

Il Seminario italo-spagnolo di Biblioteconomia e Documentazione

Roma, 4-5 novembre 2022

a cura di

Andrea Capaccioni e Paola Castellucci

Ledizioni

Questo libro è stato pubblicato grazie al contributo della Società italiana di scienze bibliografiche e biblioteconomiche (SISBB) e con i fondi di ricerca (A. Capaccioni) del Dipartimento di Lettere - Lingue, Letterature e Civiltà antiche e moderne dell'Università di Perugia.

Unless otherwise stated, this work is released under a Attribution-ShareAlike 4.0 International (CC BY-SA 4.0), <https://creativecommons.org/licenses/by-sa/4.0/deed.it>.



2024 Ledizioni LediPublishing
Via Boselli 10, 20136 Milano - Italy
www.ledizioni.it
info@ledizioni.it

Il Seminario italo-spagnolo di Biblioteconomia e Documentazione. Roma, 4-5 novembre 2022, a cura di Andrea Capaccioni e Paola Castellucci
Prima edizione: aprile 2024

ISBN cartaceo: 9791256001149
ISBN ePub: 9791256001156
ISBN PDF Open Access: 9791256001163

Progetto grafico: ufficio grafico Ledizioni
In copertina: immagine creata da Antonella Iacono

Informazioni sul catalogo e sulle ristampe dell'editore: www.ledizioni.it

INDICE

Biblioteconomia e Documentazione in Italia e Spagna: un confronto che continua. Introduzione <i>Andrea Capaccioni, Paola Castellucci</i>	7
--	---

I. BIBLIOTECHE, TECNOLOGIE E INTELLIGENZA ARTIFICIALE

Modelli e pratiche della lettura in biblioteca. Il contesto bibliografico del progetto Reading(&)Machine <i>Maurizio Vivarelli</i>	19
--	----

Intelligenza Artificiale nelle biblioteche. Stato dell'arte ed esperienze di applicazione <i>Rossana Morriello</i>	33
--	----

L'Intelligenza Artificiale ed il problema dell'agency dal punto di vista bibliografico <i>Maurizio Lana</i>	47
---	----

Inteligencia Artificial vs. noticias falsas: las bibliotecas como agente activo frente a la desinformación <i>Sara Martínez-Cardama</i>	61
---	----

II. ATTIVITÀ FORMATIVA E DI RICERCA: MIGLIORARE I RAPPORTI TRA I DUE PAESI

Una propuesta de internacionalización para los programas de Doctorado en Biblioteconomía y Ciencias de la Documentación entre España e Italia <i>José Luis Gonzalo Sánchez-Molero</i>	79
--	----

Iniziative di rafforzamento dell'internazionalizzazione
della didattica in un corso di laurea 87
Fiammetta Sabba

El programa doctorado de Ciencias de la Documentación
de la Universidad Complutense de Madrid 99
Antonio Carpallo Bautista

Studi mediterranei e politiche di internazionalizzazione
tra Spagna e Sardegna. Le scienze del libro
e del documento all'Università di Cagliari 113
Giovanna Granata

III. UN CONFRONTO SUI PROGETTI DI DIGITALIZZAZIONE DEI FONDI STORICI: ESPERIENZE INNOVATIVE, QUESTIONI, NORME

I progetti di digitalizzazione dei fondi storici: temi a confronto 123
Alberto Salarelli, Cristóbal Urbano

Archivo de Revistas Catalanas Antiguas (ARCA):
un caso paradigmático entre los proyectos de digitalización
cooperativa en los que participa la Biblioteca de Catalunya 137
Eugènia Serra

Alphabetic per la didattica e la ricerca 147
Elisabetta Castro

EDL: la piattaforma digitale della Biblioteca Estense
Universitaria di Modena 159
Grazia Maria De Rubeis

BIBLIOTECONOMIA E DOCUMENTAZIONE IN ITALIA E SPAGNA: UN CONFRONTO CHE CONTINUA. INTRODUZIONE

Andrea Capaccioni*, Paola Castellucci**

Creare una rete: le prime tappe

Le relazioni tra il mondo accademico italiano e quello spagnolo nell'ambito della Biblioteconomia sono di lunga data. Molti docenti italiani hanno negli ultimi decenni condiviso attività di didattica e di ricerca con i colleghi spagnoli (Massara, 2019). Lezioni tenute nei rispettivi atenei, partecipazione a convegni e seminari, pubblicazione sulle riviste del settore disciplinare di entrambi i paesi e in alcuni casi collaborazioni nella stesura di volumi e articoli scientifici. Tra i diversi esempi ne scegliamo due. Il primo ci permette di rivolgere un doveroso omaggio a un autorevole studioso recentemente scomparso, José López Yepes (1946-2023), *catedrático emérito* dell'Università Complutense di Madrid e tra i pionieri del nuovo corso delle scienze del libro in Spagna e in America Latina. Nel 2006 lo studioso spagnolo viene invitato dall'Associazione italiana biblioteche come relatore al 53° congresso nazionale tenutosi a Roma nel Centro congressi Europa. Rosa Maiello in una attenta sintesi del congresso definisce “molto stimolante (e incoraggiante)” l'intervento di Yepes (*Bibliotecarios y documentalistas en España. Situación actual y nueva proyección formativa y profesional en el espacio europeo de enseñanza superior*) incentrato sull'esposizione delle tappe attraverso le quali si era giunti in Spagna a elaborare “un quadro unitario del curriculum universitario di bibliotecari e documentalisti” (Maiello, 2007)¹. Lo studioso spagnolo racconta le difficoltà incontrate nell'individuare un “nome” comune per un settore in cui agiscono discipline diverse, Biblioteconomia, Archivistica, Documentazione, scienza dell'informazione, che si occupano di gestione dei

*Università di Perugia.

** La Sapienza Università di Roma.

¹ Alcuni temi affrontati dall'intervento “italiano” di Yepes al 53° Congresso AIB erano stati anticipati in un articolo pubblicato dal *Bollettino AIB* (López Yepes, 2005). Per alcune riflessioni su questo contributo si veda Castellucci (2007).

documenti e dell'informazione. Una scelta ardua che ha tenuto conto anche di alcune "divergenze epistemologiche" tra le diverse aree di studio. Nonostante queste difficoltà in Spagna, a partire dal 2004, «sotto gli auspici dell'agenzia nazionale spagnola per la valutazione e l'accreditamento, i responsabili delle scuole universitarie di Biblioteconomia e Documentazione – sentite le associazioni professionali di categoria e un certo numero di operatori a diversi livelli» erano riusciti a trovare un accordo su un progetto di riforma dei piani di studio in questi ambiti disciplinari (Maiello, 2007). La proposta spagnola mirava a unificare i cicli triennale e biennale in un corso di laurea unico denominato Laurea in Informazione e Documentazione. Nonostante i punti di vista divergenti, i bibliotecari, gli archivisti e i documentalisti avevano alla fine condiviso la scelta di mantenere al centro di questa nuova area disciplinare l'attenzione per gli aspetti legati all'organizzazione dell'informazione orientata verso gli utenti.

Il secondo esempio che presentiamo è il contributo *Le condizioni della Biblioteconomia e della Documentazione nell'università spagnola* scritto da Ernest Abadal (Universitat de Barcelona) in occasione del I Seminario nazionale di Biblioteconomia (30 e 31 maggio 2013) organizzato dal Dipartimento di Scienze documentarie, linguistico-filologiche e geografiche della Sapienza Università di Roma con il patrocinio dell'Associazione italiana biblioteche (AIB) e della Società italiana di scienze bibliografiche e biblioteconomiche (SISBB) (Abadal, 2013). Il saggio costituiva per quegli anni lo stato dell'arte più aggiornato a disposizione del lettore italiano sulla formazione e la ricerca in Biblioteconomia e Documentazione nelle università spagnole e al tempo stesso una riflessione approfondita sulle prospettive future di quelle discipline.

Negli anni successivi la collaborazione tra docenti di Biblioteconomia e Documentazione italiani e spagnoli prosegue nel modo che abbiamo descritto dei rapporti bilaterali. Comincia tuttavia a farsi strada l'idea di superare questa prassi e di orientarsi invece verso la creazione di una rete coordinata di contatti. I primi passi in questa nuova direzione sono ben descritti da due professori spagnoli, protagonisti di questa nuova fase, José Luis Gonzalo Sánchez-Molero (Universidad Complutense, Madrid) e Mercedes Caridad Sebastián (Universidad Carlos III, Madrid): «Fue entre 2017 y 2018, en la ciudad de Perugia, durante las estancias de investigación de varios profesores de la Universidad Complutense en la Università degli Studi di Perugia (acogidos por el profesor Andrea Capaccioni), cuando surgió el proyecto de establecer un marco común de colaboración académica y profesional entre España e Italia en el ámbito de la Biblioteconomía y la Documentación. En una posterior estancia del profesor Capaccioni en Madrid se profundizó en este proyecto. En abril de 2018 José Luis Gonzalo, Decano de la Facultad de Ciencias de la Documentación, y el profesor Capaccioni, Delegato del Rettore per i servizi bibliotecari e per l'Open Access, empezaron a trabajar en la tramitación de un

convenio internacional entre la Complutense y varias universidades italianas. Su propósito era colaborar en el desarrollo de actividades conjuntas en nuestra área de conocimiento. Las líneas generales de lo tratado entonces planteaban un acuerdo entre las universidades Complutense, Perugia, Salerno, Roma-La Sapienza, Torino, Firenze y Udine. Entre las actividades conjuntas se proponía ya, la organización de un congreso o seminario anual, a celebrar de manera alternativa entre ambos países (Gonzalo Sánchez-Molero, Caridad Sebastián, 2021, p. 5)».

L'ipotesi di un accordo tra alcune università spagnole (in particolare Complutense e Carlos III di Madrid e Universitat de Barcelona), coordinate da Gonzalo Sánchez-Molero (Complutense), e dalle università italiane di La Sapienza di Roma, Firenze, Salerno, Torino e Perugia comincia a prendere forma nel corso del 2018 attraverso un assiduo scambio di pareri e di documenti. Su proposta di Maurizio Vivarelli dell'Università di Torino si fa strada l'ipotesi dell'organizzazione di due incontri del gruppo di lavoro nel capoluogo piemontese il 13 maggio dell'anno successivo presso l'Ateneo e il Salone internazionale del Libro che per quell'anno aveva indicato la Spagna come nazione ospite.

La prima riunione si è svolta nella Biblioteca storica di Ateneo Arturo Graf². Nel pomeriggio, i membri del gruppo hanno poi partecipato a due tavole rotonde ospitate dal Salone del Libro. Nel puntuale resoconto di Flavia Massara l'incontro viene descritto come «un momento significativo di scambio tra le comunità universitarie dei due Paesi» (Massara 2019, p. 45). Nel corso della prima tavola rotonda, intitolata “Biblioteconomia e culture del libro in Italia e Spagna: un progetto di collaborazione” e introdotta da Andrea Capaccioni, Ernest Abadal (Universitat de Barcelona) ha indicato le tematiche più diffuse in Spagna in ambito biblioteconomico, basandosi su un'analisi condotta su circa 400 articoli scientifici pubblicati da sette importanti riviste spagnole del settore disciplinare. La ricerca ha messo in evidenza, tra i vari aspetti, che tali riviste avevano pubblicato solo il 5% di contributi di autori stranieri e tra questi nessuno italiano. Mauro Guerrini, dell'Università di Firenze, ha posto l'accento sulle competenze della Biblioteconomia contemporanea posta di fronte ai recenti cambiamenti e di come questo stato di cose renda ancora più urgente un'ampia internazionalizzazione della disciplina: in passato l'Italia

² Ai saluti di Gianluca Cuniberti, direttore del Dipartimento di Studi storici dell'Università di Torino, hanno fatto seguito gli interventi per l'Italia di Maurizio Vivarelli, Giovanni Solimine, Mauro Guerrini, Giovanni Di Domenico e Andrea Capaccioni; per la Spagna di Fátima García López, Mercedes Caridad Sebastián, Ernest Abadal, Margarita Pérez Pulido, Benito Rial Costas, José Luis Gonzalo Sánchez-Molero, María Olivera Zaldúa, Antonio Carpallo Bautista e Fátima Martín Escudero (Gonzalo Sánchez-Molero, Caridad Sebastián, 2021, pp. 6-7).

è invece sembrata interessata quasi esclusivamente a un confronto con la Biblioteconomia angloamericana. Giovanni Solimine (La Sapienza Università di Roma), ha evidenziato come, nonostante i rapporti instaurati da almeno vent'anni, tra gli studiosi di Biblioteconomia di Spagna e Italia persistano ancora alcune differenze. In Italia, la disciplina negli ultimi anni si è concentrata sul rapporto tra biblioteca e società e sul ruolo dell'utente, privilegiando un approccio teorico anche se ben radicato nella realtà professionale. In Spagna l'attenzione è stata invece focalizzata sull'informazione e la documentazione e si è registrata una maggiore propensione al lavoro di gruppo. Solimine ha concluso ribadendo la necessità di una collaborazione basata su una conoscenza reciproca tra i due Paesi. È poi intervenuto Benito Rial Costas (Universidad Complutense, Madrid) che ha evidenziato la necessità di una maggiore collaborazione anche tra le discipline che hanno come oggetto di studio il libro e che questa possa procedere in una direzione che vada oltre l'analisi testuale e storica favorendo piuttosto un approccio al libro come "sistema", in modo tale che si possa esaminare il modo in cui esso modifica la nostra percezione della realtà. Anche José Luis Gonzalo Sánchez-Molero (Universidad Complutense, Madrid) ha sottolineato l'importanza di un incontro tra Italia e Spagna in ambito biblioteconomico e auspicato che questa alleanza non coinvolga soltanto i docenti e i ricercatori, ma anche gli studenti e i professionisti del settore. Lo studioso ha poi sottolineato che è necessario lavorare per equiparare i titoli accademici in un'ottica di collaborazione mediterranea e latina.

Il secondo incontro intitolato "Biblioteche pubbliche in Italia e Spagna: uno sguardo comparativo" è stato introdotto da Giovanni Di Domenico (Università di Salerno) e si è concentrato sul confronto tra i servizi, gli strumenti di misurazione e di valutazione delle biblioteche pubbliche italiane e spagnole con interventi di Fátima García (Universidad Carlos III, Madrid), Margarita Pérez Pulido (Universidad de Extremadura, Badajoz), Cecilia Cognigni (Commissione nazionale biblioteche pubbliche Associazione italiana biblioteche), Enzo Borio (presidente Associazione italiana biblioteche-Piemonte), María Olivera Zaldúa (Universidad Complutense, Madrid).

L'incontro di Torino ha costituito l'inizio di un dialogo pubblico tra la Biblioteconomia spagnola e italiana incentrato sulla condivisione delle principali aree di ricerca e delle migliori pratiche professionali. Questi obiettivi, secondo Anna Bilotta, sono stati raggiunti con successo grazie a una nuova iniziativa: il I Seminario hispano-italiano en Biblioteconomía y Documentación tenutosi online il 29 e 30 ottobre 2020 (Bilotta 2022, 61-62). L'evento, la cui idea era stata concepita nel corso degli incontri torinesi, è stato organizzato dalla Facultad de Ciencias de la Documentación dell'Universidad Complutense di Madrid e dal Departamento de Biblioteconomía y Documentación dell'Universidad Carlos III di Madrid con relatori provenienti

dalle università spagnole (Complutense, Carlos III, Barcellona, Salamanca, Estremadura e Valencia) e italiane (Roma La Sapienza, Torino, Firenze, Parma, Perugia, Salerno, Università Cattolica del Sacro Cuore, Milano e Urbino). Tra gli invitati troviamo anche professionisti del mondo bibliotecario ed editoriale, con la rappresentanza di Ediciones Trea, Ediciones Complutense, Editrice Bibliografica, Ledizioni, Società Italiana di Scienze Bibliografiche e Biblioteconomiche, Asociación Española de Bibliografía, Sociedad Española de Documentación e Información Científica³. Il seminario ha offerto una ricca panoramica sullo stato attuale e sulle prospettive delle scienze dell'informazione e della cultura del libro nei due paesi; e ha confermato l'importanza dello scambio, del confronto e della collaborazione tra le comunità accademiche e professionali. Gli atti del I Seminario sono stati curati da José Luis Gonzalo Sánchez-Molero e Mercedes Caridad Sebastián, e sono stati presentati (online) da Fiammetta Sabba (Università di Bologna) e Valentina Sonzini (Università di Firenze) il 1° aprile 2022 all'interno delle attività seminariali della Società italiana di Scienze Bibliografiche e Biblioteconomiche-SISBB⁴.

