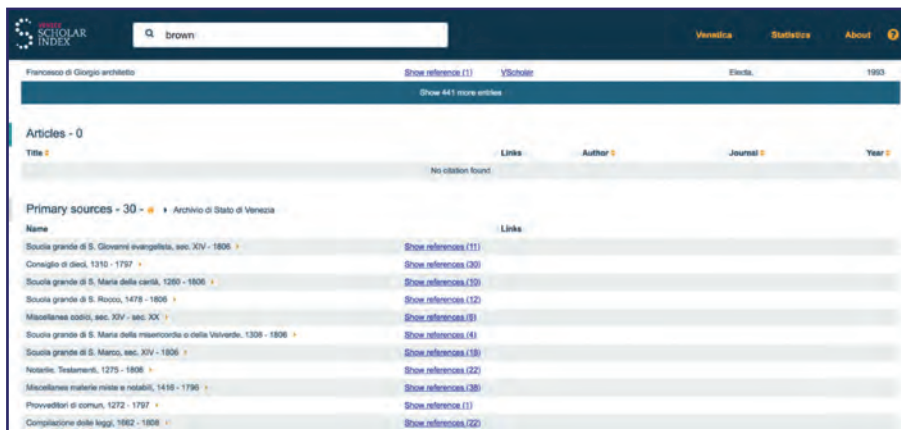


Figura 6. *Venice Scholar Index: fonti primarie citate dall'autrice Patricia Fortini Brown.*

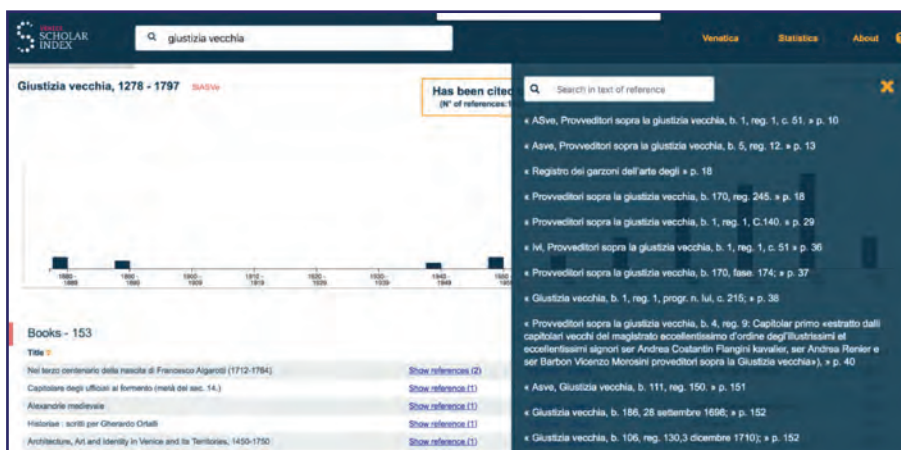


Figura 7. *Venice Scholar Index: profilo citazionale del fondo "Giustizia Vecchia" dell'Archivio di Stato di Venezia.*

Per ciascuna fonte il VSI fornisce a) una *timeline* che mostra la frequenza delle citazioni ricevute per decennio (Fig. 7) e b) una lista dettagliata delle pubblicazioni citanti, suddivise per tipologia (articolo di rivista, monografie, ecc.).

A concludere questa carrellata, una funzionalità sperimentale che è stato possibile sviluppare ed aggiungere al VSI grazie ad un finanziamento messo a disposizione da *Europeana Research* nel quadro del suo *Research Grants Programme 2018*<sup>29</sup>. Per ogni risorsa disponibile nel VSI (autore moderno, fonte secondaria, fonte primaria), nella pagina ad essa dedicata, un pulsante con il logo di *Europeana* consente di accedere ad un menù laterale con oggetti rilevanti a tale risorsa e provenienti da *Europeana* (Fig. 8). Tale lista di risorse è creata dinamicamente aggre-

<sup>29</sup> <https://pro.europeana.eu/post/developing-the-venice-scholar-matteo-romanello-europeana-research-grant-winner>.

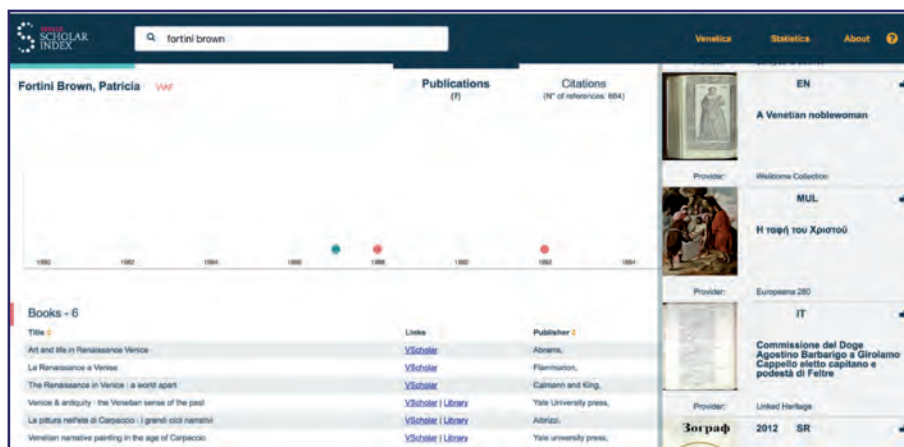


Figura 8. Venice scholar index: Europeana sidebar.

gando e filtrando dati provenienti dalla *Europeana Search API*<sup>30</sup>. Tra i milioni di oggetti digitali contenuti nel meta-aggregatore, un sistema basato sulle parole chiave estratte dai titoli delle pubblicazioni seleziona gli oggetti potenzialmente più rilevanti per l'utente. Tale funzionalità ci ha permesso, da un lato, di mostrare come strumenti diversi possano venire integrati insieme in modo produttivo e, dall'altro, di creare una funzionalità capace di amplificare l'aspetto di serendipità<sup>31</sup> inerente ad ogni ricerca effettuata attraverso un motore di ricerca.

## Conclusione

Il progetto *Linked Books* costituisce un'esplorazione della fattibilità e delle potenzialità degli indici citazionali nelle discipline umanistiche. Pensato con l'obiettivo di collegare cataloghi bibliotecari, sistemi informativi archivistici e rispettive piattaforme digitali per il tramite della storiografia, il progetto ha permesso di porre ed affrontare una serie di questioni chiave rispetto al futuro dell'indicizzazione di risorse bibliografiche e documentarie in un futuro sempre più digitale. Prima di tutto, la definizione di un corpus rappresentativo, seguita da digitalizzazione e analisi in un contesto legale appropriato. Cruciale, in questa fase, non solo la presenza di collezioni bibliografiche adeguate, ma anche di competenze pluridisciplinari (esp. *subject librarians*). A seguire, l'uso di metodi automatizzati per estrarre citazioni dalla letteratura, che richiede la disponibilità di un rapido accesso ai cataloghi o sistemi informativi di riferimento per quanto riguarda identificativi univoci e meta-dati (e.g. SBN). Per concludere, lo sviluppo di prototipi di applicativi: una biblioteca digitale (*Venice Scholar - Library*) e un indice citazionale (*Venice Scholar*).

<sup>30</sup> <https://pro.europeana.eu/resources/apis/search>.

<sup>31</sup> Wikipedia s.v. serendipità: «Il termine serendipità indica la fortuna di fare felici scoperte per puro caso e, anche, il trovare una cosa non cercata e impreveduta mentre se ne stava cercando un'altra».

Merita di essere sottolineato quello che è forse ad oggi l'ostacolo principale per progetti come *Linked Books*: la possibilità (o sua assenza) di accedere programmaticamente (per esempio tramite API aperte) a metadati archivistici e bibliotecari strutturati in maniera uniforme e secondo stabiliti standard internazionali. Idealmente, sotto forma di *linked data*. Costruire dei servizi che si basino su questi metadati e sulla loro aggregazione sarà sempre difficoltoso, se non irrealizzabile, senza un maggiore sforzo istituzionale rivolto non solo alla cura di metadati e alla loro pubblicazione tramite interfacce dedicate, ma anche alla loro accessibilità per un uso programmatico/automatico. Anche per queste ragioni, l'ICCU sta avviando un progetto denominato Sistema di Ricerca Integrato (SRI) che porterà alla realizzazione di un accesso *linked open data* ai dati del catalogo SBN.

Il progetto *Linked Books* ha infine permesso di concepire un progetto di più ampio respiro, lo *Scholar Index*: un indice citazionale aperto, collaborativo e distribuito<sup>32</sup>. Seguendo la tradizione dei cataloghi bibliotecari, questo progetto propone di affidare alla responsabilità di ciascuna istituzione culturale facente parte del progetto, una parte relativamente piccola della letteratura da indicizzare. Compito dell'istituzione è quindi quello di digitalizzare le pubblicazioni di sua competenza e, con l'ausilio dell'automazione permessa dagli applicativi sviluppati, curare l'estrazione delle citazioni. Questi dati citazionali possono poi essere aggregati, ad esempio in infrastruttura esistente come *Open Citations*<sup>33</sup>, e messi a disposizione dell'utenza tramite interfacce o come *linked data*. Riteniamo che il futuro dell'indicizzazione digitale di materiali documentari sarà marcato dalla connettività, e specificamente dal superamento di confini informativi istituzionali. I legami citazionali offrono in questo senso un'opportunità ancora del tutto inesplorata.