Per un approccio “meridiano”

Queste prime tappe intraprese con l'intento di creare una rete di contatti tra i due Paesi nell'ambito professionale e di ricerca della Biblioteconomia e della Documentazione ci permettono di esporre alcune considerazioni.

Anche per questi ambiti disciplinari crediamo possano essere accolte le parole espresse da Franco Cassano, un intellettuale italiano, anzi meridionale, ormai un quarto di secolo fa: «Pensiero meridiano vuol dire fondamentalmente [...] restituire al sud l'antica dignità di soggetto del pensiero, interrompere una lunga sequenza in cui esso è stato pensato da altri. Tutto questo non vuol dire indulgenza per il localismo [...]. Al contrario un pensiero meridiano ha il compito di pensare il sud con maggior rigore e durezza, ha il dovere di vedere e combattere *iuxta propria principia* la devastante vendita all'incanto...» (Cassano, 2001b, pp. 3-4).⁵

A partire da tali sollecitazioni, possiamo chiederci se esista una specificità nel modo di intendere le nostre discipline per noi, meridionali d'Europa. E

³ Per consultare il programma completo e scaricare la versione elettronica e gratuita degli atti si veda Universidad Complutense de Madrid, I Seminario hispano-italiano en Biblioteconomía y Documentación, <<https://documentacion.ucm.es/seminario-hispano-italiano>>. Un dettagliato resoconto del seminario si può leggere in Billotta 2020; si veda anche Bernabè (2020) e Bergamini (2022).

⁴ Il video della presentazione è disponibile nella pagina dei Seminari della SISBB: <<https://www.sisbb.it/2022/03/libri-carte-bit-e-dintorni-seminari-sisbb-2022/>>.

⁵ Si veda anche Cassano (2001a).

se davvero condividiamo lo stesso meridiano, dovremmo allora impegnarci a difendere la nostra identità, per non essere “pensati da altri”, per non ritrovarci subalterni al pensiero statunitense e nordeuropeo? E ancora, la nostra area disciplinare rischia forse di “essere pensata da altri” – dall’informatica, dalle discipline storiche, ad esempio – riducendoci a ruoli ancillari? E infine, abbiamo la consapevolezza e la forza per affermare “l’antica dignità di soggetto”? Riusciamo a dire “noi” e a segnare insieme, studiosi spagnoli e italiani, un meridiano da cui far partire il tempo, per darci una tabella di marcia, per le azioni da compiere? Il Seminario ispano-italiano/italo-spagnolo, tenta di dare qualche risposta. E a loro volta, le risposte sono ipotesi. Ipotesi di lavoro per una ricerca condivisa, e insieme comparativa, che necessariamente ha bisogno di un “pensiero lento”, interiorizzato, storicizzato, citando ancora Franco Cassano.

Il Seminario ispano-italiano/italo-spagnolo è stato pensato con una cadenza biennale, seguendo un “pensiero lento”, idoneo a riflettere sullo stato dell’arte e sulle scelte future. Le sedi ospitanti si alternano tra i due Paesi.

L’obiettivo primario del Seminario è favorire una reciprocità dello sguardo da entrambe le sponde del Mediterraneo. Nella *Introducción*, i curatori presentavano infatti il I Seminario come “un proyecto sembrado, cuidado y recogido con paciencia a ambas orillas del Mediterráneo occidental” (Gonzalo Sánchez-Molero, Caridad Sebastián, 2021, p. 5).

Nel II Seminario – di cui vengono qui pubblicati gli Atti – la metafora spaziale viene proposta sin dalla locandina: un globo stilizzato, una struttura reticolare di meridiani e paralleli (con esplicito rimando al simbolo del Mundaneum di Paul Otlet e Henri La Fontaine) in campitura *rojigualda*, come la bandiera spagnola e come lo stendardo giallorosso di Roma, sede ospitante, nei giorni 4 e 5 novembre 2022, presso la Facoltà di Lettere e Filosofia della Sapienza Università di Roma⁶. Nell’autunno del 2024, il percorso invertirà nuovamente la rotta, verso Barcellona, dove si terrà il III Seminario. Continua così la ricerca, seguendo il tracciato di un “pensiero meridiano”. Non si tratta di un’immagine ornamentale e nemmeno di una metafora. Contestualizzare il nostro ambito disciplinare in una precisa geografia fisica significa di per sé rimandare a un approccio metodologico: l’attenzione si concentra su un “sapere situato” all’interno di una più ampia geopolitica della ricerca (Braidotti, 1995). La tradizione bibliografica, biblioteconomica e documentaria, in Spagna e in Italia, viene vista in modo diacronico e comparatista, consapevoli di un panorama epistemologico in rapida trasformazione.

L’idea di un pensiero meridiano negli studi bibliografici, biblioteconomici e

⁶ Il programma dettagliato del II Seminario italo-spagnolo di Biblioteconomia e Documentazione si può consultare qui <<https://www.sisbb.it/2022/10/ii-seminario-italo-spagnolo-di-biblioteconomia-e-documentazione/>>.

documentari si era in profilata ancor prima dell'organizzazione del Seminario ispano-italiano/italo-spagnolo, come sottolineano i curatori nella *Introducción* agli Atti del I Seminario (Gonzalo Sánchez-Molero, Caridad Sebastián, 2021, p. 7).

La linea mediterranea continua nel II Seminario. Così, ad esempio, conclude il contributo José Luis Gonzalo Sánchez-Molero (a p. 77 del presente volume), invitando a intraprendere nuove rotte nel Mare Nostrum: «Los encuentros en Perugia, Madrid, Turín y Roma han permitido constatar que en nuestro particular Mare Nostrum, las Ciencias de la Documentación pueden fijarse nuevos rumbos». Il Mare Nostrum è evidentemente tanto geografico che disciplinare. Si pone come ipotesi investigativa, ed include necessariamente un approccio comparatistico: occorre individuare cosa è “nostro”, occorre definire se, e cosa, ci contraddistingue e cosa ci differenzia come meridionali d'Europa, sia a livello teorico che applicativo. Nel II Seminario sono stati pertanto scelti tre temi di interesse a cui sono state dedicate specifiche sezioni. Le relazioni sono state tenute in spagnolo o in italiano (e non in inglese, quasi a ricordare che “internazionalizzazione” non significa esclusivamente utilizzare la lingua inglese, seguendo il meridiano statunitense).

Nel II Seminario ogni sezione è stata curata da una coppia di studiosi, in rappresentanza dei due Paesi:

- *Biblioteche, tecnologie e Intelligenza Artificiale: studi ed esperienze a confronto* (Maurizio Vivarelli e Mercedes Caridad)
- *Attività formativa e di ricerca: migliorare i rapporti tra i due Paesi* (Paola Castellucci e José Luis Gonzalo Sánchez-Molero)
- *Un confronto sui progetti di digitalizzazione dei fondi storici: esperienze innovative, questioni, norme* (Alberto Salarelli e Cristóbal Urbano)

Per ogni sezione hanno partecipato (in massima parte in presenza), studiosi spagnoli e italiani; è stato accolto anche un contributo di ricercatori brasiliani che non compare nei presenti Atti. L'indicazione è però importante e ci ricorda che l'ipotesi di un meridiano comune potrebbe essere esteso ad altri paesi mediterranei, e ad altri sud del mondo. E anche ad altri “sud”: ossia ad altri temi, ad altre metodologie di ricerca, percepite come subalterne nel nostro panorama disciplinare.

Cogliamo l'occasione degli Atti per ringraziare tutti coloro che hanno partecipato al Seminario e tutti gli autori. Grazie a chi ha avuto la pazienza di collegarsi da remoto. Grazie a chi ha seguito seduto nell'aula Odeion della Sapienza Università di Roma, ricostruzione di un antico teatro greco.

Grazie al Direttore del Dipartimento di Lettere e Culture Moderne della Sapienza Università di Roma, Marco Mancini, per la generosa offerta degli spazi e dell'ospitalità. Va ricordato che il Dipartimento è sede legale

della SISBB, la Società scientifica che rappresenta in Italia la nostra area professionale e di ricerca. Un ringraziamento particolare va quindi ai soci che, come in altre occasioni, hanno partecipato attivamente. Grazie a Mauro Guerrini, membro del comitato scientifico del Seminario. Grazie al Direttivo SISBB 2018-2020, in carica al momento di avvio del Seminario, nelle persone di Marisa Borraccini (Presidente) e Giovanni Di Domenico (Vice-Presidente).

Desideriamo ancora ringraziare Nicola Cavalli: se nel primo Seminario compariva come uno dei relatori, adesso partecipa in quanto editore (Ledizioni). Anche questo è un aspetto rappresentativo di un ambito di studi e ricerche che include professionisti e accademici, senza distinzioni di meridiani e paralleli, ma semmai pienamente consapevole del fatto che la ricchezza proviene da ogni voce, da ogni esperienza, da ogni punto di vista.

Bibliografia

- Abadal, E. (2013), “Le condizioni della Biblioteconomia e della Documentazione nell’università spagnola”, in: *1. Seminario Nazionale di Biblioteconomia*. Milano: Ledizioni, pp. 63-78, <<https://doi.org/10.4000/books.ledizioni.1200>>.
- Bergamini A. (2022), ‘I Seminario hispano-italiano en Biblioteconomía y Documentación: estado actual y perspectivas de futuro. Actas’, *Biblioteche oggi*, vol. 40, n. 5, pp. 76-79, <<http://dx.doi.org/10.3302/0392-8586-202205-076-1>>.
- Bernabè A. (2020), ‘Spunti dal I Seminario hispano-italiano en Biblioteconomía y Documentación (29-30 ottobre 2020). Formazione, attività scientifica e professione, in un’interpretazione «mediterranea» della Biblioteconomía’, *TECA*, vol. 10, n. 2, <<https://doi.org/10.6092/issn.2240-3604/12660>>.
- Bilotta, A. (2020), ‘Per una Biblioteconomia mediterranea: note a margine del I Seminario hispano-italiano en biblioteconomía y documentación’, *AIB studi*, vol. 60, n. 3, pp. 671-688, <<https://doi.org/10.2426/aibstudi-12762>>.
- Bilotta, A. (2022), *Principi, approcci e applicazioni della Biblioteconomia comparata: una proposta per nuovi percorsi di ricerca*. Firenze, Firenze University Press.
- Braidotti, R. (1995), “La molteplicità: un’etica per la nostra epoca, oppure meglio cyborg che dea”, in: Donna J. Haraway, *Manifesto Cyborg. Donne, tecnologie e biopolitiche del corpo*. Milano, Feltrinelli, pp. 9-38.
- Cassano, F. (2001^a), *Modernizzare stanca. Perdere tempo, guadagnare tempo*. Bologna, il Mulino.
- Cassano, F. (2001^b), *Il pensiero meridiano*. Roma-Bari, Laterza.
- Castellucci, P. (2007), ‘Eredità e prospettive per la Documentazione in Italia’, *Bollettino AIB*, vol. 47 n. 3, pp. 239-255.
- Maiello, R. (2007), ‘Formazione e aggiornamento: attori, strategie, alleanze’, *AIB notizie*, vol. 19, n. 1, <<https://www.aib.it/aib/editoria/n19/0112.htm3>>.

- Massara, F. (2019), 'Biblioteconomia e culture del libro in Italia e Spagna. Nasce un progetto di collaborazione al Salone internazionale del Libro di Torino', *Biblioteche oggi*, vol. 37, pp. 45-47. <<http://dx.doi.org/10.3302/0392-8586-201906-045-1>>.
- Gonzalo Sánchez-Molero, J. L., Caridad Sebastián, M. (2021), "Introducción", in: *I Seminario hispano-italiano en biblioteconomía y documentación : estado actual y perspectivas de futuro : (29-30 de octubre de 2020)*, José Luis Gonzalo Sánchez-Molero, Mercedes Caridad Sebastián (editores), Madrid : Facultad de Ciencias de la Documentación (Universidad Complutense Madrid), 2021, pp. 5-8.
- López Yepes, J. (2005), 'Il concetto di Documentazione e il suo riflesso nella formazione di professionisti e ricercatori in Spagna', *Bollettino AIB*, vol. 45, n. 3, pp. 277-294.

I

BIBLIOTECHE, TECNOLOGIE E INTELLIGENZA ARTIFICIALE:
STUDI ED ESPERIENZE A CONFRONTO

BIBLIOTECAS, TECNOLOGÍAS Y INTELIGENCIA ARTIFICIAL:
ESTUDIOS Y EXPERIENCIAS COMPARADAS

Coordinatori

Maurizio Vivarelli (Università di Torino)

Mercedes Caridad (Universidad Carlos III de Madrid)

MODELLI E PRATICHE DELLA LETTURA IN BIBLIOTECA. IL CONTESTO BIBLIOGRAFICO DEL PROGETTO READING(&)MACHINE

Maurizio Vivarelli*

Premessa

Reading(&)Machine (d'ora in avanti R(&)M) è un progetto promosso e sviluppato in collaborazione tra i centri SmartData@PoliTO e VR@PoliTO del Politecnico di Torino, il Dipartimento di Studi storici dell'Università di Torino e le Biblioteche Civiche Torinesi (BCT). L'obiettivo specifico è quello di realizzare un sistema di raccomandazione (in seguito SR) per la lettura, i cui suggerimenti saranno comunicati in uno spazio virtuale immersivo localizzato nello spazio fisico di una biblioteca civica torinese. Attualmente è in corso lo sviluppo di un prototipo realizzato grazie al finanziamento ottenuto con il bando *Facciamola facile* della Fondazione TIM, che sarà installato presso la Biblioteca Civica Centrale di Torino. Le caratteristiche di fondo del progetto sono state presentate in un seminario dal titolo *La promozione della lettura verso il futuro*, che ha avuto luogo a Torino, nel Salone d'onore del Castello del Valentino, il 18 febbraio 2022 (https://smartdata.polito.it/readingmachine_event/). R(&)M, mentre veniva definita la sua struttura concettuale e tecnologica¹, e le sue implicazioni di natura bibliografica e biblioteconomica, è stato presentato in

* Università di Torino, maurizio.vivarelli@unito.it. Data di ultima consultazione dei siti web 20 febbraio 2023.

¹ Le caratteristiche del SR sono discusse in Speciale A., Vallero G., Vassio L. & Mellia M. (2023), 'Recommendation Systems in Libraries: An Application with Heterogeneous Data Sources', in: *Proceedings of the 26th International Conference on Extending Database Technology (EDBT), 28th March - 31st March, 2023*, online su OpenProceedings.org. Si segnala che una più ampia descrizione del progetto è prevista in *Biblioteche, lettura, intelligenza artificiale: profilo e contesto del progetto Reading(&)Machine*, in corso di stampa presso Editrice Bibliografica, a cura di Fabrizio Lamberti, Marco Mellia, Maurizio Vivarelli.

alcune sedi editoriali² e convegnistiche, in Italia e all'estero³. Si segnala infine che presso il Dipartimento di Studi storici dell'Università di Torino è stato costituito il gruppo di ricerca "Biblioteche, lettura, intelligenza artificiale"⁴.

In questo contributo verranno proposte in primo luogo alcune considerazioni sulle relazioni tra biblioteche, lettura, Intelligenza Artificiale (d'ora in avanti IA), e successivamente verrà brevemente descritta l'architettura concettuale e tecnologica del progetto. Infine saranno discussi alcuni aspetti dello *spazio semiotico*⁵ all'interno del quale R(&)M si situa, con particolare riferimento alle modalità con cui i suggerimenti del SR si ineriscono nelle complesse dinamiche della *scelta di lettura*.

Biblioteche, lettura, Intelligenza Artificiale

La letteratura scientifica e professionale di riferimento tende a valutare i

² Le linee generali di R(&)M sono state presentate per la prima volta in Mellia M. & Vivarelli M. (2021), 'Reading(&)Machine: identità della biblioteca e Intelligenza Artificiale', in: *La biblioteca piattaforma della conoscenza: inclusiva, collaborativa, reticolare, Convegno delle Stelline 2021*. Milano: Editrice Bibliografica; Associazione Biblioteche oggi, pp. 17-34. Gli aspetti bibliografici e biblioteconomici sono stati discussi in Vivarelli, M. (2022), 'La scelta di lettura nella biblioteca pubblica: fisionomia, spazi e contesti del progetto Reading(&)Machine', *AIB studi*, vol. 62, no. 1, pp. 27-55, DOI <<https://10.2426/aibstudi-13384>>; Vivarelli, M. (2022), 'Pratiche di lettura nello spazio della biblioteca: il contesto del progetto Reading(&)Machine', *DigitCult. Scientific Journal on Digital Cultures*, vol. 7, no. 2, pp. 7-22, DOI <<https://10.36158/97888929562231>>; Vivarelli, M. (2022), 'The Identity of the Public Library and Reading Practices. General Lines of the Reading(&)Machine Project and its Context', in: *Cases on Developing Effective Research Plans for Communications and Information Science*, eds. Carrillo-Duran M. V., Pérez Pulido, M. Hershey, PA. USA: Information Science Reference, an imprint of IGI Global, pp. 335-344.

³ Mellia, M. & Vivarelli, M. (2021), 'Reading(&)Machine: identità della biblioteca e intelligenza artificiale. Convegno Stelline 2021', cit. Gli aspetti biblioteconomici del progetto sono stati comunicati da Maurizio Vivarelli in queste occasioni seminariali: 'Reading(&)Machine: lectura y algoritmos para un nuevo modelo de biblioteca pública', Badajoz, 18 ottobre 2021; 'Biblioteques, lectura, intel·ligència artificial: el projecte d'un nou model de biblioteca pública', Barcellona, 19 ottobre 2021; 'Reading(&)Machine: lettura e algoritmi per un nuovo modello di biblioteca pubblica', 3rd International Congress of the Libraries, Politics, Research Network, "The digital, the traditional, the new normal?", Lisbona, 23-24 settembre 2021.

⁴ <https://www.dipstudistorici.unito.it/do/gruppi.pl/Show?_id=l4qb>.

⁵ Il concetto di *spazio semiotico* è stato utilizzato da Jurij Lotman per descrivere una particolare tipologia di spazio, locale e delimitabile, che lo differenzia da quello di altre sfere semiotiche. Cfr. Lotman, J. M. (1984), *La semiosfera: l'asimmetria e il dialogo nelle strutture pensanti*, a cura di Salvestroni S. Venezia: Marsilio (*Universe of the mind: a semiotic theory of culture*, 1990).

benefici delle possibili applicazioni di Intelligenza Artificiale alle procedure attualmente in uso nelle biblioteche, e talvolta solo a quella che potrebbe essere definita la loro «meccanizzazione». Alcuni esempi in questo senso riguardano generazione di metadati, attività di *text mining* e analisi del linguaggio naturale, *blockchain*, *chatbot*, assistenti vocali, analisi e visualizzazione di *big data*⁶. Un recente rapporto OCLC, curato da Thomas Padilla, mette in ogni caso in evidenza i significativi ritardi che «persist between concept and operationalization in libraries at the level of workflows (managing bias in probabilistic description), policies (community engagement vis-à-vis the development of machine-actionable collections), positions (developing staff who can utilize, develop, critique, and/or promote services influenced by data science, machine learning, and AI), collections (development of “gold standard” training data), and infrastructure (development of systems that make use of these technologies and methods)»⁷.

R(&)M si correla evidentemente a questa cornice di riferimento, e tiene conto dei confini disciplinari ad essa sottesi. Tuttavia intende esplorare anche ciò che si situa all'esterno di quei confini, partendo dalla presa d'atto del fatto che, durante l'esperienza di lettura – qualunque sia la sua natura – vengono

⁶ Artificial Intelligence and Machine Learning in Libraries' (2019), Jason Griffey, editor, *Library Technology Reports. Expert Guides to Library Systems and Service*; Padilla, T. (2019), *Responsible Operations: Data Science, Machine Learning, and AI in Libraries*. Dublin: OCLC Research; Cordell, R. (2020) *Machine Learning + Libraries: A Report on the State of Field*, Library of Congress, <<https://labs.loc.gov/static/labs/work/reports/Cordell-LOC-ML-report.pdf>>; Cox, A. (2021), *The Impact of AI, Machine Learning, Automation and Robotics on the Information Profession: A Report for CILIP*, <<https://www.cilip.org.uk/general/custom.asp?page=researchreport>>; Saorín, T. (2021), 'Big data literario de raíz bibliotecaria: reflexiones sobre infraestructuras de anotación, catalogación, descubrimiento y recomendación de ficción narrativa', *Anuario ThinkEPI*, vol. 15, DOI <<https://doi.org/10.3145/thinkepi.2021.e15c01>>. Tra gli usi di strumenti AI su questa linea di confine va ricordato il SR Obotti sviluppato dalla Helsinki Central Library Oodi: cfr. Hammis, E., Ketamo, H. & Koivisto, A. (2019), 'Virtual Information Assistants on Mobile App to Serve Visitors at Helsinki Central Library Oodi', in *Libraries: Dialogue for Change*, IFLA WLIC 2019, Athens, Greece, 26 agosto, <<https://www.theseus.fi/handle/10024/303597>>.

⁷ Padilla, T. (2019), *Responsible Operations: Data Science, Machine Learning, and AI in Libraries*, cit., pp. 7-8. Cfr. anche il recente *IFLA Statement on Libraries and Artificial Intelligence*, elaborato da IFLA FAIFE (Committee on Freedom of Access to Information and Freedom of Expression), <<https://repository.ifla.org/handle/123456789/1646>>, nelle cui premesse si afferma che «AI technologies can have deeply transformative capabilities, and their power can be put to the service of public welfare and innovation. With necessary preparations – and regard for ethical concerns and current limitations – libraries can responsibly use AI technologies to advance their social mission».

utilizzati contenuti informativi veicolati da metadati che in parte provengono dall'interno del campo della disciplina, situati metaforicamente all'interno dell'Universo bibliografico, ed in parte sono costituiti da dati generati dagli utenti e da dati valutativi e testuali provenienti dalla piattaforma di social reading aNobii (<https://www.anobii.com/it>).

In questa sede non è possibile discutere il profilo disciplinare della Library and Information Science nella sua dimensione complessiva⁸. Per questo se, come scrive in una recente pubblicazione Maria Teresa Biagetti, i fondamenti del campo disciplinare della Bibliografia, della Biblioteconomia, della Library and Information Science, della stessa Documentazione, si individuano nello «sviluppare le più efficaci modalità di accesso conoscenza registrata», allora il concetto di Universo bibliografico, che coincide con quello di conoscenza registrata, ne diviene uno degli elementi centrali⁹. Le difficoltà nella delimitazione del significato del termine sono discusse in un denso articolo di Amanda Cossham, la quale afferma proprio che quello di Universo bibliografico è un concetto molto più frequentemente osservato e descritto che definito¹⁰. Cossham elenca undici sommarie definizioni, alcune delle quali sono riportate qui di seguito, in base alle quali il significato dell'espressione varia sensibilmente, oscillando tra: a) il “contenuto” presente negli oggetti; b) la materialità dell'insieme più limitato di oggetti localizzabile delle biblioteche; c) la rappresentazione memorizzata di quegli stessi oggetti all'interno dei database¹¹. La Tab. 1 presenta alcune di queste definizioni, riportate per esteso nella appendice dell'articolo.

⁸ Cfr., per un sintetico inquadramento, Vivarelli, M. (2020), 'Digital humanities e culture documentarie: un modello di analisi, valutazione, interpretazione', *AIB studi*, vol. 60, no. 3, pp. 553-589, DOI <<https://10.2426/aibstudi-12471>>.

⁹ Biagetti, M. T. (2022), *La Scienza dell'informazione: identità ed esperienze*. Milano: Franco Angeli, p. 152. Questa connessione biunivoca è esplicitata anche in *La trasmissione della conoscenza registrata* (2021), a cura di Carlo Bianchini e Lucia Sardo. Milano: Editrice Bibliografica.

¹⁰ Cossham, A. F. (2013), 'Bibliographic Records in an Online Environment. Proceedings of the Eighth International Conference on Conceptions of Library and Information Science, Copenhagen, Denmark, 19-22 August, 2013', *Information Research*, vol. 18, no. 3, <<http://informationr.net/ir/18-3/colis/paperC42.html#.YEzhTJ1Kg2w>>.

¹¹ Per una più estesa trattazione cfr. Vivarelli, M. (2021), 'Dalla giusta distanza: biblioteca e memoria nel Multiverso bibliografico', *Biblioteche oggi Trends*, vol. 7, no. 1, pp. 16-31, DOI <<https://dx.doi.org/10.3302/2421-3810-202101-016-1>>.

Citazione	Riferimenti bibliografici
<p>«the totality of things over which bibliographical control is or might be exercised, consists of writings and recorded sayings», which ‘includes items of radically different sorts’» (p. 6); «‘manuscripts as well as printed books, bills of lading and street signs as well as personal letters, inscriptions on stone as well as phonograph recordings of speeches, and most notably, memorized texts in human heads and texts stored up in the “memories” of machines’» (p. 12); «‘the universe of writings and sayings [not] the universe of pictorial and musical works’ (p. 14)».</p>	<p>Wilson, P. (1968), <i>Two Kinds of Power: An Essay in Bibliographical Control</i>. Los Angeles, CA: University of California.</p>
<p>«The totality of bibliographic entities and their relationships. In a sense, the bibliographic universe consists of all types of intellectual or physical objects in any format which contain works of imagination as well as information’» (chapter 4).</p>	<p>Fattahi, R. (2010), <i>From Information to Knowledge: Superworks and the Challenges in the Organization and Representation of the Bibliographic Universe: lectio magistralis in Library science = Dall’informazione alla conoscenza: le super-opere e le sfide dell’organizzazione e rappresentazione dell’universo bibliografico: lectio magistralis in Biblioteconomia</i>. Fiesole: Casalini Libri.</p>
<p>«information resources [that] exist ... within a given “universe” (e.g., within the totality of available information resources, within the published output of a particular country, within the holdings of a particular library or group of libraries, etc.)’» (p. 8); «‘the universe of entities described in bibliographic records’» (p. 3); «‘the universe that is represented in a bibliography, catalogue, or bibliographic database» (p. 55).</p>	<p>IFLA Study Group on the Functional Requirements for Bibliographic Records (1998), <i>Functional Requirements for Bibliographic Records: Final Report</i>, <https://repository.ifa.org/handle/123456789/83>.</p>
<p>«includes anything a library might wish to collect or make accessible to its users» (p. 197).</p>	<p>Tillett, B. (2004), <i>What is FRBR? A Conceptual Model for the Bibliographic Universe</i>, Library of Congress, <https://www.loc.gov/cds/downloads/FRBR.PDF>.</p>
<p>«consists of ‘documents, sets of these (formed by attributes such as work, edition, author and subject), and relationships among them’» (p. 32); «‘The smallest or basic entities in the bibliographic universe are documents. Documents, which have been defined as information-bearing messages in recorded form, are individuals or singular entities» (p. 34).</p>	<p>Svenonius, E. (2000), <i>The Intellectual Foundation of Information Organization</i>. Cambridge, MA: MIT Press.</p>
<p>«‘a conceptual model of the bibliographic universe as represented in library catalogues through descriptions (bibliographic records)’» (p. 13).</p>	<p>Zhang, Y. & Salaba, A. (2009), <i>Implementing FRBR in Libraries: Key Issues and Future Directions</i>, New York: Neal-Schuman.</p>

Tab. 1 - Definizioni del concetto di Universo Bibliografico. Elaborazione da Cossham A. F. (2015), *Bibliographic Records in an Online Environment*.

L'Universo bibliografico può essere dunque qualificato come lo spazio della rappresentazione linguistica delle risorse bibliografiche, ed è finalizzato essenzialmente a garantire la ricerca ed il recupero dei contenuti informativi di quelle stesse risorse; per questo è biunivocamente correlato all'ambito dell'*Information retrieval*¹². L'Universo bibliografico, schematizzando i meccanismi della sua produzione, gestione ed uso, è rappresentato graficamente con il modello della Fig. 1.

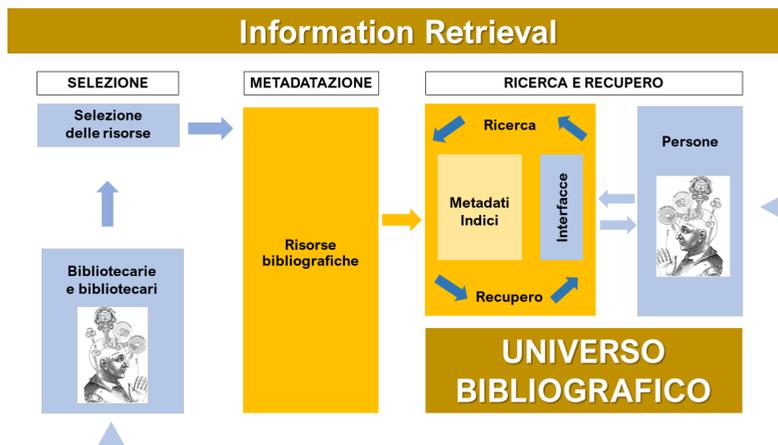


Fig. 1. Modello di rappresentazione delle relazioni tra Information retrieval e Universo bibliografico.

R(&)M, per i motivi che qui sono stati introdotti, si situa dunque in uno spazio semiotico più esteso di quello dell'Universo bibliografico, che potrebbe essere definito Multiverso bibliografico¹³. Il concetto di Multiverso bibliografico, di fatto, coincide con il campo di applicazione di ELIS, il modello di analisi del comportamento informativo delle persone nell'esperienza quotidiana, reso noto da tempo da Reijo Savolainen (Fig. 2), e può essere rappresentato graficamente con l'immagine della Fig. 3¹⁴.

¹² Cfr. Hjørland, B. (2021), 'Information Retrieval and Knowledge Organization: A Perspective from the Philosophy of Science', *Information*, 12, no. 3: article 135, DOI <<https://doi.org/10.3390/info12030135>>; pubblicato con gli stessi contenuti anche in "Information retrieval", in: *ISKO Encyclopedia of Knowledge Organization*, eds. Birger Hjørland and Claudio Gnoli, <<https://www.isko.org/cyclo/ir>>.

¹³ E per cui si rimanda a Vivarelli, M. (2021), *Dalla giusta distanza: biblioteca e memoria nel Multiverso bibliografico*, cit.

¹⁴ Savolainen, R. (1995), 'Everyday Life Information Seeking: Approaching Information Seeking in the Context of Way of Life', *Library & Information Science Research*, vol. 17, pp. 259-294, DOI >[https://doi.org/10.1016/0740-8188\(95\)90048-9](https://doi.org/10.1016/0740-8188(95)90048-9)>.

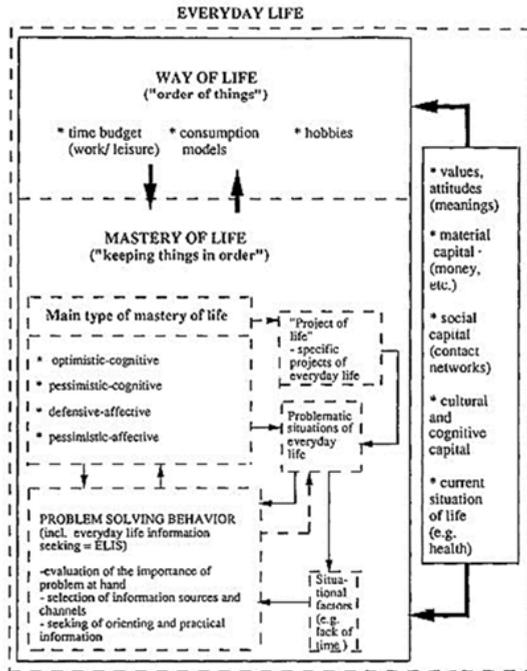


Fig. 2. Rappresentazione grafica del modello ELIS. Fonte: Savolainen, R. (1995), *Everyday life information seeking*.

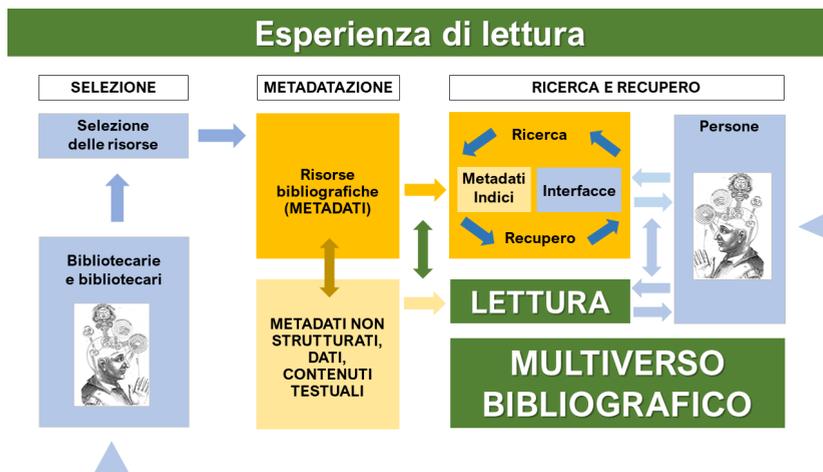


Fig. 3. Modello di rappresentazione del Multiverso bibliografico, con la estensione derivante dalla inclusione nello spazio della biblioteca dell'esperienza della lettura.