*We present the outcomes of the Linked Books project, resulting in a prototype citation index interlinking the Italian national library catalog (Opac SBN) with the information system of the State Archive of Venice and international authority records or "metaengines" such as VIAF.org and Europeana. Our prototype includes 3.850.581 citations extracted from a corpus of 2.475 volumes, of which 1.905 monographs, and 552 journal volumes, or 5.496 articles therein. The corpus is focused on the history of Venice. The Linked Books project allowed us to explore the feasibility and desirability of a citation index for the humanities, and to face and solve technical challenges including: the selection of a thematically representative corpus from bibliographic resources and expertise, the digitization of these materials within the bounds of copyright, the automatic extraction of citations and the development of public search interfaces.*

<sup>32</sup> <<https://scholarindex.eu>>, il progetto è in discussione come potenzialmente facente parte del progetto europeo Time Machine (<<http://timemachine.eu>>).

<sup>33</sup> <<http://opencitations.net>>. Silvio Peroni – Alexander Dutton – Tanya Gray – David Shotton, *Setting Our Bibliographic References Free: Towards Open Citation Data*, «Journal of Documentation», 71 (2015), n. 2, p. 253-77.

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

# Phaidra, un archivio digitale FAIR per la disseminazione e l'accesso integrato a testi, testimonianze, immagini e storie del patrimonio culturale

**Lorisa Andreoli — Elena Bianchi — Linda Cappellato — Yuri Carrer — Gianluca Drago — Giulio Turetta — Antonella Zane e Laura Tallandini**

*Università degli Studi di Padova, Sistema Bibliotecario di Ateneo*

*Dalla fine degli anni Ottanta del secolo scorso le tecnologie dell'informazione hanno permesso la costruzione dei cataloghi online delle biblioteche. L'Università di Padova ha velocemente realizzato il catalogo ad accesso libero, oggi con oltre 2 milioni di record, comprensivo delle biblioteche accademiche e di tutte le istituzioni culturali della città. Ciò ha messo in evidenza la continuità storica e la ricchezza del patrimonio librario a partire dall'origine dell'Ateneo nel 1222.*

*Dalla realizzazione del catalogo online, che partecipa al Servizio Bibliotecario Nazionale (SBN), ha preso avvio un lavoro di digitalizzazione di fondi librari antichi e di collezioni archivistiche e museali al fine di rappresentare personaggi e sviluppi salienti della storia di Padova e delle Scienze. Questa attività si è inserita nella cosiddetta Terza Missione, ovvero l'insieme delle attività con le quali gli atenei possono entrare in interazione diretta con la società in un costante dialogo tra l'ateneo e i cittadini.*

*Lo sviluppo richiedeva un'infrastruttura che garantisse l'archiviazione a lungo termine, la descrizione e l'indicizzazione degli oggetti digitali, l'accesso pubblico da un'interfaccia web, con ricerca semplice e avanzata, condizioni d'uso e modelli di licenza per gli oggetti digitali. Nel 2010 il Sistema Bibliotecario di Ateneo, dopo attenta analisi, ha scelto Phaidra, il repository dalle caratteristiche FAIR sviluppato all'origine dall'Università di Vienna su software open source Fedora.*

*La piattaforma Phaidra ospita attualmente 400.000 oggetti digitali in ambito GLAM. Nel corso di quasi dieci anni di lavoro, Phaidra si è arricchita di strumenti utili a favorire la disseminazione e il riuso delle risorse depositate (tra questi, book viewer, collection viewer, image viewer e API) e ha inoltre costantemente ricercato e attivato il dialogo con altre piattaforme, quali Europeana e Internet Archive, e sistemi software, quali Movio, il kit open source dell'ICCU per la creazione di mostre virtuali.*

## Perché la piattaforma Phaidra

Dalla fine degli anni Ottanta le tecnologie dell'informazione (ICT) hanno rivoluzionato le prospettive di accesso alla documentazione scientifica e umanistica e ai patrimoni librari portando alla convinzione che, a breve, tutto si sarebbe trovato nella rete. Si può parlare dell'avvio di una nuova "società della conoscenza".

Le università sono storicamente attori nei processi di sviluppo culturale, sociale ed economico, in quanto nodo di reti di relazioni comprendenti imprese, enti culturali, istituti di ricerca, parchi tecnologici, archivi, musei e biblioteche.

Le nuove prospettive, *ICT-compliant*, hanno messo in evidenza la continuità storica degli atenei come imprese che istituzionalmente investono sull'aggiornamento dell'informazione scientifico-accademica e sui prodotti della ricerca. Le università sono divenute punto di riferimento aperto non solo alle realtà scientifiche ma anche al cittadino.

Così negli atenei, oltre alla didattica superiore e alla ricerca, si sono andate rinforzando, e sono state riconosciute, anche in via istituzionale, le attività per la "terza missione" ovvero «l'insieme delle attività con le quali gli atenei possono entrare in interazione diretta con la società», con la produzione di beni pubblici di natura sociale, educativa e culturale, in un costante dialogo con i cittadini, grazie allo strumento digitale che annulla le distanze fisiche e rende partecipi delle realizzazioni online tutti quanti stiano nella rete.

I sistemi bibliotecari universitari, da sempre essenziale infrastruttura di servizio per didattica e ricerca, hanno colto sollecitamente l'enorme potenzialità dei nuovi strumenti procedendo negli anni Novanta alla informatizzazione dei loro cataloghi. Così in Padova si è realizzato un catalogo unico di tutte le biblioteche del territorio<sup>1</sup>, ad accesso universale, ricco di oltre 2 milioni di documenti, condiviso con il catalogo SBN e anche con il catalogo mondiale WorldCat<sup>2</sup>.

Grazie alla informatizzazione dei cataloghi i contenuti delle biblioteche, accademiche e non, sono apparsi in tutta la loro estensione e i fondi storici, tradizionalmente punto di arrivo e ricerca di studiosi specializzati, hanno acquisito una visibilità prima impensabile attivando una vasta aspettativa di illustrazione di personaggi, eventi, attività scientifiche e processi storici.

Di qui, già dai primi anni Duemila, è iniziata da parte del Sistema Bibliotecario di Ateneo dell'Università di Padova l'attività di digitalizzazione di importanti fondi librari per renderli fruibili senza comprometterne la conservazione. Questa attività ha ulteriormente accresciuto l'attenzione verso un patrimonio culturale ricchissimo da parte di un pubblico ampio che, rilevandone l'importanza, ne riconosce il valore in una sorta di riscoperta della propria storia di comunità, volta a volta locale, regionale, nazionale, universale.

<sup>1</sup> Catalogo del Sistema Bibliotecario Padovano <<https://catalogo.unipd.it>>.

<sup>2</sup> Limitatamente al patrimonio delle biblioteche dell'Università.

Le biblioteche accademiche hanno quindi esteso la digitalizzazione anche a materiali di diversa natura, dalle collezioni di fotografie storiche alle tavole didattiche; di qui, per immediata estensione si è aperta la necessità di collegare tali materiali con quelli degli archivi e dei musei. Questo richiedeva lo sviluppo di un'infrastruttura tecnologica, in grado di gestire i sistemi semantici delle informazioni di fonti eterogenee del patrimonio culturale e di consentirne l'integrazione con strumenti trasparenti all'utente, quali ad esempio i *Linked Open Data*.

Si imponeva in definitiva la necessità di promuovere un sistema di servizi digitali che non riguardasse solo il patrimonio librario. Il Sistema Bibliotecario di Ateneo decideva allora di condurre un'analisi sui requisiti architetturali e funzionali di un sistema di archiviazione e presentazione di oggetti e collezioni digitali. La valutazione di software, sia commerciali che open source, per le collezioni digitali di università e altre istituzioni culturali portava a prendere in considerazione Phaidra (Permanent Hosting, Indexing and Archiving of Digital Assets and Resources)<sup>3</sup>, la piattaforma DAMS (Digital Asset Management System) per l'archiviazione a lungo termine di oggetti digitali, creata e sviluppata dall'Università di Vienna<sup>4</sup>.

Lo studio da vicino del progetto e il confronto con i colleghi austriaci sugli indirizzi di sviluppo del sistema hanno portato alla scelta di adottare Phaidra a Padova. Tale scelta, formalizzata nel 2010 con un accordo di collaborazione e sviluppo tra le due università, ha tenuto conto della presenza di una tecnologia consolidata per l'archiviazione degli oggetti digitali, del contenimento dei costi, dell'aspetto innovativo della piattaforma e della possibilità di sviluppi modulari in una prospettiva europea. La collaborazione tra i due atenei oltre alla condivisione di Phaidra ha permesso lo scambio di esperienze e competenze in un contesto europeo, ciascuno con la propria identità e specificità riconoscibile nella diversa presentazione dei siti web di Padova e di Vienna.

La *community* di Phaidra si è allargata includendo nuovi partner fino alla sua attuale configurazione<sup>5</sup>. A livello italiano, le università Ca' Foscari e Luav di Venezia hanno aderito al sistema e dal 2014 pubblicano le loro collezioni digitali nell'installazione di Phaidra a Padova.

Il Sistema Bibliotecario, sviluppando la prospettiva GLAM (Galleries, Libraries, Archives and Museums), ha inoltre potuto instaurare una collaborazione con i musei e gli archivi che, insieme alle biblioteche, sono deputati alla tutela e valorizzazione del patrimonio culturale. Possono archiviare e gestire i loro oggetti digitali in Phaidra, non solo gli enti, ma anche bibliotecari, conservatori di museo, archivisti nonché ricercatori e docenti. Questi ultimi usano la piattaforma come ambiente di deposito, organizzazione e condivisione dei propri materiali<sup>6</sup>.

<sup>3</sup> <https://phaidra.cab.unipd.it/>.

<sup>4</sup> <https://phaidra.univie.ac.at/>.

<sup>5</sup> <https://www.phaidra.org/community/phaidra-partners/>.

<sup>6</sup> Le guide relative alla digitalizzazione e all'archiviazione in Phaidra sono accessibili all'indirizzo: <[https://phaidra.cab.unipd.it/help\\_long#guide](https://phaidra.cab.unipd.it/help_long#guide)>.