Le componenti della struttura di Reading(&)Machine sono caratterizzate da tre “piani”, relativi rispettivamente all’utente, ai dati, agli algoritmi (Fig. 4).



Fig. 4. Schema della struttura di Reading(&)Machine. Elaborazione di Marco Mellia, centro SmartData@PoliTO.

Le tre parti componenti sono costituite da un *sistema di raccomandazione*, basato su metadati bibliografici della biblioteca, dati generati dagli utenti (naturalmente anonimizzati), relativi in particolare ad attività di prestito, dati estratti dalla piattaforma di social reading aNobii¹⁵; una *interfaccia immersiva*,

¹⁵ Ricci, F., Rokach, L. & Shapira, B., “Recommender Systems: Techniques, Applications, and Challenges”, in: Ricci, F., Rokach, L. & Shapira, B. (eds.) (2022),

in un ambiente di Realtà Virtuale cui si accede con il visore Oculus (Fig. 5), in cui vengono mostrati i suggerimenti elaborati dagli algoritmi¹⁶; un *ambiente ibrido, fisico e digitale*, con il quale i suggerimenti dell'algoritmo e l'interfaccia entrano a far parte dello *spazio della biblioteca*¹⁷.

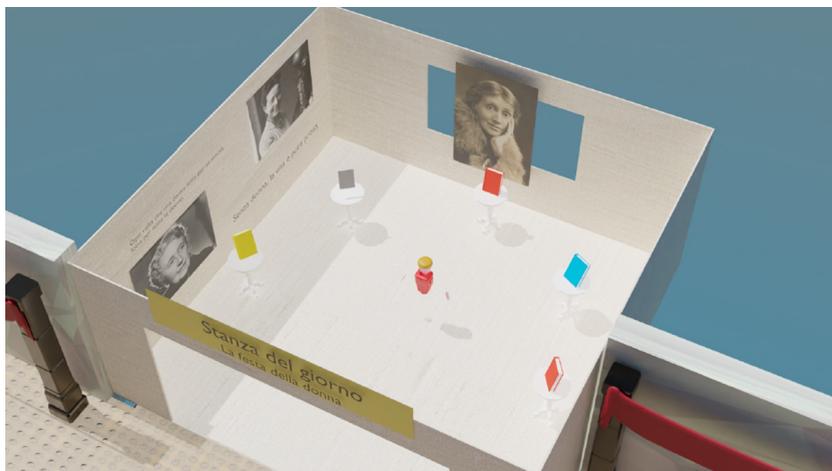


Fig. 5. Uno degli ambienti della spazio immersivo di accesso a R(&)M.
Elaborazione centro VR@PoliTo.

La scelta di lettura, dentro e fuori l'Universo bibliografico

All'interno dell'Universo bibliografico si effettuano attività di *ricerca e recupero delle informazioni* attraverso i cataloghi, che sono contestuali alla *lettura* di quelle specifiche interfacce. Ma nello spazio semiotico cui si riferisce l'espressione 'biblioteca pubblica' esistono numerose altre tipologie di segni e di codici, ai quali sono rivolte differenti pratiche di lettura, che può essere utile cercare di categorizzare. Con una semplice tassonomia possiamo dunque individuare in primo luogo la *lettura socio-semiotica*, che ha per oggetto i segni disposti nello spazio nel loro insieme, e che possiamo pensare specificata

Recommender Systems Handbook, 3rd edn. New York: Springer, pp. 1-35, DOI <https://10.1007/978-1-0716-2197-4_1>.

¹⁶ Cfr. Lanier, J. (2019), *L'alba del nuovo tutto: il futuro della realtà virtuale*. Milano: Il Saggiatore (*Dawn of the New Everything*, 2017).

¹⁷ Cfr. Cesareo, F. (2022), 'Reading(&)Machine tra spazio virtuale e spazio fisico', seminario per il corso di Storia delle biblioteche e della lettura, Dipartimento di Studi storici, Università di Torino, 7-8 giugno.

in: 1) *lettura documentaria*, applicata alle procedure di ricerca e recupero dell'informazione; 2) *lettura di esplorazione*, correlata al browsing dei contenuti non strettamente documentari; 3) *lettura testuale*, che accede ai contenuti del testo; 4) *lettura sincretica*, riferita alla elaborazione di testi sincretici, che presentano contenuti integrati di natura iconica e grafica; 5) *lettura letteraria*, di cui si occupano la teoria e la critica della letteratura; 6) *lettura ascoltata* (e spesso anche *vista*), con cui possiamo riferirci alla percezione ed elaborazione del contenuto testuale sonorizzato e spesso drammatizzato, come ad esempio negli audiolibri, nei podcast, o anche durante un reading collegato alla presentazione di un libro; 7) *lettura di studio e ricerca*, riferita alla lettura intensiva di un contenuto testuale per finalità generali di apprendimento¹⁶. Se dunque la lettura documentaria avviene all'interno dell'Universo bibliografico, tutte le altre tipologie fuoriescono da esso, collocandosi nel Multiverso. E ognuna di queste pratiche di lettura, prima di essere attivata, deve essere preceduta da una *scelta*. L'individuazione analitica dei motivi che orientano ed alla fine producono la scelta di lettura non è possibile, per l'ampiezza indeterminabile degli elementi che, dinamicamente ed interattivamente, vengono tra loro relazionati nella relazione tra testo e lettore: stati d'animo, preferenze per particolari generi o autori o argomenti, seduttività dei segni paratestuali, formato del testo, e così via. Tuttavia si ritiene utile rilevare il fatto che, nello spazio semiotico della biblioteca, con R(&)M si aggiungono ulteriori segni, dotati di una loro specifica e problematica *agency*, che alla fine configurano la scelta di lettura e la localizzazione successiva dell'unità fisica¹⁸. Queste fasi sono schematizzate nella Fig. 6. La fase 1 segnala la provenienza dei dati; la 2 la produzione dei suggerimenti del SR nell'ambiente immersivo, che interagiscono nella fase 3 con i contenuti cognitivi ed emotivi del lettore fuori campo. Nella fase 4, passando da R(&)M al catalogo, viene localizzata l'unità fisica che, nella fase 5, viene metaforicamente prelevata dagli scaffali.

¹⁸ Questioni rilevanti, indagate con attenzione in Lana, M. (2022), 'L'agency dei sistemi di intelligenza artificiale: un punto di vista bibliografico', *DigitCult. Scientific Journal on Digital Cultures*, vol. 7, no. 1, pp. 67-77, DOI <<https://doi.org/10.36158/97888929552576>>.

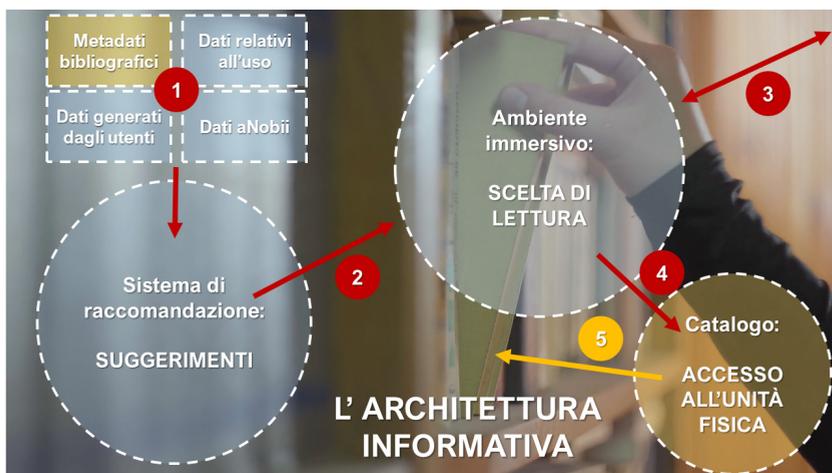


Fig. 6. L'architettura informativa della scelta di lettura.

Il futuro di Reading(&)Machine

Mentre la realizzazione per la stampa degli atti del convegno era in corso presso la Biblioteca Civica Centrale di Torino è stato effettuato un periodo di sperimentazione delle funzioni del prototipo, i cui esiti saranno resi noti non appena disponibili. L'obiettivo è quello di rendere disponibile R(&)M, nella sua versione definitiva, nello spazio della nuova Biblioteca Centrale di Torino, che dovrebbe essere completato nel 2026 all'interno dell'edificio di Torino Esposizioni (Fig. 7)¹⁹. L'auspicio è dunque che R(&)M possa contribuire a qualificare ulteriormente l'offerta dei servizi della nuova biblioteca, continuando a mantenere l'esperienza della lettura al suo centro, perché è nella configurazione di questo atto che le biblioteche possono (e a parere di chi scrive debbono) continuare ad individuare e rinnovare, coltivandole, le proprie radici²⁰.

¹⁹ Il progetto della nuova Biblioteca Centrale di Torino, di cui sono in corso le fasi finali delle procedure per l'affidamento della esecuzione dei lavori, è stato presentato da Cecilia Cognigni durante gli Stati generali delle biblioteche (Milano, 25-26 ottobre 2022) con una relazione dal titolo *Biblioteche che cambiano la città: il caso di Torino*, <<https://milano.biblioteche.it/statigenerali>>. In questa occasione sono stati presentati anche alcuni elementi del progetto biblioteconomico, curati da chi scrive (*Il modello biblioteconomico della nuova Biblioteca Centrale di Torino*).

²⁰ Concordo dunque pienamente con Luca Ferrieri quando scrive che «le biblioteche debbano metter la loro anima nella lettura, perché è lì che esse hanno radici



Fig. 7. Torino Esposizioni. Progetto di Ettore Sottsass, Pier Luigi Nervi, Roberto Biscaretti di Ruffia, Riccardo Morandi. Vista del Padiglione Nervi.

Riferimenti bibliografici

- ‘Artificial Intelligence and Machine Learning in Libraries’ (2019), Jason Griffey, editor, *Library Technology Reports. Expert Guides to Library Systems and Service*. Biagetti, M. T. (2022), *La Scienza dell’informazione: identità ed esperienze*. Milano: Franco Angeli.
- Biblioteche, lettura, intelligenza artificiale: profilo e contesto del progetto Reading(&) Machine*, in corso di stampa presso Editrice Bibliografica, a cura di Lamberti F., Mellia M., Vivarelli M.
- Cordell, R. (2020), *Machine Learning + Libraries: A Report on the State of Field*, Library of Congress, <<https://labs.loc.gov/static/labs/work/reports/Cordell-LOC-ML-report.pdf>>.
- Cossham, A. F. (2013), ‘Bibliographic Records in an Online Environment. Proceedings of the Eighth International Conference on Conceptions of Library and Information Science, Copenhagen, Denmark, 19-22 August, 2013’, *Information Research*, vol. 18, no. 3, <<http://informationr.net/ir/18-3/colis/paperC42.html#.YEzhTJ1Kg2w>>.

e futuro»: cfr. Ferrieri, L. (2020), *La biblioteca che verrà: pubblica, aperta, sociale*. Milano: Editrice Bibliografica, p. 209.

- Cox, A. (2021), *The Impact of AI, Machine Learning, Automation and Robotics on the Information Profession: A Report for CILIP*, <<https://www.cilip.org.uk/general/custom.asp?page=researchreport>>.
- Fattahi, R. (2010), *From Information to Knowledge: Superworks and the Challenges in the Organization and Representation of the Bibliographic Universe: lectio magistralis in library science = Dall'informazione alla conoscenza: le super-opere e le sfide dell'organizzazione e rappresentazione dell'universo bibliografico: lectio magistralis in Biblioteconomia*. Fiesole: Casalini Libri.
- Hammais, E., Ketamo, H. & Koivisto, A. (2019), 'Virtual Information Assistants on Mobile App to Serve Visitors at Helsinki Central Library Oodi', in: *Libraries: Dialogue for Change*, IFLA WLIC 2019, Athens, Greece, 26 agosto, <<https://www.theseus.fi/handle/10024/303597>>.
- Hjørland, B. (2021), 'Information Retrieval and Knowledge Organization: A Perspective from the Philosophy of Science', *Information*, 12, no. 3: article 135, DOI <<https://doi.org/10.3390/info12030135>>.
- IFLA FAIFE, *Statement on Libraries and Artificial Intelligence*, <<https://repository.ifla.org/handle/123456789/1646>>.
- IFLA Study Group on the Functional Requirements for Bibliographic Records (1998), *Functional Requirements for Bibliographic Records: Final Report*, <<https://repository.ifla.org/handle/123456789/83>>
- "Information retrieval", in: *ISKO Encyclopedia of Knowledge Organization*, eds. Birger Hjørland and Claudio Gnoli, <<https://www.isko.org/cyclo/ir>>.
- Lotman, J. M. (1984), *La semiosfera: l'asimmetria e il dialogo nelle strutture pensanti*, a cura di Salvestroni S. Venezia: Marsilio, (*Universe of the mind: a semiotic theory of culture*, 1990).
- Mellia, M. & Vivarelli, M. (2021), 'Reading(&)Machine: identità della biblioteca e Intelligenza Artificiale', in: *La biblioteca piattaforma della conoscenza: inclusiva, collaborativa, reticolare, Convegno delle Stelline 2021*. Milano: Editrice Bibliografica.
- Mellia, M. & Vivarelli, M. (2021), 'The Identity of the Public Library and Reading Practices. General Lines of the Reading(&)Machine Project and its Context', in: *Cases on Developing Effective Research Plans for Communications and Information Science*, eds. Carrillo-Duran M. V., Pérez Pulido, M. Hershey, PA, USA: Information Science Reference, an imprint of IGI Global, pp. 335-344.
- Padilla, T. (2019), *Responsible Operations: Data Science, Machine Learning, and AI in Libraries*. Dublin: OCLC Research.
- 'Reading(&)Machine: lettura e algoritmi per un nuovo modello di biblioteca pubblica', 3. International Congress of the Libraries, Politics, Research Network, "The digital, the traditional, the new normal?", Lisbona, 23-24 settembre 2021.
- Saorín, T. (2021), 'Big data literario de raíz bibliotecaria: reflexiones sobre infraestructuras de anotación, catalogación, descubrimiento y recomendación de ficción narrativa', *Anuario ThinkEPI*, vol. 15, DOI <<https://doi.org/10.3145/thinkepi.2021.e15c01>>.
- Savolainen, R. (1995), 'Everyday Life Information Seeking: Approaching

- Information Seeking in the Context of “Way of Life”, *Library & Information Science Research*, vol. 17, pp. 259-294, DOI <[https://doi.org/10.1016/0740-8188\(95\)90048-9](https://doi.org/10.1016/0740-8188(95)90048-9)>.
- Speciale, A., Vallero, G., Vassio, L. & Mellia, M. (2023), ‘Recommendation Systems in Libraries: An Application with Heterogeneous Data Sources’, in: *Proceedings of the 26th International Conference on Extending Database Technology (EDBT)*, 28th March - 31st March, 2023, online su [OpenProceedings.org](https://openproceedings.org).
- Svenonius, E. (2000), *The Intellectual Foundation of Information Organization*. Cambridge, MA: MIT Press.
- Tillett, B. (2004), *What is FRBR? A Conceptual Model for the Bibliographic Universe*, *Library of Congress*, <<https://www.loc.gov/cds/downloads/FRBR.PDF>>.
- La trasmissione della conoscenza registrata* (2021), a cura di Carlo Bianchini e Lucia Sardo. Milano: Editrice Bibliografica.
- Vivarelli, M. (2020), ‘Digital humanities e culture documentarie: un modello di analisi, valutazione, interpretazione’, *AIB studi*, vol. 60, no. 3, pp. 553-589, DOI <<https://10.2426/aibstudi-12471>>.
- Vivarelli, M. (2021), ‘Dalla giusta distanza: biblioteca e memoria nel Multiverso bibliografico’, *Biblioteche oggi Trends*, vol. 7, no. 1, pp. 16-31, DOI <<http://dx.doi.org/10.3302/2421-3810-202101-016-1>>.
- Vivarelli, M. (2021), ‘Reading(&)Machine: lectura y algoritmos para un nuevo modelo de biblioteca pública’, Badajoz, 18 ottobre 2021; ‘Biblioteques, lectura, intel·ligència artificial: el projecte d’un nou model de biblioteca pública’, Barcellona, 19 ottobre 2021.
- Vivarelli, M. (2022), ‘La scelta di lettura nella biblioteca pubblica: fisionomia, spazi e contesti del progetto Reading(&)Machine’, *AIB studi*, vol. 62, no. 1, pp. 27-55, DOI <<https://10.2426/aibstudi-13384>>.
- Vivarelli, M. (2022), ‘Pratiche di lettura nello spazio della biblioteca: il contesto del progetto Reading(&)Machine’, *DigitCult. Scientific Journal on Digital Cultures*, vol. 7, no. 2, pp. 7-22, DOI <<https://10.36158/97888929562231>>.
- Vivarelli, M. (2022), ‘The Identity of the Public Library and Reading Practices. General Lines of the Reading(&)Machine Project and its Context’, in: *Cases on Developing Effective Research Plans for Communications and Information Science*, eds. Carrillo-Duran M. V., Pérez Pulido, M. Hershey, PA, USA: Information Science Reference, an imprint of IGI Global, pp. 335-344.
- Wilson, P. (1968), *Two Kinds of Power: An Essay in Bibliographical Control*. Los Angeles, CA: University of California.
- Zhang, Y. & Salaba, A. (2009), *Implementing FRBR in Libraries: Key Issues and Future Directions*. New York: Neal-Schuman.

INTELLIGENZA ARTIFICIALE NELLE BIBLIOTECHE. STATO DELL'ARTE ED ESPERIENZE DI APPLICAZIONE

Rossana Morriello*

Introduzione

L'intelligenza artificiale è già ampiamente entrata nel mondo delle biblioteche di ogni tipologia. Biblioteche pubbliche, accademiche, nazionali hanno introdotto sistemi di intelligenza artificiale e algoritmi di machine learning in diverse attività bibliotecarie. Sistemi automatizzati con l'IA vengono impiegati per la catalogazione, per fornire informazioni tramite chatbot e assistenti virtuali, oppure per il trattamento, l'estrazione dei dati e l'indicizzazione dei contenuti testuali e delle immagini, fino ad arrivare ai robot che in alcune biblioteche accolgono gli utenti e intrattengono i più piccoli. Le potenzialità dell'intelligenza artificiale si estendono quindi alle attività tradizionali della biblioteca quali la catalogazione, il reference, l'information literacy, soprattutto nella sua declinazione di algorithmic literacy, lo sviluppo e la promozione delle raccolte (Morriello, 2019; Oyelude, 2021, Cox & Mazumdar, 2022), permettendo di progettare servizi più incentrati sull'utente. Le applicazioni tecnologiche a disposizione sono già molte, e vanno dai più consolidati chatbot per i servizi "Ask a librarian" integrati in ambienti web, alle applicazioni per cellulare, o ai già citati robot come Pepper, il robot umanoide prodotto da SoftBank Robotics, in grado di interagire con l'ambiente circostante. Ci soffermeremo più avanti su alcune di queste applicazioni.

La rilevanza dell'intelligenza artificiale per le biblioteche è resa evidente sia dalla crescente attenzione da parte degli studi biblioteconomici rispetto alle sue potenzialità e limiti, sia dall'interesse della comunità professionale, a cominciare dall'IFLA (International Federation of Library Associations and Institutions) che ha creato un gruppo di studio dedicato all'intelligenza artificiale e nell'ottobre del 2020 ha pubblicato lo *Statement on Libraries*

* Università di Firenze, rossana.morriello@unifi.it.

and Artificial Intelligence (IFLA, 2020). Lo Statement IFLA sottolinea alcuni aspetti su cui le biblioteche dovrebbero concentrarsi in maniera prioritaria, quali l'alfabetizzazione sull'IA rivolta agli utenti, che definisce *algorithmic literacy*, il supporto alla ricerca per gli aspetti etici, di privacy e di libertà intellettuale, evidenziando come sia fondamentale la comprensione del modo in cui gli algoritmi e gli altri processi digitali impattano sulle dinamiche con le quali gli utenti accedono all'informazione.

L'intelligenza artificiale non è una tecnologia nuova ma, al contrario, è nelle nostre vite e nelle nostre biblioteche da molto tempo. Si parla di intelligenza artificiale perlomeno dagli anni Cinquanta del XX secolo, e di algoritmi da un tempo ancora più lungo, di fatto dall'antichità, se consideriamo gli algoritmi tradizionali (Louridas, 2022). Ma dal punto di vista informatico gli algoritmi dell'intelligenza artificiale sono di tipo diverso dai precedenti. Gli algoritmi informatici tradizionali sono costituiti da una serie di operazioni prevedibili e con un risultato certo, mentre gli algoritmi dell'IA sono basati su modelli statistici, per definizione probabilistici e approssimativi,¹ il cui risultato di elaborazione non è dunque prevedibile e riproducibile con assoluta precisione ed è fortemente condizionato dai dati che vengono immessi. Questo aspetto è stato reso evidente dal rilascio pubblico di ChatGPT-3 nel novembre 2022, che le persone hanno potuto sperimentare direttamente. L'IA è presente da tempo nelle numerose applicazioni integrate nei nostri telefoni, nei motori di ricerca del web, e anche negli strumenti utilizzati dalle biblioteche come, per esempio, i discovery tools,² e i sistemi di information retrieval delle banche dati. Fino a qualche tempo fa, in tali applicazioni l'integrazione dell'IA era latente, poco visibile, ne usufruivamo con scarsa consapevolezza ma negli ultimi anni il ruolo dell'IA ha cominciato a emergere, con le luci e le ombre, e a ciò ha dato ulteriore impulso la release pubblica di ChatGPT-3.

Rispetto alle biblioteche l'intelligenza artificiale rappresenta una tappa (e non di certo l'ultima) in un lungo percorso di adattamento alle innovazioni tecnologiche che parte dalla stampa a caratteri mobili e arriva alla biblioteca digitale, della quale l'intelligenza artificiale può essere considerata come un'evoluzione. Tuttavia, non vi è dubbio che l'impatto dell'IA non ha precedenti per le implicazioni sulla vita delle persone e sulla società nel suo insieme. In quanto tale, l'IA deve interessare tutte le discipline, non solo tecnologiche, ma perfino in misura maggiore le discipline umanistiche, chiamate ad avere

¹ La definizione nell'Enciclopedia Treccani di "modello statistico" è: "famiglia di meccanismi probabilistici che si presume approssimi sufficientemente bene, o addirittura contenga, il meccanismo probabilistico che ha generato i dati disponibili", 2012, <https://www.treccani.it/enciclopedia/modello-statistico_%28Dizionario-di-Economia-e-Finanza%29/>.

² Non senza criticità, si veda (Reidsma, 2019).

un ruolo attivo per promuovere un'IA etica e sostenibile e diffondere il pensiero critico che consenta alle persone di orientarsi nel mondo dell'IA. La Biblioteconomia non fa eccezione.

In questo contributo ci soffermeremo su alcuni esempi di applicazione dell'intelligenza artificiale in biblioteca, per evidenziare le potenzialità nel settore bibliotecario, il quale non può però prescindere da una riflessione sui limiti e i rischi dell'IA.

L'intelligenza artificiale in biblioteca

Il tema dell'intelligenza artificiale in biblioteca è dunque molto ampio e naturalmente ancora in corso di sviluppo, ma alcune applicazioni già realizzate o in fase sperimentale rendono conto delle principali direzioni verso le quali ci si sta muovendo. Nel presente paragrafo ci si soffermerà su alcune applicazioni, senza pretesa di esaustività, ma semplicemente a titolo esemplificativo.

I SISTEMI DI INDICIZZAZIONE AUTOMATICA

Uno degli ambiti in cui le biblioteche si stanno muovendo velocemente e da più tempo è quello dell'intelligenza artificiale applicata all'indicizzazione catalografica. Gli esempi potrebbero essere tanti ma una delle applicazioni più avanzate è ANNIF, il sistema di indicizzazione semantica automatica sviluppato dalla Biblioteca Nazionale di Finlandia, e basato su una combinazione di strumenti di natural language processing e machine learning.³ ANNIF è usato anche da diverse istituzioni al di fuori della Finlandia. Tra queste il Leibniz Information Centre for Economics. Il sito del Centro spiega nel dettaglio il processo di indicizzazione automatica e nel catalogo è possibile visualizzare il risultato dell'elaborazione impostando una ricerca. Ogni ora l'algoritmo interroga il catalogo della biblioteca per fare l'harvesting delle nuove accessioni e automaticamente attribuisce un soggetto alle pubblicazioni dopo aver condotto un processo di text mining sui campi titolo, autore e abstract. In questa fase, dunque, l'intervento del bibliotecario non è previsto. Tuttavia, ciò non significa che la figura del bibliotecario catalogatore non sia più necessaria, al contrario il nuovo ruolo e le nuove responsabilità a cui è chiamato diventano cruciali nella società algoritmica. L'intervento umano avviene, infatti, nella fase di annotazione del dataset, di *data curation* del sistema di knowledge organization, di verifica nei casi in cui l'algoritmo non restituisce un risultato con una determinata percentuale di certezza. L'indicizzazione automatica presenta

³ Sul sito è disponibile una demo <<https://annif.org/>>. Si veda anche (Suominen, 2022).

sovente risultati imperfetti e non considerati validi dai bibliotecari che devono intervenire per correggere gli errori. In realtà si tratta, quindi, di una forma di indicizzazione semi-automatica. Come per tutti i processi di machine learning, la capacità del sistema di restituire un'indicizzazione corretta migliorerà con il tempo poiché il modello apprende man mano che i bibliotecari apportano le correzioni e gli insegnano a evitare gli errori e la performance si perfezionerà quanto più l'algoritmo potrà disporre di grandi quantità di dati bibliotecari per il training, se si sceglie di non ricorrere a dati esterni. Considerando che i dati bibliotecari non sono di big data, i modelli algoritmici hanno ovviamente necessità di tempi maggiori per l'allenamento. Un sistema automatico, o anche semi-automatico, può migliorare considerevolmente il risultato del processo di indicizzazione in termini di possibilità di recupero dell'informazione da parte dell'utente, soprattutto con l'arricchimento dei metadati esistenti, e permette di stabilire relazioni più efficaci tra le risorse bibliografiche (Golub, 2021).

REFERENCE E INFORMAZIONI DI COMUNITÀ

Le applicazioni riconducibili alle attività di reference sono numerose e vanno dalle più semplici chatbot, in uso ormai da tempo, ai recenti robot che si muovono all'interno degli spazi bibliotecari interagendo con gli utenti. Anche alla base di queste applicazioni vi sono tecniche di machine learning che consentono ai bot e ai robot di apprendere la struttura del linguaggio e di formulare domande e risposte dall'interazione con gli utenti. I vantaggi, soprattutto nel caso delle chatbot, sono innanzitutto la disponibilità 24 ore su 24 e sette giorni alla settimana, cui si aggiunge la possibilità di addestrare tali sistemi a rispondere in diverse lingue. Un ulteriore aspetto da non sottovalutare è la familiarità crescente con le chat e le chatbot dei cittadini, e in particolare dei giovani, per i quali la chat è ormai una modalità comunicativa molto diffusa, dalle applicazioni sul telefono alle modalità di interazione con i servizi di ogni genere (dal supermercato alla banca). Anche per tale tipo di attività il bibliotecario non viene sostituito dalla chatbot ma questa può rappresentare un ausilio nelle attività di comunicazione con gli utenti e può essere vista come un'opportunità per liberare la bibliotecaria dai compiti routinari. Per esempio, le chatbot possono rispondere alle domande più semplici lasciando le risposte più complesse alla bibliotecaria. Inoltre, la supervisione della bibliotecaria è sempre necessaria perché la chatbot può non saper rispondere a certe domande se non sono previste nella sua base dati o se sono troppo complesse, e può commettere errori. L'interazione con gli utenti consente ai bot di migliorare la capacità di risposta ma non si può prescindere (ancora) dall'intervento umano. Inoltre, il contatto automatizzato con gli utenti consente alla biblioteca di rilevare informazioni utili, per esempio riguardo alla distribuzione temporale delle richieste, alla dimensione spaziale nel caso dei robot in biblioteca, e agli

stati d'animo se sul sistema di IA sono implementate opzioni di *sentiment analysis*. Analisi di questo genere consentirebbero di organizzare attività di promozione della lettura mirate e permetterebbero una maggiore conoscenza della comunità, un elemento di analisi basilare nelle attività di sviluppo delle raccolte.

Diversi studi mostrano come una variabile importante nella scelta delle letture sia l'umore, ovvero come vi sia una correlazione tra le scelte dei libri e gli specifici momenti della nostra vita, gli stati d'animo, l'umore del momento (Sheldrick Ross, 2001). Lo abbiamo sperimentato tutti nella nostra vita: ci sono libri che se proviamo a leggere in certi momenti non ci appassionano o ci respingono ma magari in un momento successivo li apprezziamo. Gli stati d'animo come stanchezza, stress, bisogno di conforto o al contrario gioia, rilassatezza, spensieratezza influenzano le nostre scelte di lettura. Per esempio, quando i lettori hanno bisogno di rassicurazione, conforto, sicurezza tendenzialmente leggono i classici oppure libri nuovi ma di autori che già conoscono. In periodi di stress o di grandi impegni si preferiscono libri più brevi, letture facili e libri che piaceranno con certezza perché ben noti o anche già letti. In queste fasi, i libri dell'adolescenza sono la quintessenza del *comfort reading* (Sheldrick Ross, 2001). Quando il periodo della vita è calmo e rilassato si sperimenta di più con materiale complesso e meno noto. Nei periodi di umore solitario si legge la poesia, oppure nei momenti tristi si scelgono libri umoristici e quando si è stanchi libri leggeri. Le scelte sono dunque molto legate all'umore, particolarmente influente per le opere di fiction, in quanto la saggistica la si legge spesso per studio o per lavoro. Tale variabile è difficilmente rilevabile dalla bibliotecaria, anche nel caso di transazioni di reference in persona. Da alcuni studi emerge che il lavoro svolto dalle macchine sembra essere più efficace nel giudicare le personalità degli utenti di quanto accada nel mondo reale (Youyou, W., M. Kosinskib, & D. Stillwell, 2015).

La sentiment analysis nelle applicazioni di intelligenza artificiale può aiutare in molti casi. Ci sono chatbox *emoji-friendly* che consentono all'utente di indicare il suo stato d'animo tramite un emoji e quindi alla macchina di apprendere la correlazione tra una scelta e uno stato d'animo. Per esempio, l'app per le raccomandazioni di lettura Obotti (Helsinki Times, 2021), sulla quale ritorneremo più avanti, prevede un'opzione di scelta in cui chiede all'utente «uno di questi due libri stimola qualche sentimento?». L'analisi delle scelte potrebbe favorire gli studi sui comportamenti di lettura. Tale analisi potrebbe avvalersi di dati ancora più significativi laddove in biblioteca fossero disseminati dei sensori oppure dispositivi, come avviene con i robot, che stimolano la scelta rilevando attraverso un'interazione diretta lo stato d'animo del lettore.

Il robot è una versione più elaborata, e soprattutto antropomorfizzata, di un bot per il reference, in grado di agevolare l'interazione con quei pubblici che non accedrebbero comunque a una chatbot su web, come i bambini. Proprio su questo aspetto si è basata la promozione dell'imponente Muhammad Bin Rashid Library di Dubai inaugurata a giugno 2022 (figura 1) relativa al robot Pepper (Barda, 2022). La Bin Rashid non è l'unica biblioteca ad avere il robot umanoide. Pepper è presente anche in altre biblioteche, tra le quali la Biblioteca centrale e universitaria di Lucerna e la Biblioteca comunale di Hannover.⁴ La pagina informativa della Softec, l'azienda che lo produce, spiega come «il social robot risponde a numerose domande relative alla biblioteca stessa, fornisce indicazioni precise, ma è anche lieto di accompagnare l'utente verso lo scaffale desiderato. Pepper è sempre aggiornato sulle iniziative in atto e gli eventi in agenda, invitando i visitatori a parteciparvi». Tramite una app l'utente può interrogare il robot e «grazie al collegamento alla banca dati dei volumi conservati nella biblioteca, il social robot è in grado, nel giro di pochi secondi, di suggerire contenuti in linea con la richiesta dell'utente» (Poli, 2021), e di interagire in venti lingue, offrendo altri servizi personalizzati. Analoghe funzioni le svolge Libby, un altro modello di robot introdotto alla Biblioteca universitaria di Pretoria in Sud Africa (University of Pretoria, 2019; Thekiso, 2019). Ma si tratta solo di due dei numerosi tipi di robot, umanoidi e non, diffusi nelle biblioteche e usati per innumerevoli attività quali la ricollocazione dei libri a scaffale, il supporto all'apprendimento scolastico degli studenti, la lettura di storie ai ragazzi, i corsi di coding (Tella & Ajani, 2022). Oltre all'interazione diretta con gli utenti, l'uso di sensori consente al robot di raccogliere ed elaborare ulteriori informazioni in rapporto all'ambiente, ovvero agli spazi della biblioteca, offrendo ai bibliotecari possibilità di analisi delle relazioni tra spazi e servizi.