## Come è oggi Phaidra

La piattaforma ospita ad oggi 400.000 oggetti digitali, quali immagini, documenti di testo, libri e video, per lo più derivanti da digitalizzazioni di originali analogici. Le diverse tipologie di beni culturali rappresentate includono libri antichi, manoscritti, fotografie, tavole didattiche parietali, mappe, oggetti museali, materiale d'archivio, pergamene.

Questa eterogeneità di beni corrisponde ad un'eterogeneità di origini: il patrimonio culturale proviene da biblioteche, musei, archivi, centri e uffici dell'Università di Padova e dalle altre istituzioni cooperanti. L'avvio di collaborazioni con enti e istituzioni culturali ha permesso di arricchire notevolmente il numero di oggetti digitali ospitati nella piattaforma e di creare collezioni virtuali, superando la barriera fisica e amministrativa dovuta alla localizzazione differenziata degli oggetti analogici.

L'organizzazione dei contenuti di Phaidra in collezioni e la creazione di relazioni tra gli oggetti rendono facilmente esplorabile la piattaforma anche al pubblico più ampio (Fig. 1). L'interfaccia web di Phaidra riflette tale scelta, che posiziona le collezioni al centro dell'esperienza di navigazione e ricerca, anche al seguito dei principi per la costruzione di buone collezioni digitali espressi dal National Information Standards Organization (NISO)<sup>7</sup>.



Figura 1. Organizzazione delle collezioni in Phaidra. Da sinistra a destra: Home page, pagina di tutte le collezioni, presentazione di una singola collezione.

<sup>7</sup> <http://framework.niso.org/8.html>.

La complessità del processo che va dal continuo sviluppo della piattaforma alla digitalizzazione, all'archiviazione fino alla conservazione a lungo termine richiede l'intervento di diverse figure professionali: *repository manager*, *data librarian*, tecnici informatici con competenze di programmazione, *web designer*, responsabili di progetto di digitalizzazione, referenti scientifici delle collezioni, esperti nella conservazione dei documenti, consulenti legali, personale amministrativo.

### L'ambiente digitale di Phaidra

Le caratteristiche di Phaidra rispondono alle raccomandazioni dell'OAIS (Open Archival Information System), il modello di riferimento per gli archivi digitali riconosciuto a livello mondiale<sup>8</sup>.

La piattaforma è costruita su Fedora (Flexible Extensible Digital Object Repository Architecture)<sup>9</sup>, software open source dedicato alla costruzione di biblioteche e archivi digitali<sup>10</sup> con supporto nativo ai Linked Data.

Phaidra è dotata di API (Application Programming Interface) che consentono di implementare diverse categorie di servizi quali ricerca, gestione dei contenuti, estrazione e riuso dei metadati. Le API sono pubbliche e liberamente utilizzabili<sup>11</sup>.

La Fig. 2 rappresenta l'ambiente digitale di Phaidra:

- Caricamento (*Ingestion*). Il caricamento degli oggetti digitali avviene per mezzo dell'interfaccia web o dell'applicazione Phaidra Importer<sup>12</sup> da parte dell'utenza autorizzata. È inoltre possibile effettuare caricamenti massivi tramite le API pubbliche.
- Gestione (*Archival*). Gli oggetti digitali con i relativi metadati, e le relazioni che tra essi intercorrono, sono archiviati nella piattaforma Fedora e organizzati secondo la logica applicativa di Phaidra per essere messi a disposizione e ricercabili attraverso l'interfaccia web<sup>13</sup>, disponibile in italiano, inglese e tedesco.
- Interoperabilità (*Dissemination*). L'utilizzo delle API permette il trasferimento delle risorse digitali da Phaidra ad altre piattaforme e l'integrazione di oggetti e collezioni in siti web e mostre virtuali, favorendo la disseminazione dei contenuti. I

<sup>8</sup> Brian Lavoie, *The Open Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition)*, 2014 <<https://www.dpconline.org/docs/technology-watch-reports/1359-dpctw14-02/file>>.

<sup>9</sup> <https://duraspace.org/fedora/>.

<sup>10</sup> Esempi di altri *repository* digitali aderenti al modello OAIS costruiti su Fedora sono il Digital Repository of Ireland <<http://www.dri.ie/>> e DANS - Electronic Archiving SYstem (EASY) l'archivio del Netherlands institute for permanent access to digital research resources <<https://dans.knaw.nl>>.

<sup>11</sup> <https://github.com/phaidra/phaidra-api>.

<sup>12</sup> <https://github.com/phaidra/java-phaidra-importer>.

<sup>13</sup> La ricchezza dei metadati, la presenza di strumenti come i visualizzatori di libri (*Book Viewer*), di immagini ad alta risoluzione (*Image Viewer*) e di collezioni (*Collection Viewer*) ampliano le possibilità di fruizione degli oggetti digitali. Le opzioni di ricerca prevedono filtri a faccette, ricerca avanzata e ricerca nel full-text dei documenti.



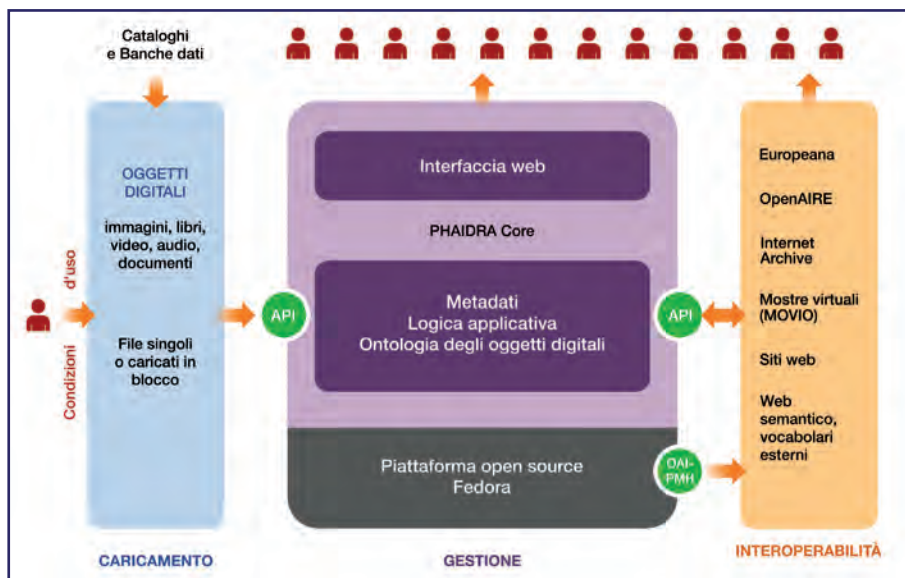


Figura 2. L'ambiente digitale di Phaidra

metadati nello schema Dublin Core vengono esposti tramite il protocollo OAI-PMH (Open Archive Initiative Protocol for Metadata Harvesting) per lo scambio di dati in ambiente distribuito.

### Phaidra come *repository* FAIR

Phaidra è una piattaforma che consente un'ampia fruizione delle sue collezioni digitali nel rispetto dei principi FAIR (Findable, Accessible, Interoperable, Reusable) previsti dalla Commissione Europea per favorire la condivisione e il riuso dei dati nella maniera più aperta possibile. È stata misurata la "FAIRness" di Phaidra verificando la sua conformità ai principi FAIR e come evidenziato in Tab. 1 la maggior parte dei principi delle faccette FAIR viene soddisfatta.

Dato che un *repository* affidabile per la conservazione e il riutilizzo dei dati favorisce lo sviluppo di servizi FAIR, si è intrapreso il processo di certificazione con l'ente certificatore internazionale CoreTrustSeal (CTS)<sup>14</sup>.

La certificazione garantisce l'affidabilità degli archivi, la condivisione a lungo termine dei dati e assicura che siano rispettati tutti gli standard per la conservazione nel tempo. Inoltre Phaidra è registrata nelle directory internazionali di *data repository* Re3Data.org<sup>15</sup> e OpenDOAR<sup>16</sup>.

<sup>14</sup> <https://www.coretrustseal.org/>.

<sup>15</sup> <https://www.re3data.org/>.

<sup>16</sup> <http://v2.sherpa.ac.uk/opensoar/>.

F	FINDABLE (Rintracciabile)	Valutazione
F1	Ai (meta)dati viene assegnato un identificatore univoco e persistente	✓
F2	I dati sono descritti con ricchi metadati	✓
F3	I metadati includono in modo chiaro ed esplicito l'identificatore dei dati descritti	✓
F4	I (meta)dati sono registrati o indicizzati in una risorsa ricercabile	✓
A	ACCESSIBLE (Accessibile)	
A1	I (meta)dati sono recuperabili attraverso il loro identificatore utilizzando un protocollo di comunicazione standardizzato	✓
A1.1	Il protocollo è aperto, gratuito e universalmente implementabile	✓
A1.2	Il protocollo consente una procedura di autenticazione e autorizzazione, ove necessario	✓
A2	I metadati sono accessibili, anche quando i dati non sono più disponibili	X
I	INTEROPERABLE (Interoperabile)	
I1	I (meta)dati utilizzano un linguaggio formale, accessibile, condiviso e ampiamente applicabile per la rappresentazione della conoscenza	✓
I2	I (meta)dati utilizzano i vocabolari che seguono i principi FAIR	X
I3	I (meta)dati includono riferimenti qualificati ad altri (meta)dati	✓
R	REUSABLE (Riusabile)	
R1	I (meta)dati sono riccamente descritti con una pluralità di attributi accurati e rilevanti	✓
R1.1	I (meta)dati vengono rilasciati con una licenza di utilizzo chiara e accessibile	✓
R1.2	Ai (meta)dati è associata una specifica provenienza	✓
R1.3	I (meta)dati soddisfano gli standard rilevanti per specifiche comunità scientifiche	✓

Tabella 1. Misurazione della FAIRness di Phaidra

## Strategie per l'integrazione

Per far convergere in Phaidra risorse provenienti da istituzioni GLAM è stato necessario effettuare analisi dettagliate dei modelli di dati e delle metodologie proprie dei differenti domini di conoscenza, la cui descrizione si basa su schemi di metadati diversificati: UNIMARC<sup>17</sup> per le risorse presenti nel Catalogo del Sistema bibliotecario padovano, standard ICCD<sup>18</sup> per i beni museali, EAD<sup>19</sup> per i documenti di archivio. A tal fine si sono elaborate mappature esaustive tra gli schemi di meta-

<sup>17</sup> <https://www.ifla.org/publications/unimarc-formats-and-related-documentation>.