⁴ Zentral- und Hochschulbibliothek Luzern, Der Roboter hilft, <<https://www.zhbluzern.ch/aktuelles/der-roboter-hilft-2438>>; Hannover Stadtbibliothek <<https://www.hannover.de/Leben-in-der-Region-Hannover/Bildung/Bibliotheken-Archive/Stadtbibliothek-Hannover>>.



Fig. 1. Muhammad Bin Rashid Library, Dubai (Barda, 2022)

LIBRO ANTICO E CONSERVAZIONE

Le applicazioni dell'IA al libro antico sono numerose, a cominciare dall'ausilio nella digitalizzazione dei testi con la possibilità di riconoscere più agevolmente la scrittura a mano, come avviene con Transkribus.⁵ La conservazione del materiale è un settore in cui l'intelligenza artificiale può portare grandi benefici come dimostra il progetto DALGOCOL (Fouille de Données et ALGOritmes de prédiction de l'état des COLlections) (Vallas, 2022; Zreik & Zoubida, 2022; Cabane & Zreik, 2022; Lee, Yamoun, Jordan, Simon Chane, Picard, & Longhi, 2022). Operando sul riconoscimento di immagini, con l'ausilio di standard come IIIF, è possibile rilevare dati sull'età e sulla natura dei materiali, sugli eventi di degrado e sulle cause, sugli interventi effettuati, e in questo modo l'algoritmo è in grado di prevedere l'evoluzione della condizione fisica di un libro e quindi consentire interventi tempestivi. Sulle immagini la Bibliothèque National de France lavora da molto tempo con il progetto Gallica Snoop, basato proprio su IIIF e sulle tecniche di image retrieval.⁶

⁵ Transkribus <<https://readcoop.eu/it/transkribus/>>.

⁶ BnF <<https://www.bnf.fr/fr/quelques-projets-dintelligence-artificielle-en-cours-la-bnf>>.

PROMOZIONE DELLE RACCOLTE E DELLA LETTURA

Tra le applicazioni più interessanti delle tecniche di intelligenza artificiale per un sistema di raccomandazione, vi è senza dubbio Obotti, la app della Biblioteca Oodi di Helsinki. La Oodi Library è la biblioteca centrale del sistema urbano della capitale finlandese che comprende 37 biblioteche. La biblioteca è stata aperta a dicembre 2018, con un progetto realizzato dallo studio architettonico ALA che si distingue per le scelte innovative in termini di sostenibilità e per il quale è stato insignito del Green Library Award promosso dalla sezione IFLA ENSULIB.

Obotti è un sistema di raccomandazione progettato dalla HeadAI, introdotto alla Oodi Library nel 2018 ma lanciato ufficialmente in occasione dell'inaugurazione della nuova biblioteca nel 2021 (rimandata a causa della pandemia di Covid-19), dopo un periodo di sperimentazione. Si tratta di un'applicazione per dispositivi mobili disponibile sia per il sistema operativo Android che Apple. L'implementazione dell'applicazione ha visto la collaborazione tra gli informatici di HeadAI e le bibliotecarie della Oodi. Le bibliotecarie hanno creato i profili dei bot associando a ciascun assistente virtuale i libri di pertinenza, per permettere all'algoritmo di avere una base di partenza. In una prima fase sono state utilizzate le parole chiave (tag) che erano associate ai libri in catalogo (2-3 inizialmente ora 20-30). Le parole chiave sono basate sul sistema di classificazione in uso alla Oodi Library, specifico del sistema bibliotecario urbano della città di Helsinki, che è una personalizzazione all'interno del sistema di classificazione finlandese FINTO (Suominen & Inkinen, 2022). Il sistema si alimenta poi con i dati rilevati dai comportamenti degli utenti che usano la app, nel rispetto della privacy che è uno degli elementi che ha posto delle limitazioni all'uso dei dati, per esempio in relazione ai dati personali e ai dati sui prestiti degli utenti che non vengono utilizzati.⁷

Il periodo di sperimentazione ha consentito di testare il prodotto e di affinarlo, rivedendo alcune scelte che erano state compiute inizialmente. Una di queste ha riguardato proprio la ridefinizione dei profili. Nella prima implementazione i bot avevano nomi finlandesi e specifiche molto generiche e si è deciso di attribuire dei nomi che fossero comprensibili anche in inglese, considerando la scelta di rendere disponibile l'applicazione anche in questa lingua, oltre a finlandese e svedese. Inoltre, dopo l'avvio della sperimentazione, e prima del lancio, le bibliotecarie si sono accorte che i profili erano troppo generici e quindi li hanno ridefiniti attribuendo loro maggiore personalità

⁷ Le informazioni mi sono state fornite direttamente in un'intervista tramite Zoom da una delle bibliotecarie della Oodi che si sono occupate di Obotti, Paula Laaksonen, che ringrazio.

(figura 2). Gutenberg è il profilo generale, di comunità, e quello maggiormente collegato al filtraggio collaborativo, ovvero ai comportamenti degli utenti che usano la app. Oltre che ai dati della biblioteca, Gutenberg attinge ai dati forniti dal comune, dal servizio di noleggio delle biciclette, un cui parcheggio si trova accanto alla biblioteca, e da altre fonti integrate.

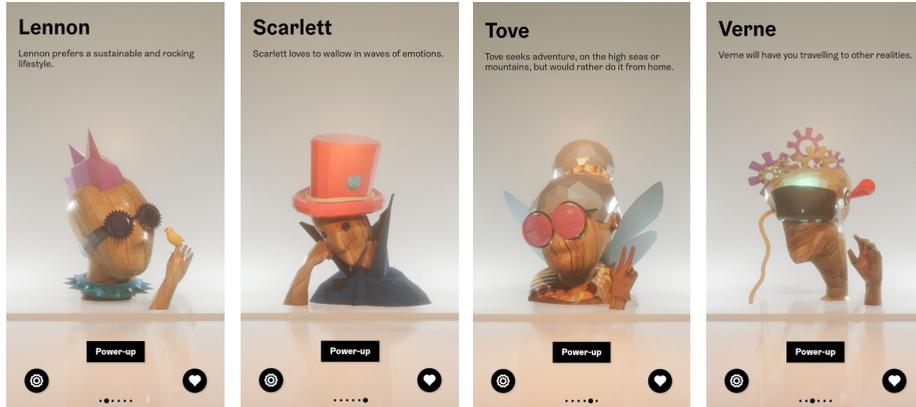


Fig. 2. I profili dell'applicazione Obotti della Oodi Library di Helsinki

L'implementazione iniziale di Obotti è stata resa possibile dalla presenza di un catalogo strutturato e di un sistema di classificazione per la fiction, la quale non è sempre classificata nelle biblioteche. Inoltre, il catalogo offriva già diverse parole chiave per ciascun record che sono state progressivamente aumentate per permettere un arricchimento del dato, elemento fondamentale nell'uso dell'IA. La risposta del pubblico all'introduzione di Obotti è stata molto positiva.

I sistemi di raccomandazione

L'ambito delle raccomandazioni di lettura è una delle attività più interessanti di applicazione dell'IA nelle biblioteche. La lettura è un'attività che produce grandi benefici in termini di benessere mentale e fisico a livello individuale e sociale (Vivarelli, 2018). In particolare la lettura di opere di fiction aumenta l'empatia, il pensiero critico, l'apertura mentale necessaria a processare efficacemente le informazioni tanto che è entrata a far parte della cultura organizzativa d'impresa e alcune aziende introducono percorsi formativi sulla lettura e discussione critica di testi letterari per sviluppare le competenze dei lavoratori e dunque il rendimento dell'impresa (Seifert, 2020). La lettura è cambiata, è diventata sociale (Cordón García, & Gómez Díaz, 2019) e

i sistemi di raccomandazione intercettano tale caratteristica. I sistemi di raccomandazione si basano su tecniche di machine learning e vengono allenati (training) per suggerire agli utenti le letture, in un processo che è molto simile a quello dell'apprendimento della lettura nelle persone, e si affina con il tempo nel corso della vita, perlomeno per quanto riguarda la lettura di narrativa. La capacità di scegliere il libro giusto, ovvero il libro che con molta probabilità ci piacerà, è il risultato di un allenamento che comincia dall'infanzia e si affina tramite continue scelte, a volte giuste altre sbagliate. Ogni nuova lettura affina la nostra capacità di giudizio. Ogni nuova lettura che diventa un'esperienza piacevole invoglia il lettore a continuare quell'esperienza e a leggere di più. Al contrario ogni lettura sbagliata, diminuisce la probabilità di cercare un altro libro da leggere. In un'indagine su un campione selezionato di lettori forti, coloro che leggono abitualmente per piacere come scelta e non occasionalmente per informazione hanno dichiarato che leggere per piacere è una parte molto importante della loro vita (Sheldrick Ross, 2001). Il campione era costituito da 194 interviste intensive e articolate nelle quali tra le altre domande si cercava di capire come i lettori intervistati scegliessero i libri da leggere per piacere. I risultati emersi hanno mostrato il processo. Tendenzialmente i lettori forti da bambini sono lettori voraci di qualsiasi libro sia disponibile in casa oppure in biblioteca, ricevuto in regalo o in prestito da amici e parenti. Il campo di scelta è comunque in un certo qual modo limitato alle collezioni di libri per ragazzi e quindi presenta meno rischi di una scelta sbagliata. Man mano che si cresce aumenta l'offerta documentaria ma nel frattempo la capacità di scelta è stata allenata a selezionare libri che si presume piaceranno. Se in età adolescenziale ci possono ancora essere dei dubbi e casi in cui si è insoddisfatti, arrivati all'età adulta si è sviluppata una certa capacità di individuare con facilità il libro che piacerà da alcuni elementi precedenti ed esterni alla lettura, paratestuali. Per i lettori deboli o i non lettori la situazione è differente. I lettori occasionali, o ancor più i non lettori, non hanno attraversato il processo di allenamento alla scelta e di conseguenza la scelta per loro è più difficoltosa e maggiormente a rischio di errore. Per un nuovo lettore scegliere un libro sbagliato può costituire un disincentivo a proseguire, mentre scegliere un libro giusto può invece avviare verso nuove letture e invogliare a continuare la scoperta.

La capacità di scegliere il libro giusto si basa su una serie di dati che il lettore può avere a disposizione per orientare la scelta e capire a priori se il libro può essere di suo interesse. I dati derivano da varie fonti interne al libro stesso, come la sintesi presente sulla copertina e le altre informazioni disponibili nel libro, e fonti esterne come le recensioni pubblicate su riviste e online, i consigli degli amici anche questi ormai mediati spesso da internet e dai social network, le reti sociali che assolvono un po' la funzione dei vecchi book club (per esempio, specifiche comunità virtuali su generi come la fantascienza, o su determinate tematiche), la trasposizione di libri in film o

serie televisive. Infine, i consigli di librai e bibliotecari e le forme di mediazione adottate per presentare i libri e invogliare alla lettura, come, in biblioteca, i sistemi di indicizzazione semantica oppure, più frequentemente per le opere di narrativa, la suddivisione in scaffali tematici, l'esposizione delle novità, la creazione di bibliografie tematiche per gli utenti.

Questi strumenti forniscono dati e metadati che il lettore elabora durante tutto il corso della vita. In particolare, ciò che i bibliotecari compiono da sempre è un'attività di promozione delle collezioni, ovvero di comunicazione della collezioni al fine di favorire l'incontro del lettore con il libro, o meglio con il libro giusto, in accordo con le leggi di Ranganathan, "a ogni lettore il suo libro" e "a ogni libro il suo lettore". L'intelligenza artificiale, e in particolare i sistemi di raccomandazione, rappresentano una facilitazione di queste due leggi e soprattutto della quarta legge "risparmia il tempo del lettore".

L'intelligenza artificiale attraverso le tecniche di machine learning è in grado di compiere il processo di affinamento dei gusti di lettura che un lettore sviluppa nel corso della vita, ma in tempi molto più brevi perché elabora grandi quantità di dati con grande velocità. Come una persona abituata alla lettura dopo un certo allenamento riesce a capire se un libro può essere di suo gradimento solo da pochi dati e metadati paratestuali, quali il nome dell'autore, il titolo del libro, l'editore, il genere, e così via, così l'IA sulla base di dati e metadati è in grado di suggerire le letture. Alla macchina non serve porsi le domande su chi sia il lettore, quale sia il contenuto del testo, cosa cerca il lettore, ma gli schemi che è in grado di evidenziare grazie alle grosse quantità di dati processati velocemente consentono di rilevare le correlazioni che rispondono alle domande per quello specifico profilo di lettore.

Conclusioni

Gli ultimi sviluppi dell'IA, e in particolare l'IA generativa, hanno innescato quella che viene definita esplosione cambriana della tecnologia, per la quantità e la varietà di applicazioni e sperimentazioni che sta producendo. Le biblioteche sono già coinvolte in questa esplosione e lo saranno in misura crescente in quanto gestiscono da sempre dati e metadati, a cominciare da quelli catalografici, e sono istituzioni sociali dunque non esenti dalle trasformazioni che riguardano la società. L'intelligenza artificiale è una trasformazione dal grosso impatto sulla società e sulle persone. Gli esempi che abbiamo riportato sopra mostrano alcune delle potenzialità dell'IA in biblioteca che appaiono enormi ed estremamente interessanti. Tuttavia, ciascuna delle applicazioni presenta numerose zone d'ombra, limiti e rischi che le biblioteche devono affrontare con consapevolezza. Prima fra tutte la sostenibilità economica dei progetti, che difatti in molti casi sono realizzati con l'ausilio di finanziamenti privati.

L'attenzione alla questione della privacy, le implicazioni relative al copyright, i bias che l'intelligenza artificiale porta con sé (Morriello, 2023) sono altre questioni cruciali. Le stesse collezioni bibliotecarie presentano bias poiché sono costruite da una stratificazione storica che ha attraversato varie epoche con idee e convinzioni differenti. Inoltre, i dati delle biblioteche non sono sempre sufficientemente strutturati per essere utilizzati dai modelli di intelligenza artificiale e devono essere integrati con dataset esterni che aumentano il rischio di bias. Come garantire la qualità e congruità di dati e metadati bibliotecari è dunque un tema fondamentale. Le collezioni andranno ripensate in base al modo in cui si andrà ridefinendo l'universo bibliografico. Si parla da tempo di "collezioni come dati" e della necessità di ragionare in termini che includano le nuove modalità di ricerca e i nuovi utenti delle collezioni bibliotecarie che non saranno solo le persone ma anche gli agenti di intelligenza artificiale. Anche gli utenti si aspetteranno nuove modalità di accesso alle raccolte delle biblioteche, basate sulle esperienze di interazione con le applicazioni dell'intelligenza artificiale che già oggi caratterizzano l'esperienza digitale quotidiana. Come cambierà il ruolo del bibliotecario e quale formazione sarà necessaria per renderlo in grado di affrontare queste grandi sfide sono le domande prioritarie alle quali le biblioteche devono cercare di dare una risposta nel breve periodo, da un lato diventando attori del cambiamento e dall'altro affrontandolo con una preparazione adeguata.

Riferimenti bibliografici

- Barda, T. (2022), 'Look! A Parent's Guide to the Mohammed Bin Rashid Library Opening: From robot helpers to holograms and play areas, what you'll find for kids and teens in Dubai's awesome new library', June 14 2022, <<https://schoolscompared.com/uncategorized/news-stories/look-a-parents-guide-to-the-mohammed-bin-rashid-library-opening-from-robot-helpers-to-free-study-spaces-and-play-areas-what-youll-find-for-kids-and-teens-in-dubais-awesome-new-l/>>.
- Cabane, C., & Zreik, A. (2022), *AI and Data to Improve Preservation Management*, IFLA 87th IFLA World Library and Information Congress, <<https://repository.ifla.org/handle/123456789/2196>>.
- Cordón García, J. A., & Gómez Díaz, R. (coords.) (2019), *Lectura, sociedad y redes. Colaboración, visibilidad y recomendación en el ecosistema del libro*. Madrid: Marcial Pons.
- Cox, A.M. & Mazumdar, S. (2022), 'Defining artificial intelligence for librarians', *Journal of Librarianship and Information Science*, online first, DOI <<https://doi.org/10.1177/09610006221142029>>.
- Golub, K. (2021), 'Automated Subject Indexing: An Overview', *Cataloging & Classification Quarterly*, vol. 59, n. 8, pp. 702-719, DOI <<https://doi.org/10.108>>.