<sup>18</sup> <http://www.iccd.beniculturali.it/index.php?en/269/metadati>.

<sup>19</sup> <https://www.loc.gov/ead>.

dati sorgente e quello interno alla piattaforma<sup>20</sup> e si sono definite le procedure di esportazione da tali basi di dati a Phaidra<sup>21</sup>.

Per i tre schemi di metadati citati gli esempi sono, rispettivamente, la collezione dei libri antichi e di pregio del polo bibliotecario giuridico<sup>22</sup>, la collezione di statue antiche Mantova Benavides ospitata nel Museo di Scienze archeologiche e d'arte<sup>23</sup> e l'Archivio Metelli conservato nella Biblioteca Centrale di Psicologia<sup>24</sup>.

### Phaidra in dialogo con altri ambienti digitali

In una piattaforma DAMS è di fondamentale importanza la disseminazione e il riuso dei dati in essa archiviati anche al di fuori del sistema.

Oltre alla già citata esposizione dei metadati attraverso le API e il protocollo OAI-PMH, l'interoperabilità si concretizza nel dialogo sempre più stretto con piattaforme internazionali di archiviazione digitale, nell'evoluzione semantica del modello di dati di Phaidra<sup>25</sup> e nell'integrazione tra Phaidra e Movio, il software open source dell'ICCU per la creazione di mostre virtuali<sup>26</sup>.

Utilizzando Movio, il Sistema Bibliotecario ha realizzato varie mostre virtuali che permettono di includere gli oggetti digitali entro veri e propri percorsi narrativi, più ricchi, articolati e interattivi rispetto alla loro presentazione entro una collezione. Fin dalle prime mostre, Phaidra ha rappresentato una miniera da cui attingere immagini, documenti, libri e talvolta intere collezioni. Il Sistema Bibliotecario ha quindi sviluppato un *plug-in* per consentire l'inserimento di oggetti e collezioni di Phaidra in una pagina della mostra virtuale attraverso una semplice maschera integrata in Movio.

Sfruttando l'interoperabilità tra sistemi, i contenuti di Phaidra si sono aperti ulteriormente alla visibilità internazionale, in particolare alcune collezioni della piattaforma sono pubblicate in Europeana e in Internet Archive.

La pubblicazione di collezioni in Europeana è avvenuta nel contesto del progetto europeo Linked Heritage<sup>27</sup>. È in corso di analisi l'invio di altre collezioni a Cultu-

<sup>20</sup> Gli oggetti di Phaidra sono descritti attraverso uno schema di metadati chiamato UWMetadada (Universität Wien Metadada), che estende lo schema LOM (Learning Object Metadata) integrando campi di schemi differenti.

<sup>21</sup> Sono attivi i link da catalogo a Phaidra e viceversa.

<sup>22</sup> <https://phaidra.cab.unipd.it/collections/ladag>.

<sup>23</sup> [https://phaidra.cab.unipd.it/collections/mantova\\_benavides](https://phaidra.cab.unipd.it/collections/mantova_benavides).

<sup>24</sup> <https://phaidra.cab.unipd.it/collections/metelli>.

<sup>25</sup> Anna Bellotto, Cristiana Bettella, *Metadata as Semantic Palimpsests: The Case of PHAIDRA@unipd*, in: *Communications in Computer and Information Science Digital Libraries: Supporting Open Science*, 2019, p. 167-84. DOI:10.1007/978-3-030-11226-4\_14.

<sup>26</sup> <https://www.movio.beniculturali.it/>.

<sup>27</sup> <[https://www.europeana.eu/portal/it/search?f\[DATA\\_PROVIDER\]\[\]=Università%20degli%20Studi%20di%20Padova%20%20University%20of%20Padua&view=grid](https://www.europeana.eu/portal/it/search?f[DATA_PROVIDER][]=Università%20degli%20Studi%20di%20Padova%20%20University%20of%20Padua&view=grid)>. Una collezione di esempio è l'*Iconoteca dei botanici*: <<https://phaidra.cab.unipd.it/o:4714>>, inclusa in "Photography" di Europeana Collections: <<https://www.europeana.eu/portal/it/collections/photography>>.

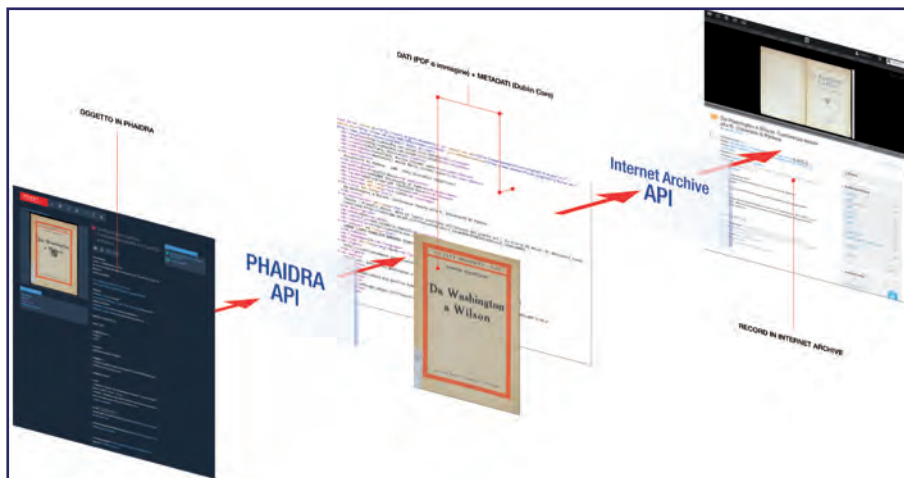


Figura 3. Esportazione di dati da Phaidra a Internet Archive via API

raltalia<sup>28</sup>, il portale del MiBACT, al quale Phaidra aderisce, per l'accesso alle risorse digitali della cultura italiana nonché aggregatore nazionale per Europeana.

Il dialogo tra Phaidra e Internet Archive avviene mediante le rispettive API (Fig. 3), veicolando collezioni digitali e relativi metadati in Internet Archive<sup>29</sup>. L'approccio intrapreso è bidirezionale, in quanto consente di arricchire Phaidra stessa con i file derivati da Internet Archive, quali full text, EPUB e PDF ricercabile, originariamente non disponibili nella piattaforma.

Per migliorare l'interoperabilità della piattaforma, nel contesto del web semantico, un significativo ostacolo è rappresentato dall'assenza di un modello di risorsa basato sullo standard RDF (Resource Description Framework) e dall'impiego di vocabolari non interoperabili il cui spazio di azione è limitato alla realtà locale o a reti preesistenti. Per superare questo limite è in corso di sviluppo un nuovo modello semantico per descrivere le risorse in RDF tramite vocabolari controllati standard e ontologie e l'implementazione di un sistema di gestione di vocabolari basato su SKOS<sup>30</sup>.

### Casi di studio: collezioni digitali e mostre virtuali

Obiettivo di Phaidra è realizzare una base di conoscenza duttile, che fornisca diverse tipologie di risorse per rispondere agli interessi di un pubblico che può accedere attraverso molteplici vie, dai motori di ricerca, dal portale della Biblioteca Digitale di Ateneo, dal catalogo online del Sistema Bibliotecario Padovano e muoversi liberamente con gli strumenti di navigazione.

<sup>28</sup> <http://www.culturaitalia.it/>.

<sup>29</sup> <https://archive.org/details/universityofpadovalibrarysystem>.

<sup>30</sup> <https://www.w3.org/TR/skos-reference/>.

Le collezioni coprono un ampio spettro di soggetti, quali una personalità e il suo tempo, la storia di una scuola con la sua eredità culturale, la raccolta dei materiali relativi a uno studioso. Per restare nell'ambito della vita dell'Ateneo padovano esempi di collezioni sono l'archivio dello psicologo Fabio Metelli<sup>31</sup>, l'eredità culturale della Scuola Medica Padovana<sup>32</sup>, i materiali relativi al botanico Pier Andrea Saccardo<sup>33</sup>. L'accesso ad ogni collezione è facilitato da una presentazione in cui sono evidenziati e direttamente linkabili gli elementi salienti, sia presenti nel *repository* sia collegati da altri siti, in modo da creare un ambiente che favorisca l'esplorazione, invitando all'approfondimento e aprendo a nuovi anche inaspettati percorsi.

A titolo di esempio per le mostre virtuali si ricordano: "Giovanni Marsili, la biblioteca del prefetto dell'Orto botanico di Padova"<sup>34</sup>, una mostra che approfondisce l'ambiente culturale di un erudito del Settecento in Padova e "La Bellezza della biodiversità"<sup>35</sup>, un'illustrazione scientifica delle tavole parietali del Dipartimento di Biologia.

### Conclusioni

Lo sviluppo dell'ambiente digitale aiuta due attività che fanno parte della *mission* di ogni ateneo: da un lato la tutela e conservazione dei beni culturali e dall'altro la loro valorizzazione.

Per la prima è fondamentale guardare alla sicurezza, non possono essere consentite *défaillance* nella conservazione e nel recupero dei dati. Per la seconda è necessario realizzare modalità di accesso e di navigazione che tengano conto delle diverse richieste che un'utenza ampia e differenziata può avanzare. Devono essere disponibili strumenti di navigazione trasparenti e fortemente intuitivi e in grado di integrarsi con realtà differenti.

Tenendo conto di questi requisiti nello sviluppo di Phaidra si è mirato a realizzare un ambiente solido e amichevole, in grado di rispondere velocemente alle diverse esigenze. Per questo Phaidra è una piattaforma dinamica, in continua evoluzione, che dialoga senza difficoltà con realtà, soggetti e servizi esterni. Per la sua realizzazione si sono dimostrati di fondamentale importanza la disponibilità di un codice aperto, la presenza di metadati ricchi, l'esistenza di API pubblicamente accessibili e il fatto che il *repository* sia aderente ai requisiti FAIR. Con queste caratteristiche Phaidra fa parte e valorizza un ecosistema le cui componenti sono beni culturali di natura diversa, e tuttavia testimoni di

<sup>31</sup> <https://phaidra.cab.unipd.it/collections/metelli>.

<sup>32</sup> [https://phaidra.cab.unipd.it/collections/scuola\\_medica\\_padovana](https://phaidra.cab.unipd.it/collections/scuola_medica_padovana).

<sup>33</sup> <https://phaidra.cab.unipd.it/collections/saccardo>.

<sup>34</sup> <https://mostre.cab.unipd.it/marsili/>.

<sup>35</sup> <https://mostre.cab.unipd.it/la-bellezza-della-biodiversita/>.

un'unica grande storia che nell'ambiente digitale può essere raccontata, analizzata e ricomposta, portando così a un valore aggiunto in termini di rappresentazione, accesso e possibilità di larga condivisione dell'eredità culturale.

*Since the end of the Eighties of last century, information technologies have enabled the construction of online library catalogues. The University of Padova has quickly set up a free access catalogue, nowadays with more than 2 million records, including academic libraries and all the cultural institutions of the city. The catalogue has highlighted the historical continuity and the richness of the library heritage, beginning from the origin of the Atheneum in 1222. From the creation of the online catalogue, which participates in the National Library Service, work began on the digitization of book collections and archival and museum collections in order to portray the personalities and developments that pertain to the history of Padova and Sciences. This activity is part of the so-called Third Mission, that is, the set of activities with which universities can directly interact with society in a constant dialogue between the university and citizens.*

*This development required an infrastructure that would guarantee long-term archiving, the description and identification of digital objects, public access from a web interface, with simple and advanced research, terms of use and licensing models for digital objects. In 2010, the University Library System, after careful analysis, chose Phaidra, the repository with FAIR features developed originally by the University of Vienna on Fedora open source software. The platform Phaidra currently hosts 400.000 digital objects within the GLAM framework. During these almost ten years of work, Phaidra has been enriched with useful tools to favour the dissemination and re-use of the deposited resources (for example, Book Viewer, Collection Viewer, Image and API Viewer) and has also constantly researched and activated dialogue with other platforms, such as Europeana and Internet Archive, and software systems, such as Movio, the ICCU open source kit for creating virtual exhibitions.*

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

# Da un oggetto racconta la tua scuola

**Maria Teresa Natale**

ICCU

*“Da un oggetto racconta la tua scuola” è un progetto rivolto alle scuole, promosso dall’Istituto Centrale per il Catalogo Unico delle Biblioteche italiane (ICCU), in collaborazione con la rete Diculther.*

*Si tratta di un primo passo per la nascita di un Museo digitale delle scuole italiane, contenitore di oggetti digitali rappresentativi di alcune scuole italiane e di narrazioni ad essi connessi, frutto del lavoro partecipato delle classi. In questo contributo vengono illustrate le caratteristiche del progetto, le modalità di partecipazione e i primi risultati.*

“Da un oggetto racconta la tua scuola” è un progetto rivolto alle scuole, promosso dal Ministero per i Beni e le Attività Culturali attraverso l’Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche (ICCU), in collaborazione con la rete Diculther<sup>1</sup>. In questo contributo vengono illustrati gli obiettivi del progetto, la piattaforma di raccolta dei dati, i primi risultati raggiunti.

In Italia ci sono circa 20.000 scuole di ogni ordine e grado, pubbliche e private, ognuna delle quali conserva un patrimonio unico di persone, di storie, di edifici, di progetti, di documenti, di archivi, di biblioteche e di oggetti stratificatosi nel tempo: mobili, vecchi quaderni, registri, libri, fotografie, pagelle, strumenti e materiali didattici del passato. Ognuno di questi, da solo e nel complesso, costituisce un piccolo tesoro che racconta una storia: la storia della didattica, la storia della scuola, la storia del territorio, la nostra storia, la nostra memoria.

Poche sono le scuole che hanno valorizzato, anche parzialmente, questi patrimoni eccezionali, allestendo spazi museali stabili o temporanei. Si tratta di “beni invisibili”, quasi mai valorizzati, che non trovano una collocazione adeguata e che spesso sono conservati in magazzini chiusi al pubblico, scatoloni polverosi, scaffali e cassetti inaccessibili, tutt’al più nella stanza del dirigente scolastico.

<sup>1</sup> L’ICCU ha lanciato questo progetto in occasione dell’iniziativa Hackcultura 2019, promossa dalla rete Diculther volta allo sviluppo di progetti digitali da parte degli studenti delle scuole italiane per favorire nei giovani, in un’ottica di “titolarità culturale”, la conoscenza e la “presa in carico” del patrimonio culturale nazionale. Diculther aggrega oltre sessanta organizzazioni tra università, enti di ricerca, scuole, istituti tecnici superiori, istituti di cultura, associazioni e imprese pubbliche e private con l’obiettivo comune di favorire lo sviluppo delle competenze digitali tra gli studenti.



Figura 1.

Oggi, attraverso le tecnologie digitali, non è difficile ipotizzare la costruzione di un immenso museo scolastico digitale che raccolga al suo interno una selezione di oggetti significativi di ogni scuola: per la loro forma, per la loro complessità, per la storia a cui sono collegati, per le persone che li hanno utilizzati.

Se ogni scuola italiana, in un solo anno, contribuisse al Museo scolastico digitale con un solo oggetto, in un anno si avrebbero 20.000 oggetti e 20.000 storie, se contribuisse con una selezione di 10 oggetti digitali, nel solo primo anno si avrebbero 200.000 oggetti pari a 200.000 storie e così via. E ogni oggetto/storia sarebbe un unicum.

Chi sarebbero i visitatori di questo museo scolastico digitale? La popolazione scolastica che all'interno di questo immenso museo – se ben organizzato – troverebbe una quantità infinita di materiali su cui impostare nuovi progetti didattici, gli studiosi e appassionati di storia locale, i giornalisti, gli storici, gli storici dell'arte, gli artigiani, le guide turistiche, gli studiosi di pedagogia e così via.

L'ICCU, in occasione dell'iniziativa Hackcultura 2019 lanciata in collaborazione con la rete DiCulther, ha voluto avviare le basi per la nascita di questo Museo scolastico digitale, contenitore di oggetti digitali rappresentativi di alcune scuole italiane e di narrazioni ad essi connessi, frutto del lavoro partecipato delle classi.



La piattaforma di raccolta degli oggetti digitali<sup>2</sup> è stata realizzata con il software MOVIO distribuito dall'ICCU, assai versatile, nato per la costruzione di mostre digitali ma personalizzabile anche per altri progetti grazie ai numerosi strumenti a disposizione, che consentono l'impostazione e la relazione tra schede catalografiche, la creazione di storie, di mappe geolocalizzate e molto altro.

Chi può partecipare a questo progetto? Ogni scuola localizzata in Italia, di ogni ordine e grado, con una o più classi che possono anche proporre storie diverse legate a uno stesso oggetto.

Le classi che intendono partecipare devono selezionare uno o più oggetti che raccontino un aspetto della propria scuola, che siano una componente del patrimonio culturale della scuola stessa e candidare l'oggetto attraverso un modulo online reso disponibile dall'ICCU<sup>3</sup>. Il lavoro da effettuare con la classe consiste nella schedatura dell'oggetto attraverso la compilazione di una serie di metadati descrittivi (alcuni obbligatori, altri consigliati, altri discrezionali) e nella realizzazione di una narrazione fatta da una successione di testi, audio, video, gallerie fotografiche, frutto di un'attività di ricerca ed elaborazione creativa da parte degli studenti, con il coordinamento degli insegnanti.

Possono essere candidati al Museo digitale oggetti di ogni tipo: da opere d'arte, strumenti scientifici, strumenti musicali, strumenti didattici, oggetti di uso comune, oggetti d'arredo, vestiario, mobili, cimeli, premi ecc., a fotografie, materiale grafico, documenti d'archivio, stampati, registrazioni sonore, filmati storici.

Non vi sono vincoli, ad eccezione di eventuali diritti legati all'oggetto candidato o alle risorse digitali prodotte. La scuola e, per conseguenza, la classe devono essere consapevoli che consegnando i contenuti prodotti al Museo scolastico digitale, li rilasciano con una licenza aperta per favorirne la diffusione e la riusabilità. In particolare, i testi vengono rilasciati con licenza CC BY, le risorse digitali prodotte vengono rilasciate con una licenza CC, a scelta della scuola, per favorirne il riuso.

Il valore didattico del progetto è assai ampio e in linea con le competenze europee: dall'utilizzo del patrimonio lessicale ed espressivo della lingua italiana per raccontare gli oggetti e le storie ad essi associati, alle competenze matematiche per raccontare le storie legate a oggetti dalla valenza scientifica e tecnica; dall'uso di strumenti di comunicazione visiva e multimediale per produrre immagini, audio, video, gallerie fotografiche all'apprendimento delle norme relative all'uso e riuso delle risorse digitali e all'applicazione di licenze adeguate, dalla partecipazione attiva alle attività portando il proprio contributo personale in un lavoro di gruppo al reperimento di informazioni da fonti diverse, dal riconoscimento del valore e delle potenzialità dei beni artistici e ambientali, per una loro corretta fruizione e valoriz-

<sup>2</sup> <https://movio.beniculturali.it/iccu/daunoggettoraccontalatuascuola/>.