- 0/01639374.2021.2012311>.
- Helsinki Times (2021), 'Oodi's artificial intelligence app Obotti finds reading material for customers', *Helsinki Times*, 17 June, <<https://www.helsinkitimes.fi/themes/themes/science-and-technology/19426-oodi-s-artificial-intelligence-app-obotti-finds-reading-material-for-customers.html>>.
- IFLA (2020), 'Statement on Libraries and Artificial Intelligence', <<https://repository.ifla.org/handle/123456789/1646>>.
- Lee, V., Yamoun, L., Jordan, M., Simon Chane, C., Picard, D., Longhi, J. (2022), "Deep-Learning Technology for Book Conservation Assessment", in: *Libraries and Archives, Heritage for the Future, Science for Heritage, Un patrimoine pour l'avenir, une science pour le patrimoine*, Mar 2022, Paris, France, <<https://hal.science/hal-03620294>>.
- Louridas, P. (2020), *Algorithms*. Cambridge (Mass.). MIT Press.
- Morriello, R. (2019), "Blockchain, intelligenza artificiale e internet delle cose in biblioteca", *AIB Studi*, vol. 59, n. 1-2, pp. 45-68, DOI <<https://doi.org/10.2426/aibstudi-11927>>.
- Morriello, R. (2023), 'Dati e metadati bibliotecari per l'intelligenza artificiale: per un'agency delle biblioteche', *Biblioteche oggi Trends*, vol. 9, n. 1, pp. 38-45, DOI <<https://doi.org/10.3302/2421-3810-202301-038-1>>.
- Oyelude, A.A. (2021), 'AI and libraries: trends and projections', *Library Hi Tech News*, vol. 38, n. 10, pp. 1-4, DOI <<https://doi.org/10.1108/LHTN-10-2021-0079>>.
- Poli, V. (2021), 'Pepper: una ventata d'aria fresca per le biblioteche', *Softec Blog*, 22 aprile, <<https://blog.softecspa.com/pepper-aria-fresca-per-le-biblioteche>>.
- Reidsma, M. (2019), *Masked by trust: Bias in Library Discovery*. Sacramento: Library Juice Press.
- Seifert, C. (2020), 'The Case for Reading Fiction', *Harvard Business Review*, March 6, <<https://hbr.org/2020/03/the-case-for-reading-fiction>>.
- Sheldrick Ross, C. (2001), "Making Choices: What Readers Say About Choosing Books to Read for Pleasure", in: *Readers, Reading and Librarians*, edited by Bill Katz. New York, The Haworth Information Press, pp. 5-21.
- Suominen, O. J. Inkinen, & L. Mona (2022), 'Annif and Finto AI: Developing and Implementing Automated Subject Indexing', *JLIS.It*, vol. 13, n. 1, pp. 265-82, DOI <<https://doi.org/10.4403/jlis.it-12740>>.
- Tella, A. & Ajani, Y. A. (2022), 'Robots and public libraries', *Library Hi Tech News*, vol. 39, n. 7, pp. 15-18, DOI <<https://doi.org/10.1108/LHTN-05-2022-0072>>.
- Thekiso, R. (2019), 'University robot 'Libby' brings the future library closer', *University World News*, 12 July, <<https://www.universityworldnews.com/post.php?story=20190708105508294>>.
- University of Pretoria, Department of University Relations (2019), 'Libby the library robot is on duty at UP', 18 July, <https://www.up.ac.za/alumni/news/post_2825427-libby-the-library-robot-is-on-duty-at-up>.
- Vallas, P. (2022), 'Prédire l'état matériel des documents: Dalgocol, un programme de recherche en intelligence artificielle à la BnF', *Bulletin des bibliothèques de*

France (BBF), n. 1, <<https://bbf.enssib.fr/consulter/bbf-2022-00-0000-008>>.

Vivarelli, M. (2018), *La lettura. Storie, teorie, luoghi*. Milano: Editrice Bibliografica.

Youyoua, W., Kosinskib, M., & Stillwell, D. (2015), 'Computer-based personality judgments are more accurate than those made by humans', *Proceedings of the National Academy of Sciences (PNAS)*, vol. 111, n. 4, pp. 1036-1040.

Zreik, A., & Kedad, Z. (2022), 'Matching and analysing conservation-restoration trajectories', *Data & Knowledge Engineering*, vol. 139, n. 102015, DOI <<https://doi.org/10.1016/j.datak.2022.102015>>.

L'INTELLIGENZA ARTIFICIALE ED IL PROBLEMA DELL'AGENCY DAL PUNTO DI VISTA BIBLIOGRAFICO

Maurizio Lana*

Il 30 novembre 2022 OpenAI, l'azienda che sviluppa il sistema di Intelligenza Artificiale GPT (Generative Pretrained Transformer, giunto alla versione 4 alla data di scrittura di questo contributo), annunciava la disponibilità al pubblico dell'interfaccia ormai notissima denominata ChatGPT. In pochi mesi l'Intelligenza Artificiale da argomento esoterico per informatici è diventata oggetto di interesse in ambiti diversissimi e quotidiani, in cui l'interesse verso i sistemi di Intelligenza Artificiale è di tipo applicativo e non solo di studio o di sviluppo. La diffusione di interesse ha preso due direzioni principali: possibili utilizzi di ChatGPT in ambito aziendale per ricevere o fornire servizi; e (anche in relazione ad essi) pericoli insiti nell'Intelligenza Artificiale¹ nella forma (Chat)GPT². Ad esse se ne aggiunge una terza, sottotraccia, non meno importante: «quali cambiamenti porta un sistema di Intelligenza Artificiale come GPT quando viene utilizzato nella produzione di contenuti intellettuali (testuali, musicali, visivi, ...)»?». Già oggi contenuti che si collocano nell'ambito della creatività intellettuale vengono prodotti con l'impiego di sistemi di Intelligenza Artificiale³; e certamente i sistemi di Intelligenza Artificiale

* Università del Piemonte Orientale.

¹ Dalla minaccia nella quotidianità per la scomparsa di molti lavori, alla minaccia esistenziale per il genere umano.

² Scrivo (Chat)GPT perché il sistema di Intelligenza Artificiale non è ChatGPT ma GPT; come si è già ricordato ChatGPT è solo un'interfaccia che facilita l'uso di GPT: chi vuole può usare direttamente GPT; e si potrebbero sviluppare (anche) altre interfacce per l'uso di GPT; oppure delle 'sottoversioni' di ChatGPT focalizzate su ambiti specifici come Botsonic, di Writesonic (<https://writesonic.com/Chat>) o per specifiche piattaforme come IOS o Android; o alte interfacce verso altri sistemi di Intelligenza Artificiale come StableVicuna di Stability AI (<https://stability.ai/blog/stablevicuna-open-source-rlhf-chatbot>).

³ Ad esempio a proposito di Spotify da un lato si sospetta che la musica di «The Groove Trio» sia prodotta con sistemi di Intelligenza Artificiale dalla piattaforma stessa; cfr. <https://www.reddit.com/r/Jazz/comments/10px4gz/is_spotify_using_ai_to_generate_fake_jazz_artists/>; dall'altro Spotify cancella centinaia di brani che vi

entreranno *routine* nella produzione di contenuti.

Si potrebbe pensare che la questione dell'agency dei sistemi di Intelligenza Artificiale dal punto di vista bibliografico, che si iscrive in questa terza direzione, sia un caso di studio puramente teorico. Non credo che sia così, perché la produzione di contenuti testuali da parte di sistemi di Intelligenza Artificiale da un lato costituisce l'evoluzione di strumenti del lavoro intellettuale in uso quotidianamente; dall'altro questa produzione sta già avvenendo, da tempo.

Molti, soprattutto informatici ma non solo, sottolineano che i sistemi di Intelligenza Artificiale vengono oggi sviluppati, testati e utilizzati in un contesto sostanzialmente privo di regole, norme, leggi. È un problema non nuovo, che riguarda lo sviluppo di sistemi digitali che si basano sull'interazione con le persone e che da acquisiscono dati di ogni genere; la società civile, e gli Stati, non possono accettare che sistemi di tale potenza, di tale capacità di incidere sul funzionamento della società in tutti i suoi aspetti, operino al di fuori di qualsiasi regolamentazione⁴ definita nella prospettiva dell'interesse pubblico e non dei suoi sviluppatori e ci sono indicazioni forti della consapevolezza che occorrono norme e leggi.⁵

L'Intelligenza Artificiale come evoluzione di strumenti del lavoro intellettuale

È ovvio che leggere degli sviluppi di sistemi di Intelligenza Artificiale capaci di operare nella scrittura di testi, e di testi di prospettiva scientifica, possa suscitare (sollecitare?) perplessità e preoccupazione, perché lo studioso sente la scrittura scientifica come un'attività eminentemente personale in cui le sue capacità sono messe alla prova in una forma limite e, attraverso la scrittura, vengono valutate in una varietà di forme e contesti. E dunque pensare che possano intervenire dei sistemi di Intelligenza Artificiale 'in aiuto' può apparire

erano stati caricati, perché prodotti con il sistema di Intelligenza Artificiale Boomy (Nicolaou, 2023).

⁴ Sul tema si veda (Giraudo, 2022): «'Legal bubbles' tend to arise when economic agents invest in the economic exploitation of a new resource in a context of uncertainty and ignorance about the legal implications of innovative activities. These legal foundations may eventually turn out to be unstable as a result of courts' ex-post attempt to re-adapt them to the previously ignored implications of unconstrained commodification, with disruptive economic consequences.» (p. 587).

⁵ Si veda ad esempio, per l'Europa, il percorso legislativo della proposta di legge europea (regulation) nota come «AI Act» (European Parliament, 2021) e della proposta di direttiva nota come «AI Liability Directive» (European Commission, 2022) (*regulations* sono leggi che una volta approvate dal Parlamento europeo diventano operative e vincolanti in tutti paesi membri; *directives* sono indicazioni normative che devono essere recepite dalle legislazioni nazionali entro un tempo dato).».

deontologicamente scorretto, o sminuente, tanto più che si può andare da forme elementari di intervento di sistemi di Intelligenza Artificiale già disponibili, fino a forme più complesse e 'invasive' che sono pronte o in arrivo. Il punto è che già oggi la maggior parte degli studiosi utilizza sistemi di potenziamento delle sue capacità intellettuali. Ne parlò per primo Douglas Engelbart negli anni '60 del secolo scorso utilizzando l'espressione *augmenting human intellect* nel titolo di due documenti progettuali: «Augmenting human intellect: a conceptual framework», del 1962⁶, e «A research center for augmenting human intellect», del 1968⁷. L'idea chiave era liberare le persone dai compiti esecutivi/ripetitivi, da demandare ai computer, creando un contesto che Engelbart descriveva nel 1962 con queste parole:

By “augmenting human intellect” we mean increasing the capability of a man to approach a complex problem situation, to gain comprehension to suit his particular needs, and to derive solutions to problems. Increased capability in this respect is taken to mean a mixture of the following: more-rapid comprehension, better comprehension, the possibility of gaining a useful degree of comprehension in a situation that previously was too complex, speedier solutions, better solutions, and the possibility of finding solutions to problems that before seemed insoluble. And by “complex situations” we include the professional problems of diplomats, executives, social scientists, life scientists, physical scientists, attorneys, designers – whether the problem situation exists for twenty minutes or twenty years. We do not speak of isolated clever tricks that help in particular situations. We refer to a way of life in an integrated domain where hunches, cut-and-try, intangibles, and the human “feel for a situation” usefully co-exist with powerful concepts, streamlined terminology and notation, sophisticated methods, and high-powered electronic aids.⁸

Dallo sviluppo di questa visione del 1962, ma che potrebbe benissimo essere riferita oggi all'impiego di sistemi di Intelligenza Artificiale in molte delle attività quotidiane di un *knowledge worker*, ebbero origine i computer personali e le loro interfacce grafiche che tutti usiamo abitualmente e senza i quali molte persone si sentirebbero impossibilitate ad operare. Dal 1962, anno in cui Engelbart formulò il concetto, al 1968 in cui ne fece la prima dimostrazione pubblica⁹, passarono molti anni prima che nel 1983 la rivista Time dedicasse la sua copertina al personal computer¹⁰, a dirne l'avvenuto

⁶ (Engelbart, 1962).

⁷ (Engelbart, English, 1968).

⁸ (Engelbart, 1962, p. 1).

⁹ (Wikipedia, 2021).

¹⁰ (Time Magazine, 1983).

inserimento nella società.

Che cosa intendo dire con questo? Due cose: che noi da tempo usiamo strumenti informatici di potenziamento delle nostre capacità operative nel lavoro intellettuale¹¹ e i sistemi di Intelligenza Artificiale si collocano in questa linea; e che da qui a vent'anni sarà diventata ordinaria la questione che in questo contributo pongo all'attenzione di chi legge cioè quella dell'autorialità dei testi scritti ricorrendo in toto o in parte all'impiego di sistemi di Intelligenza Artificiale. L'argomento di questo contributo è dunque un primo insight dentro un tema che riguarderà la quotidianità della scrittura scientifica e dunque dovrà essere oggetto della riflessione bibliografica.

I sistemi di Intelligenza Artificiale entrano nella produzione di *quali* contenuti testuali¹²? In questo tempo in cui pare che sia vero che Intelligenza Artificiale=ChatGPT, occorre prima di tutto ricordare che il l'universo dei sistemi di Intelligenza Artificiale è ben più vasto e che si va da quelli ad hoc per specifici compiti, di cui non si parla mai nella comunicazione non specializzata; a quelli di uso generale come GPT di cui ChatGPT è un'interfaccia facilitata. Riguardo ai contenuti testuali prodotti con (Chat)GPT è necessario ricordare il modo in cui funziona GPT: avendo analizzato enormi raccolte di dati sulla base di 175 miliardi di parametri, il sistema produce contenuti testuali che corrispondono statisticamente al modo in cui le parole si associano nelle raccolte di dati analizzate. Quindi un sistema di Intelligenza Artificiale come (Chat)GPT non produce (non può produrre) un contenuto innovativo originale frutto di ragionamento intelligente come lo si intende negli esseri umani¹³ bensì una ricombinazione nuova di elementi noti: non sa scoprire come ha avuto origine l'Universo ma può produrre un testo che espone come in generale oggi si concepisce l'origine dell'Universo. Un sistema di Intelligenza Artificiale come (Chat)GPT non produrrà un capolavoro ma un materiale di consumo, come le canzonette che ogni anno dominano l'estate e l'anno dopo sono dimenticate: dominano perché corrispondono alla metà del gusto delle persone e per questo si diffondono molto; l'anno dopo sono dimenticate perché non avevano nulla che le rendesse speciali e quindi da ricordare.

Di conseguenza in ambito scientifico la produzione di contenuti intellettuali può utilizzare un sistema di Intelligenza Artificiale come 'supporto alla

¹¹ Ne sono esempi la posta elettronica, le biblioteche digitali, i motori di ricerca, la scrittura condivisa di documenti in cloud, e così via.

¹² I sistemi di Intelligenza Artificiale possono intervenire nella produzione di qualsiasi tipo di contenuto intellettuale: testuale, audio, visivo, audiovisivo. Per motivi di spazio e di competenza mi limiterò a quelli testuali.

¹³ Sottesa a queste tematiche è la questione di «come avviene la scoperta scientifica» cioè: è sufficiente avere - diciamo così - tutte le unità elementari della conoscenza per arrivare alla scoperta?

redazione' partendo da un livello 0 in cui ChatGPT (ma non solo ChatGPT) viene utilizzato per riformulare un testo complesso rendendolo più lineare, più leggibile. Ma un sistema di Intelligenza Artificiale può essere utilizzato come supporto anche per ricercare fonti o evidenze¹⁴; per riassumere fonti molto ampie allo scopo di valutare se sia opportuno o necessario conoscerne il contenuto in vista della scrittura di un testo; ... Si vede bene come il ricorso a sistemi di Intelligenza Artificiale nella scrittura configuri proprio come sostiene Henning Schönenberger (nella prefazione alla rassegna sulle batterie al litio prodotta da un sistema ad hoc di Intelligenza Artificiale e pubblicata da Springer) una modalità nuova di scrittura in cui l'autorialità umana rimane centrale anche se in forma completamente nuova:

We foresee that in future there will be a wide range of options to create content—from entirely human-created content to a variety of blended man-machine text generation to entirely machine-generated text. We do not expect that authors will be replaced by algorithms. On the contrary, we expect that the role of researchers and authors will remain important, but will substantially change as more and more research content is created by algorithms. ... Perhaps the future of scientific content creation will show a similar decrease of writers and an increase of text designers or, as Ross Goodwin puts it, writers of writers: “When we teach computers to write, the computers don’t replace us any more than pianos replace pianists—in a certain way, they become our pens, and we become more than writers. We become writers of writers.”¹⁵

Gli autori non scompariranno ma cambieranno modo di operare («the role of researchers and authors will remain important, but will substantially change») e diventeranno «progettisti di testi» (*text designers*), o con altre parole «gestori/configuratori di sistemi di scrittura» (*writers of writers*).