<sup>3</sup> <https://goo.gl/forms/OHYSmWDaKcC2ivBE>.

zazione alla capacità di riconoscere aspetti geografici, ecologici, territoriali dell'ambiente naturale e antropico e le loro connessioni con le strutture demografiche, economiche, sociali.

Subito dopo l'avvio l'iniziativa è stata presentata in sedi diverse: presso l'Università Roma Tre nel corso di un evento organizzato in collaborazione con il MusEd, in un webinar organizzato in collaborazione con l'Associazione Flipnet e naturalmente attraverso i diversi canali ICCU e Diculther.

Ma quali sono stati i risultati in questi primi mesi di avvio del progetto? Sette le scuole che hanno partecipato da quattro regioni (Campania, Lazio, Puglia, Toscana), 5 scuole primarie, una scuola secondaria di primo grado, un istituto di istruzione superiore. Gli oggetti candidati sono stati di tipologie diverse.

La primaria Filippo Corridoni di Bari ha ricostruito la storia della scuola attraverso lo studio di carte topografiche; la primaria Niccolò Piccinni, anch'essa di Bari, ha candidato una vecchia lavagna quale protagonista del racconto della storia dell'edificio scolastico; la primaria di Campi Bisenzio (FI) ha candidato una lavagna e un pallottoliere. La classe ha espresso la creatività attraverso la tecnica del fumetto per raccontare le storie di questi due oggetti.

La Scuola primaria Ferraboschi di Calcinato (BR) ha raccontato l'evoluzione dell'edificio scolastico a partire da una tempera su cartoncino della seconda metà del XIX secolo. Molto originale la proposta di una classe della Scuola primaria G. Parodi Delfino di Colleferro (FR) che ha narrato la storia di un'antica statua rinvenuta negli anni Cinquanta nella Valle Macerina e ritrovata casualmente nel 2003 nei depositi della scuola.

Tutti gli alunni delle classi seconde della Scuola secondaria di primo grado Simonetta Salacone, ubicata in un quartiere multiculturale della periferia orientale di Roma, hanno realizzato una vera e propria mostra digitale dal titolo "Parole e immagini: muri o ponti?" a partire dall'opera d'arte "La difesa della razza" facente parte di una collezione di opere afferenti alla mostra permanente "L'arte della guerra", curata dal maestro Bruno Canova. Attraverso questo contributo la Scuola ha voluto dimostrare come, se anche nel Ventennio la scuola fosse stata strumento di propaganda velenosa, oggi essa può essere «un ponte di riflessività, dialogo, incontro... non di verità indissolubili ma di domande alle quali ne seguiranno altre che vivificheranno la Memoria di ciò che è stato» e come «in questo viaggio della conoscenza gli alunni si facciano promotori di senso».

Infine, la storia dell'istituto Alfonso Casanova di Napoli è stata raccontata attraverso l'analisi di un album fotografico, databile al 1928, contenente 30 positivi in bianco e nero aventi per soggetto il "R. Istituto A. Casanova": aule, laboratori, prodotti delle scuole, macchinari, strumenti, beni e arredi, fotografie di gruppo e di spazi architettonici della scuola all'interno della sua sede monumentale nel convento di San Domenico Maggiore.

Questi primi contributi sono stati un'ottima base di partenza per mettere a punto il portale, nel quale è possibile effettuare ricerche per denominazione della scuola, tipologia, regione, personaggi, mostre, storie, geolocalizzazione.

Questo è solo un inizio e con l'avvio del nuovo anno scolastico verrà avviata una nuova campagna di adesione al Museo scolastico digitale<sup>4</sup>.

*“Da un oggetto racconta la tua scuola” (From an object tell your school) is a project targeted to schools, promoted by the Union Catalogue of Italian Libraries (ICCU), in collaboration with the Diculther network. It represents a first step for the building of a Digital Museum of Italian schools. This digital museum should include a selection of digital objects sent by Italian schools, together with stories connected to them. This contribution illustrates the features of the project, the guidelines to participate, and the first results.*

<sup>4</sup> Sulla piattaforma sono disponibili il modulo di adesione, l'elenco degli oggetti candidabili, le linee guida per l'elaborazione delle schede e delle storie. Per la comunicazione sono stati ideati gli hashtag: #daunoggettoraccontalatua scuola e #museodigitaledellescuoleitaliane. Per qualsiasi informazione, scrivere a: [musd.iccu@gmail.com](mailto:musd.iccu@gmail.com).

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

## Segnalazioni



# La Fototeca Tifernate On Line

**Alba Ghelli**

*Responsabile Biblioteca "Carducci" – Comune di Città di Castello*

La Fototeca Tifernate On Line<sup>1</sup>, popolarmente chiamata "Archiphoto", è stata costituita il 23 marzo 2001, con il sostegno del Comune, del Lions Club e della Fondazione Cassa di Risparmio di Città di Castello e dell'Opera Pia "Officina Operaia G. O. Bufalini". Da allora viene gestita dal Centro Fotografico Tifernate (CFT) a nome degli enti promotori e a beneficio dell'intera comunità. È sorta sullo slancio delle tante mostre retrospettive curate da Alvaro Tacchini e allestite dal CFT dal 1983 al 1996 ("Ieri, il Tevere", "L'industria

tipografica a Città di Castello", "Il carnevale tifernate", "La storia del movimento dei lavoratori nell'Alta Valle del Tevere", "Le trasformazioni del paesaggio urbano", "L'Alta Valle del Tevere in cartolina", "La Scuola Operaia G. O. Bufalini", "Il fascismo e la Liberazione a Città di Castello"). Dopo la fondazione della Fototeca, Alvaro Tacchini ha realizzato altre mostre fotografiche e documentarie retrospettive, sempre con l'allestimento del CFT: "Enrico Hartmann fotografo nell'Altotevere", "Il fascismo a Città di Castello", "L'Alta Valle del Tevere e la Grande



Figura 1. *Fausto Coppi a Città di Castello nel 1954* (Fototeca Tifernate On Line, Archivio Giuseppe Tacchini)

<sup>1</sup> <http://www.archiphoto.it/new/it/main.php>.

Guerra”, “Il Risorgimento a Città di Castello”, “La Fattoria Autonoma Tabacchi nei suoi 100 anni”, “Il centenario della Pinacoteca Comunale di Città di Castello” e la serie “Chi èno? Aiutateci a riconoscerli”.

La ricerca di documentazione fotografica da parte di Tacchini ha permesso di raccogliere e rendere pubblico un ingente patrimonio di immagini, in gran parte pubblicate nei suoi libri di carattere storico.

Tale patrimonio è riconducibile a due fonti principali: l’archivio dello Studio Fotografico Giuseppe Tacchini e l’archivio del Centro Fotografico Tifernate, dove Alvaro Tacchini (nipote di Giuseppe, ma senza rapporti commerciali con il suo Studio Fotografico) ha raccolto le immagini reperite in fondi privati nel suo prolungato e capillare lavoro di ricerca storica. Entrambi gli archivi nel 2003 sono stati dichiarati di notevole interesse storico dalla Soprintendenza Archivistica per l’Umbria.

Giuseppe Tacchini continuò lo Studio Fotografico avviato nel 1927 da Francesco Pais

(1892-1959). Ebbe un rapporto privilegiato e continuo con i principali committenti pubblici (Comune, Ospedale, Opere Pie), con gli istituti religiosi, con le scuole, con le aziende più importanti (su tutte la Fattoria Autonoma Tabacchi e la “Francesco Nardi & Figli”) e con le più autorevoli associazioni cittadine. In virtù del rapporto di fiducia con una clientela tanto vasta, l’archivio dello Studio rappresenta uno straordinario patrimonio per la memoria storica di Città di Castello e del suo territorio. Un patrimonio che Giuseppe Tacchini incrementò documentando la cronaca locale, il paesaggio urbano e rurale, gli eventi sociali e i beni artistici e culturali ben oltre le esigenze delle commesse di lavoro. Inoltre catalogò con cura le immagini prodotte, così che tutto il lavoro svolto dallo Studio dalla fine degli anni Quaranta del secolo scorso è facilmente rintracciabile per committente, periodo storico e argomento. Nel contempo avviò un’estesa ricerca delle fotografie d’epoca conservate nei fondi d’archivio famigliari, riproducendole in bianco e nero. Ciò



Figura 2. Enzo Ferrari a Monza, fotografato da un pilota di Città di Castello (Fototeca Tifernate On Line, Archivio CFT, Fondo Minciotti)



Figura 3. *Bambino su autovettura-giocattolo militarizzata, fine anni '30* (Fototeca Tifernate On Line, Archivio CFT, Fondo Riccittelli)

gli permise di rintracciare una cospicua quantità di immagini prodotte dai fotografi tifernati tra '800 e primo '900, soprattutto Aristide Villosi ed Enrico Hartmann. Tale lavoro trovò compimento nella mostra retrospettiva "Città di Castello ieri 1860-1950", allestita nell'ottobre 1977 per il 50° anno di attività dello Studio. Dopo la morte di Giuseppe Tacchini, avvenuta il 10 maggio 1980 a 65 anni di età, lo Studio Fotografico fu portato avanti fino al 2009 dal suo collaboratore Franco Ballini.

L'altra considerevole fonte di immagini della Fototeca Tifernate è costituita dal lavoro di ricerca di Alvaro Tacchini – insegnante, storico e ispettore onorario agli archivi per l'Umbria dal 1995 –, avviato nel 1983 e concretizzatosi nelle iniziative espositive ed editoriali sopra ricordate. La varietà dei temi affrontati ha reso possibile raccogliere fotografie conservate da famiglie altotiberine di ogni ceto sociale, i cui membri hanno recitato un ruolo in eventi di carattere politico, sindacale, religioso, culturale, economico, commerciale, sportivo, associativo e via

dicendo. Una miriade di fondi di proprietà di privati che generosamente hanno permesso la duplicazione delle loro immagini proprio perché essa rendeva possibile una ricostruzione storica della storia locale il più possibile esauriente e condivisa. Anche per questo la Fototeca ha potuto affermarsi come strumento di conservazione e di valorizzazione della Memoria storica locale: non solo è percepita come prezioso e imprescindibile ausilio dagli studiosi di storia contemporanea, ma soprattutto è sentita come "propria" dall'intera comunità altotiberina. In virtù di questa continua interazione e della facile fruizione online delle immagini, la popolazione continua a offrire a piene mani altre fotografie da inserire nell'archivio. Fino a qualche anno fa le fotografie dei fondi privati venivano duplicate in bianco e nero (esattamente come fece Giuseppe Tacchini nella sua attività di ricerca di foto d'epoca) e poi restituite ai proprietari; successivamente si è preso a scansionarle, la qual cosa ha permesso di realizzare duplicazioni a colori e di maggiore qualità.





Figura 4. *Macelleria di Città di Castello nel 1947* (Fototeca Tifernate On Line, Archivio Giuseppe Tacchini)

Dal 2001 al 2017 sono state inserite nella Fototeca, debitamente catalogate da archivisti, oltre 10.200 immagini. Questo lavoro altamente professionale è stato finanziato per lo più con risorse raccolte localmente. Circa 4.000 immagini sono state immesse e catalogate tra il 2004 e il 2005 con un finanziamento del Ministero dei Beni Culturali sotto l'egida della Soprintendenza Archivistica per l'Umbria.

Negli ultimi anni da un lato è emersa la necessità di un aggiornamento delle tecnologie informatiche; dall'altro si è prospettata la possibilità di inserire nella Fototeca una nuova ingente quantità di immagini. Esse continuano ad essere il frutto della ricerca di ulteriori fondi privati, una fonte che si sta dimostrando inesauribile. Nel contempo si è reso disponibile l'immenso patrimonio di fotografie dello Studio Fotografico Giuseppe Tacchini, donato dalla famiglia alla



Figura 5. *Giovani contadini durante la mietitura, anni '20* (Fototeca Tifernate On Line, Archivio CFT, Fondo Marinelli)

Biblioteca Comunale di Città di Castello nel febbraio 2010. Ciò permetterà l'inserimento nella Fototeca della documentazione prodotta dallo Studio fino a tutti gli anni '80 del '900. Finora è stata solo parzialmente selezionata, privilegiando naturalmente le fotografie più datate. Invece comincia a porsi la necessità di rendere fruibili anche quelle del passato più recente (gli anni '60, '70 e '80 del secolo scorso).

È questo un periodo, tra l'altro, che vede l'entrata in scena di una nuova generazione di fotografi che, con la loro attività amatoriale, hanno documentato la realtà sociale in ogni suo aspetto e con una sensibilità non condizionata da risvolti commerciali. Si tratta di un ulteriore fonte di "approvvigionamento" di fotografie che a Città di Castello risulta alquanto agevole, dal momento che diversi di questi fotoamatori gravitano intorno al Centro Fotografico Tifernate e, quindi, alla Fototeca. Dal 2017 l'archivio della Fototeca è gestito on-

line dal Comune di Città di Castello in collaborazione con il CFT, che ne ha ceduto la proprietà al Comune ma mantiene il ruolo di direzione culturale, curandone la promozione, la divulgazione con iniziative espositive ed editoriali e l'incremento, individuando i nuovi fondi privati da acquisire.

È in corso un progetto di digitalizzazione di circa 10.000 nuovi documenti fotografici che saranno inseriti nel database esistente. Tutto il materiale sarà poi pubblicato nella teca digitale dell'ICCU e nel portale Internet Culturale, con un accordo già sottoscritto tra Comune di Città di Castello ed ICCU.

Il progetto rientra tra le azioni previste da "Agenda Urbana Città di Castello ALTOTEVERE SMART – il motore dell'Umbria – POR FESR POR FSE 2014-2020. Asse prioritario ot6 – azione 6.4.1 intervento 1.5\_6.4.1 – azione integrata 1.6. servizio di digitalizzazione dell'archivio fotografico "fototeca digitale online".

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

# Patrimonio culturale: reale e virtuale

Maria Teresa Natale  
ICCU

“**U**ncommon Culture”, rivista in lingua inglese per la quale più volte l’ICCU ha collaborato, è un periodico di respiro europeo che ha l’obiettivo di raccontare esperienze, progetti e collezioni delle istituzioni culturali europee, dedicando ogni fascicolo a un tema specifico. Il numero 1/2 del 2018<sup>1</sup>, dedicato al patrimonio culturale reale e virtuale (*Cultural Heritage, Real & Virtual*) è stato pubblicato in occasione dell’Anno europeo del patrimonio culturale.

La prefazione del fascicolo è stata scritta da un membro del Parlamento europeo, Tadeusz Zwiefka, che si è soffermato proprio sull’importanza della cultura.

I saggi principali sono stati redatti da autori europei e statunitensi i quali hanno trattato la tematica del patrimonio culturale reale e digitale, spesso mettendolo in relazione.

Aprè la serie di saggi di approfondimento, l’articolo di Murtha Baca e Marissa Clifford (UCLA), “Developing a Digital Collaborative Research Environment: the Getty Scholars’ Workspace<sup>®</sup>”, nel quale si tratta della nascita e dell’evoluzione a partire dal 2009 del programma Digital Art History (DAH), condotto dal Getty Research Institute (GRI). Grazie a un’esperienza decennale nel campo della digitalizzazione il GRI ha potuto innovare il suo programma ponendo maggior enfasi sulla pianificazione progettuale e il *project management*, su una maggior consapevolezza delle criticità e delle sfide legate alla pubblicazione online di risorse digitali umanistiche, sulla prioritizzazione dell’analisi critica dei legami tra analogico e digitale anche in relazione alle

borse di studio nelle scienze umanistiche.

Nell’articolo “Is There a Future for Library and Information Work?” il professor Tomas D. Wilson (University of Sheffield) ci parla delle profezie sul futuro dei bibliotecari e dei professionisti dell’informazione. Alcune, come le previsioni di Lancaster sulla società senza carta (1978), hanno fallito completamente, dal momento che il consumo di carta è attualmente maggiore che nel passato e non mostra segni di vacillamento nonostante il declino dell’industria dei giornali; altre, come lo “scenario del giorno del giudizio” di Lewis (1980), forse, erano un po’ precoci nella loro stima della “morte delle biblioteche”.

I libri rari sono invece il tema trattato da Edward J. Valauskas (già Chicago Botanical Garden) nell’articolo “Pedagogical Strategies for Special Collections”. Esperto del tema, egli afferma che i libri rari stimolano i lettori a indagare sul loro significato, in un contesto storico più ampio. Troppo spesso questi libri, a causa del loro valore e della loro scarsità, sono accessibili solo a utenti referenziati, mentre rimangono inaccessibili per il pubblico generico che spesso li può solo ammirare in vetrina in occasione di mostre tematiche.

Trilce Navarrete (Erasmus University in Rotterdam), nel contributo “On the Economics of Physical and Digital Collections in Museums” spiega come i musei siano stati oggetto di indagine da parte degli economisti in termini di creazione di posti di lavoro, attrazione del turismo, coinvolgimento delle imprese. Nell’articolo si propone un modello organizzativo per valutare le condizioni di mer-

<sup>1</sup> <https://uncommonculture.org/ojs/index.php/UC/issue/view/603/showToc>.

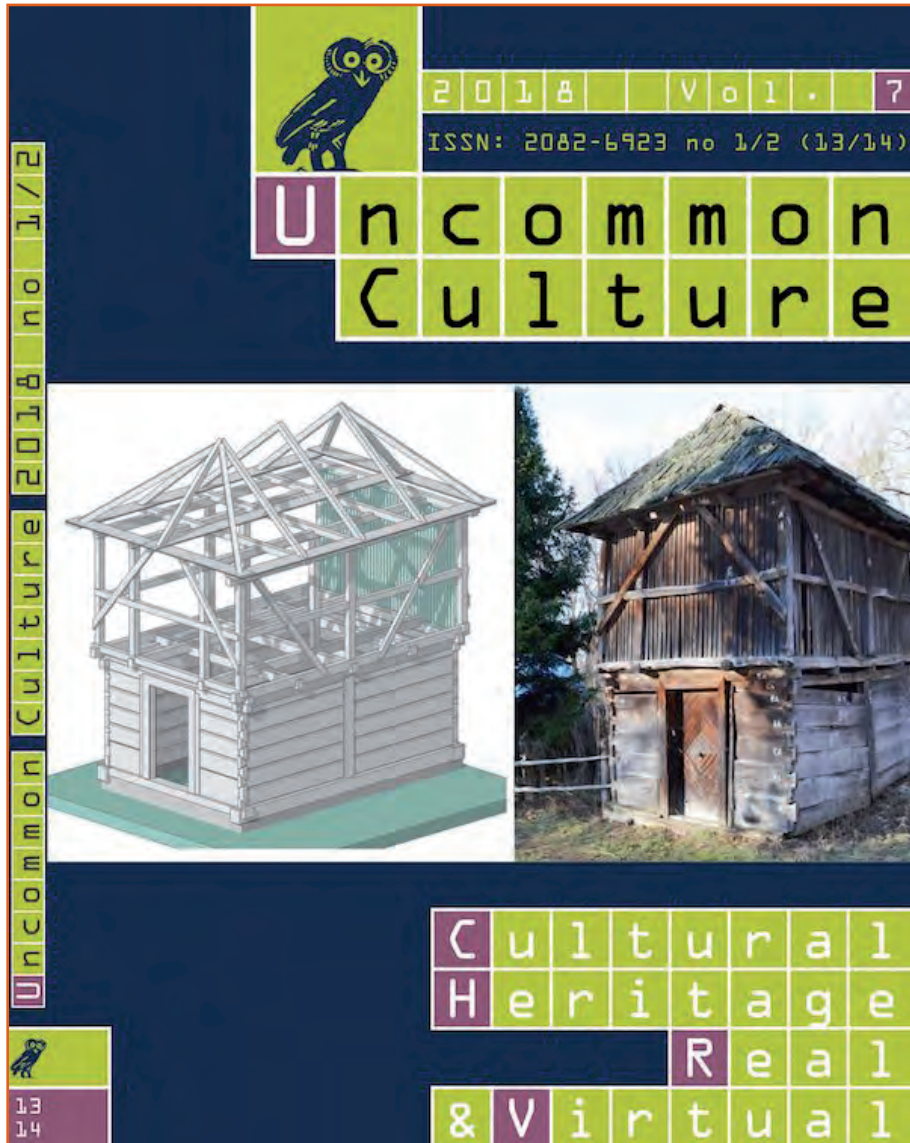


Figura 1.

cato in cui i musei operano, determinate dai costi e dall'uso delle tecnologie, le quali definiscono la concentrazione dei mercati e la differenziazione dei prodotti, nonché la capacità di innovare.

Werner Schweibenz (MusIs) e Roberto Scopigno (CNR-ISTI), nell'articolo "Documenting Past Exhibitions: Why and How Informa-

tion Technology Could Help to Preserve Dismantled Shows" riflettono sull'importanza di documentare digitalmente le mostre concluse. Queste, essendo frutto di un'interpretazione, riflettono un tempo e un luogo e quindi sono di estremo interesse per la ricerca futura. Le tecnologie dell'informazione possono essere sfruttate per conservare gli allestimenti museali.

La ricerca nelle discipline delle scienze sociali si sta focalizzando sugli aspetti dell'apprendimento nell'era digitale, in particolar modo rispetto alla generazione dei nativi digitali. Numerosi sono gli interrogativi in ambito museale: come fornire a giovani tecnologicamente sempre più abili, esperienze efficaci di apprendimento nei musei? Verso quali piattaforme e tecnologie i musei si devono indirizzare? Allo stesso tempo, il mercato dei giochi elettronici è in rapida crescita nel settore dell'edutainment, dimostrando la sua efficacia nei settori della didattica, della formazione, della ricerca e della salute. Un pool di esperti e ricercatori ucraini, nell'articolo "Gamification for Education of the Digitally Native Generation by Means of Virtual Reality, Augmented Reality, Machine Learning, and Brain-Computing Interfaces in Museums" si sofferma sul tema dei nativi digitali nei musei in relazione a interattività, dispositivi da indossare, realtà virtuale e realtà aumentata, soffermandosi su casi di studio aventi per oggetto la didattica nei musei.

Nella sezione dedicata ai progetti Monika Hagedorn-Saupe e Arlene Peukert (Stiftung Preußischer Kulturbesitz SPK) analizzano alcuni servizi sviluppati nell'ambito del progetto Europea Space per favorire il riuso creativo del patrimonio culturale digitale. L'app Blinkster consente ai musei di fornire al pubblico informazioni aggiuntive sulle mostre in modo più efficace delle tradizionali audioguide. Il prototipo "Pop-up museum" consente ai visitatori di utilizzare i propri smartphone per interagire con i contenuti delle mostre attraverso un Content Management System di facile utilizzo.

Come in ogni fascicolo monografico segue poi una rassegna di attività e progetti portati avanti nei singoli paesi europei.

Tra i tanti segnaliamo il contributo di Jordi Pons i Busquet, direttore del Museo del

Cinema di Girona, che in "Life Models Theatricals: Sources of Cultural Heritage for Digital Storytelling" si sofferma sulle collezioni speciali acquisite dall'istituto da lui diretto e in particolare sulle attività di intrattenimento del pubblico in forma di sessioni basate sulla lanterna magica, tecnologia molto popolare tra il 1870 e il 1914. Quest'antenata dell'industria cinematografica può essere riportata in auge grazie alla digitalizzazione.

Oggi sono anche attive istituzioni che pubblicano oggetti online ma non possiedono materialmente le collezioni oppure conservano solo una minima parte di oggetti reali e sfruttano tecnologie digitali aggiuntive per didattica e mostre. Tema questo trattato dalla polacca Anna Kompanowska (Contemporary Art Center a Torun) nel suo articolo "The Non-Existent Object: An Inspiring Technology".

Il giovane studente tedesco Joshua Ramon Enslin illustra la nuova funzione di un software per la gestione delle collezioni museali nell'articolo "Grasping Historical People's Relationships: Let the Objects Speak" ponendo l'attenzione sui risultati ottenuti attraverso la ricerca per nomi anche in relazione ad altri nomi ed eventi.

Fedor Boytsov e Iana Boytsova (PO Pixelated Realities company) nel contributo "The Emergency Recording and Public Crowdsourcing of Materials for Cultural Heritage Digitization in Developing Postindustrial Regions of the Historical Cities" si soffermano sullo stato di conservazione dei monumenti a Odessa. Nell'ultimo decennio più di 130 monumenti antichi sono stati demoliti o ricostruiti perdendo completamente il loro valore storico e/o artistico. Con l'aiuto di un fotografo non professionista, in breve tempo sono stati in grado di raccogliere una ricca documentazione su uno degli edifici demoliti, per poi ricostruirlo fornendo un modello 3D.

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





## **Pubblicazioni**

Futuro delle memorie digitali e patrimonio culturale: atti del convegno internazionale, Firenze, 16-17 ottobre 2003, a cura di Vittoria Tola e Cecilia Castellani, Roma, ICCU (2004)  
ISBN 978-88-7107-110-7

Memorie digitali: rischi ed emergenze, a cura di Alessandra Ruggiero e Vittoria Tola, Roma, ICCU (2005)  
ISBN 978-88-7107-111-5

Linee guida per registrazioni d'autorità e di rinvio, Roma, ICCU (2005)  
ISBN 978-88-7107-105-0

Linee guida per la digitalizzazione del materiale fotografico, a cura del Gruppo di lavoro sulla digitalizzazione del materiale fotografico, Roma, ICCU (2005)  
ISBN 978-88-7107-112-3

EAD: Descrizione archivistica codificata: dizionario dei marcatori, a cura di Giovanni Michetti, Roma, ICCU (2005)  
ISBN 978-88-7107-115-8

MAG: metadati amministrativi e gestionali: manuale utente, a cura di Elena Pierazzo, Roma, ICCU (2006)  
ISBN 978-88-7107-117-4

Linee guida per la digitalizzazione del materiale cartografico, a cura del Gruppo di lavoro sulla digitalizzazione del materiale cartografico, Roma, ICCU (2006)  
ISBN 978-88-7107-116-6

Inter omnes: contributo allo studio delle marche dei tipografi e degli editori italiani del XVI secolo, Roma, ICCU (2006)  
ISBN 978-88-7107-114-X

Le edizioni italiane del XVI secolo: censimento nazionale, Vol. 5: D, Roma, ICCU (2006)  
ISBN 978-88-7107-113-1

Linee guida per la digitalizzazione di bandi, manifesti e fogli volanti, a cura del Gruppo di lavoro sulla digitalizzazione di bandi, manifesti e fogli volanti, Roma, ICCU (2006)  
ISBN 978-88-7107-118-2

Miniature e decorazioni dei manoscritti, a cura di Emilia Ambra, Angela Dillon Bussi, Massimo Menna, Roma, ICCU (2006)  
ISBN 978-88-7107-119-0

Le edizioni italiane del XVI secolo: censimento nazionale, Vol. 6: E-F, Roma, ICCU (2007)  
ISBN 978-88-7107-125-1

Il libro italiano del XVI secolo: conferme e novità in EDIT16: atti della giornata di studio, a cura di Rosaria Maria Servello, Roma, ICCU (2007)  
ISBN 978-88-7107-124-4

OAIS: Sistema informativo aperto per l'archiviazione, a cura di Giovanni Michetti, Roma, ICCU (2007)  
ISBN 978-88-7107-126-8

ISBD(CR): International Standard Bibliographic Description for Serials and Other Continuing Resources: revisione dell'ISBD(S), Roma, ICCU (2007)  
ISBN 978-88-7107-104-2

Regole italiane di catalogazione: REICAT, a cura della Commissione permanente per la revisione delle regole italiane di catalogazione, Roma, ICCU (2009)  
ISBN 978-88-7107-127-5

Requisiti funzionali per i dati di autorità: un modello concettuale, a cura di Glenn E. Patton, Roma, ICCU (2010)  
ISBN 978-88-7107-129-9

International Standard Bibliographic Description (ISBD), IFLA, Roma, ICCU (2010)  
ISBN 978-88-7107-128-2

Guida alla catalogazione SBN. Musica: musica e libretti a stampa, registrazioni sonore, video e risorse elettroniche musicali, Roma, ICCU (2012)  
ISBN 978-88-7107-130-5

ISBD: International Standard Bibliographic Description, edizione consolidata, Roma, ICCU (2012)  
ISBN 978-88-7107-131-2

Titolo Uniforme Musicale: norme per la redazione, Roma, ICCU (2014)  
ISBN 978-88-7107-133-6

Per l'acquisto delle edizioni ICCU rivolgersi a:

*ICCU Ufficio vendite*  
Viale Castro Pretorio, 105  
00185 Roma  
T +39 06 49210430  
F +39 06 4959302  
e-mail: [venditapubbl@iccu.sbn.it](mailto:venditapubbl@iccu.sbn.it)  
<https://www.iccu.sbn.it>





# ICCU

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

Copyright © ICCU - Roma

Finito di stampare nel mese di giugno 2019  
a cura dell'Istituto Poligrafico e Zecca dello Stato S.p.A. - Roma



ICCU

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