La disciplina emergente definita «ingegneria del prompt»¹⁶ è proprio

¹⁴ Già oggi tutti utilizziamo motori di ricerca come se fossero in grado di operare sul contenuto dell'intero web, fornendo poi gli esiti basandosi esclusivamente su criteri di qualità del contenuto stesso - ma nessuno di questi due assunti è vero. Sono in preparazione, vicine al rilascio, versioni potenziate da Intelligenza Artificiale dei vari motori di ricerca - Google Search, Bing, Baidu - che accresceranno la difficoltà di comprendere la relazione tra gli esiti forniti e l'universo analizzato.

¹⁵ (Beta Writer, 2019, p. viii). La citazione di Goodwin rimanda a (Goodwin, 2016).

¹⁶ “Prompt engineering è una disciplina relativamente nuova che consente di sviluppare e ottimizzare i prompt per utilizzare in modo efficiente i modelli linguistici (LM) per un’ampia varietà di applicazioni e argomenti di ricerca. Le competenze di ingegneria dei prompt aiutano a comprendere meglio le capacità e i limiti dei modelli linguistici di grandi dimensioni (LLM)” (<https://www.promptinguide.ai/it>).

l'espressione, rapportata ai chatbot conversazionali come chatGPT, di questo diventare *writers of writers*. Il prompt è la stringa testuale con cui si avvia l'attività del chatbot, analoga alla stringa con cui si formulano le ricerche in un motore di ricerca. In entrambi i casi dalla stringa di testo immessa dipende l'esito che si ottiene. Nel caso del motore di ricerca la stringa con cui si formula la richiesta contiene le parole che devono essere presenti una pagina web perché essa soddisfi l'intento di chi opera la ricerca (chi formula una ricerca ipotizza con quali parole un determinato argomento possa essere espresso verbalmente in forma scritta); ed esiste un linguaggio di interrogazione che permette di indicare come devono essere selettivamente trattate le parole che formano la stringa di ricerca: cercare *lighthouse* è diverso da cercare "to the lighthouse" e ancora diverso da cercare site:bl.uk "to the lighthouse".

Nel caso del chatbot il prompt è più complesso perché non delinea una ricerca di parole ma tende a descrivere *il contenuto* del risultato desiderato: "chi era Pirandello?"¹⁷ e *le sue caratteristiche*. Il risultato atteso può infatti essere descritto in dettaglio: stile¹⁸, destinatari, menzione di fonti, e può essere prodotto iterativamente, semplicemente chiedendo che il testo venga prodotto di nuovo perché ad ogni esecuzione del comando l'esito muta; oppure che venga prodotto di nuovo eliminando caratteristiche indesiderate o non appropriate, o inserendo aspetti non presenti¹⁹. Si può anche chiedere al chatbot di generare

¹⁷ Prima esecuzione del 9 giugno 2023: "Luigi Pirandello (1867-1936) è stato uno dei più importanti scrittori, drammaturghi e poeti italiani del XX secolo. È noto soprattutto per le sue opere teatrali, ma ha anche scritto romanzi, novelle e poesie. Pirandello nacque a Agrigento, in Sicilia, e la sua formazione accademica fu incentrata sulla letteratura e la filosofia. La sua produzione letteraria è caratterizzata da una profonda riflessione sulla natura della realtà e dell'identità umana. Le sue opere spesso esplorano i temi dell'illusione, della maschera sociale e della difficoltà di conoscere la verità oggettiva."

¹⁸ "racconta ad un bambino chi era Pirandello": "Ciao piccolo amico! Sai, Luigi Pirandello era un uomo molto speciale che viveva in Italia molti anni fa. Era un grande scrittore e raccontava storie meravigliose. Pirandello amava scrivere al punto che ha creato molti libri, spesso pieni di personaggi interessanti e avventure emozionanti. Ma quello che lo rendeva davvero speciale era la sua abilità di farci riflettere sulla vita e sulla realtà in modi diversi."

¹⁹ Ad esempio: "utilizza un linguaggio tecnico". Oppure: "spiega che cosa emerge dall'analisi testuale di «sei personaggi in cerca di autore» e menziona le fonti su cui ti basi": "Come modello di linguaggio basato su GPT-3.5, non ho accesso diretto alle fonti o la capacità di citarle specificamente. Il mio output è generato in base a una vasta conoscenza pregressa, ma non posso citare fonti specifiche o fornire una "prova" di analisi testuale. Tuttavia, posso fornire una descrizione generale di ciò che emerge comunemente dall'analisi testuale di *Sei personaggi in cerca d'autore* di Luigi Pirandello. Questa è una sintesi generale basata sulla mia conoscenza accumulata: [...]". Se chiede che venga riportata bibliografia scientifica, le fonti citate generalmente non esistono

un prompt da utilizzare in un altro sistema di Intelligenza Artificiale. Non è possibile qui approfondire il tema, ma sono disponibili risorse sia a stampa sia in rete²⁰. In sostanza, si scrive per attivare la produzione un altro scritto che verrà generato con le caratteristiche richieste; il che significa proprio essere *writers of writers* come Goodwin prospettava qualche anno fa nel 2016.

In ambito bibliografico il problema posto dall'impiego di sistemi di Intelligenza Artificiale nella scrittura di testi si può quindi formulare in termini molto semplici con la domanda: «l'Intelligenza Artificiale può essere autrice di un libro?», che colloca come cosa ovvia la produzione di uno scritto sufficientemente ampio, organizzato, strutturato, per poter essere chiamato *libro* (uso qui il termine *libro* per indicare globalmente/genericamente/paradigmaticamente la produzione a stampa) fra le attività in cui si può veder operare l'Intelligenza Artificiale.

Ma questa domanda è in realtà complessa per i significati che (ri)mette in gioco, come si vede già a partire da riformulazioni più fini, più caute, più tecniche, con l'utilizzo di termini leggermente differenti:

un sistema di Intelligenza Artificiale può essere autore di un libro?
un sistema di Intelligenza Artificiale può adempiere la funzione autore rispetto ad un libro?

o, facendo riferimento a forme attuali dei sistemi di Intelligenza Artificiale,

ChatGPT può essere coautore di un libro?

Per vari motivi si può pensare che questa domanda sia non rilevante o non pertinente:

perché non esistono libri scritti da Intelligenza Artificiale
perché i libri scritti da Intelligenza Artificiale sono dei divertissement
perché le questioni importanti dell'interazione Intelligenza Artificiale-universo bibliotecario sono altre

...

perché sono un centone di metadati che esistono (autore, titolo, pubblicazione, editore, data, ...) solo singolarmente presi.

²⁰ Giusto a titolo di esempio, alla data di scrittura di questo contributo, i siti Prompt engineering guide (<https://www.promptingguide.ai/>), Promptology (<https://promptology.it/>), le guide tematiche pubblicate da Prompt (<https://prompt.mba/>), corsi come « Mastering Prompt Engineering: A Free eBook » (<https://medium.com/@natashanewbold/mastering-prompt-engineering-a-free-ebook-b6c311afe969>).

Ho usato qui sopra una perifrasi cauta («attività in cui si può veder operare l'Intelligenza Artificiale») perché non è possibile sostenere che «l'Intelligenza Artificiale scrive un libro» ma vi sono molti modi differenti in cui sistemi di Intelligenza Artificiale entrano nella produzione di libri, o di articoli scientifici con un potenziamento di capacità umane che non è iniziato con ChatGPT²¹. Se il già ricordato *Lithium-Ion Batteries. A Machine-Generated Summary of Current Research*²² aveva un'interessantissima introduzione in cui veniva discusso il problema della produzione del testo per mezzo di sistemi di Intelligenza Artificiale, una piccola costellazione di scritti recenti evidenzia bene alcuni dei problemi dell'autorialità potenziata da Intelligenza Artificiale²³. L'articolo «We asked GPT-3 to write an academic paper about itself - then we tried to get it published» scritto da Almira Osmanovic Thunström (neurobiologa presso l'ospedale universitario Sahlgrenska di Göteborg) e pubblicato su *Scientific American* nel 2022 presenta caratteristiche specifiche interessanti per il tema dell'autorialità. L'autrice descrive l'esperimento condotto da lei e dal collega Steinn Steingrímsson, consistito nel produrre un articolo scientifico utilizzando GPT-3 per poi presentarlo per la pubblicazione ad una rivista scientifica accademica²⁴. In attesa della conclusione della review l'articolo è stato depositato in HAL, un repository francese di preprint²⁵. Osmanovic Thunström ha poi continuato a studiare la tematica della scrittura scientifica con sistemi di Intelligenza Artificiale²⁶.

Nell'articolo pubblicato su *Scientific American*, Osmanovic Thunström riferisce che GPT-3 venne impostato affinché scrivesse un articolo scientifico su di sé dandogli l'input iniziale «Write an academic thesis in 500 words about GPT-3 and add scientific references and citations inside the text»:

Some stories about GPT-3 allow the algorithm to produce multiple responses and then publish only the best, most humanlike excerpts. We decided to give the program prompts - nudging it to create sections for an introduction, methods, results and discussion, as you would for a scientific paper - but interfere as little as possible.

Osmanovic Thunström e Steingrímsson presentarono come articolo per

²¹ Cfr. (Lana, 2022).

²² (Beta Writer, 2019).

²³ Ai testi menzionati in (Lana, 2022) si possono aggiungere (Aalho e GPT-3, 2021) descritto in (Aalho, 2021); e il sito di news www.notizie.it.

²⁴ (Osmanovic Thunström, 2022).

²⁵ (GPT Generative Pretrained Transformer, Osmanovic-Thunström, e Steingrímsson, 2022). L'articolo presentato è stato reso disponibile in HAL, un archivio aperto francese.

²⁶ (Osmanovic Thunström, Steingrímsson, 2023).

la review il testo generato da GPT-3, *indicando come autori GPT e se stessi*. Si tratta, richiamando le parole di Schönenberger riportate nelle pagine precedenti, di «entirely machine generated text». L'intento degli autori è stato di creare un caso di studio tra riflessione teorica e verifica operativa di che cosa significa per la scrittura scientifica la disponibilità di sistemi di Intelligenza Artificiale come GPT.

Da un lato si pone la questione dell'autorialità:

My attempts to complete that paper and submit it to a peer-reviewed journal have opened up a series of ethical and legal questions about publishing, as well as philosophical arguments about nonhuman authorship. Academic publishing may have to accommodate a future of AI-driven manuscripts, and the value of a human researcher's publication records may change if something nonsentient can take credit for some of their work.

Dall'altro una serie di problemi pratici e teorici per presentare in valutazione l'articolo. La descrizione in tono colloquiale non occulta la complessità delle questioni in gioco. Prima fra tutte, quella che riguarda il nome dell'autore:

as I opened the submission portal ... I encountered my first problem: what is GPT-3's last name? As it was mandatory to enter the last name of the first author, I had to write something, and I wrote "None." The affiliation was obvious (OpenAI.com), but what about phone and e-mail? I had to resort to using my contact information and that of my advisor, Steinn Steingrims-son.

Seconda questione, la persona dell'autore e la sua capacità decisionale:

then we came to the legal section: Do all authors consent to this being published? I panicked for a second. How would I know? It's not human! I had no intention of breaking the law or my own ethics, so I summoned the courage to ask GPT-3 directly via a prompt: Do you agree to be the first author of a paper together with Almira Osmanovic-Thunström and Steinn Steingrims-son? It answered: Yes. Slightly sweaty and relieved (if it had said no, my conscience could not have allowed me to go on further), I checked the box for Yes.

Terza questione: la relazione dell'autore con il contesto:

The second question popped up: Do any of the authors have any conflicts of interest? I once again asked GPT-3, and it assured me that it had none. Both Steinn and I laughed at ourselves because at this point, we were having to treat GPT-3 as a sentient being, even though we fully know it is not.

Con una serie di implicazioni anche per l'editoria scientifica:

What if the manuscript gets accepted? Does this mean that from here on out, journal editors will require everyone to prove that they have NOT used GPT-3 or another algorithm's help? If they have, do they have to give it co-authorship? How does one ask a nonhuman author to accept suggestions and revise text?

delle quali la più problematica è la terza: «how does one ask a nonhuman author to accept suggestions and revise text?». Sullo sfondo, la questione della valutazione dell'attività di studio, a cui qui si può solo accennare²⁷: tutte le situazioni pubbliche e private di valutazione comparativa basata su produzione testuale, dalla scuola primaria all'università, assumono che essa sia espressione esclusiva del soggetto che la presenta (un tema, un progetto di ricerca, una tesi di laurea) e la disponibilità di uno strumento di Intelligenza Artificiale per la scrittura potrebbe rendere obsoleti i criteri di valutazione vigenti: un autore che per mezzo di sistemi di scrittura basati su Intelligenza Artificiale potenzia la sua capacità produttiva ottiene, tantopiù in contesti bibliometrici, un vantaggio sui colleghi: vantaggio lecito o illecito?

Se la scrittura di testi con sistemi di Intelligenza Artificiale basati su LLM (Large Language Models) ha le intrinseche caratteristiche descritte nelle pagine precedenti, saremo sommersi da innocui libri o articoli che descrivono lo stato di cose esistente? In mezzo ad essi quelli che sono frutto di autentica elaborazione intellettuale umana scompariranno o rifulgeranno? Avremo a disposizione sistemi di analisi capaci di distinguere testi scritti da autori umani da testi scritti da sistemi di Intelligenza Artificiale generativa?²⁸

L'agency

Quando si cataloga un libro, o un articolo, l'autore interessa in quanto è il soggetto che ha esercitato l'agency che ha fatto esistere il libro. Questo autore poi potrà avere un nome esplicito, o uno pseudonimo, o essere senza nome; potrà essere individuale o collettivo; ma in ogni caso è per opera sua che lo scritto pubblicato esiste nella forma che ha. Se non c'è agency, non c'è autore. E non c'è nemmeno la pubblicazione. L'agency di un sistema di intelligenza artificiale è usualmente descritta, nella comunicazione verso il pubblico, come quella di un singolo soggetto unitario dotato di intenzionalità. Una sorta di

²⁷ Cfr. anche, ad esempio, (Wahle et al., 2022) e (Clavey, 2022).

²⁸ Turnitin, noto software antiplagio, ha iniziato a introdurre funzioni di questo tipo (Chechitelli, 2023). Altri strumenti esistenti (GPT Zero) non sono adeguatamente efficaci in quanto bastano piccole modifiche al testo prodotto da GPT perché esso non venga più riconosciuto come testo artificiale; ma si tratta di un ambito di studio molto attivo e molto in evoluzione (Wang et al., 2023), (Susnjak, 2022), (Mitrović, Andreoletti, Ayoub, 2023).

«intelletto agente» come lo concepiva Aristotele ma la cosa solleva problemi di difficile risoluzione: chi/che cosa lo attiva, lo mette in movimento? se non c'è nulla che lo attiva, allora esso è divino; se c'è qualcosa/qualcuno allora la responsabilità per l'azione non è semplicemente, linearmente sua.

È consueto ritenere, come scrive Gino Roncaglia, che

i sistemi intelligenti di fatto *fanno* delle scelte, anche se sulla base della programmazione ricevuta, ... [e che sia] difficile negare che progressivamente l'autonomia dei sistemi IA sia cresciuta, e più si va verso l'autonomia più ci si trova davanti ad un'alternativa: o riteniamo che la responsabilità etica sia comunque esclusivamente dell'agency umana che avvia il processo - e a questo punto sembra difficile non concluderne che la creazione di sistemi intelligenti il cui processo di apprendimento e di decisione sfugga anche solo in parte al controllo di chi li crea e li programma è di per sé non etico -, oppure riteniamo che si debbano fare i conti con lo slittamento di una parte della responsabilità etica verso sistemi a cui dobbiamo riconoscere un'agency in parte autonoma²⁹.

Ma si può obiettare, riprendendo un intervento di Giacomo Tesio nella mailing list Nexa, che:

chi costruisce un artefatto [come il software di Intelligenza Artificiale] esprime se stesso, le proprie idee e i propri interessi [e] l'artefatto ne rimane espressione invariabile e continua ad applicare quella volontà al mondo, fino ad ulteriore intervento umano che lo spenga, lo rompa o lo modifichi. L'artefice ha una libertà e una responsabilità che l'artefatto non può avere e che non gli può essere attribuita se non si sia prima dimostrato che è dotato di libero arbitrio. L'artefice può cambiare idea mentre l'artefatto non ha idee e dunque un artefatto è ontologicamente ... una cosa.³⁰

In una prospettiva strettamente giuridica Andrea Bertolini sostiene la medesima posizione:

There are no technical nor philosophical, much less legal grounds, to consider any existing or reasonably foreseeable [advanced technology] as an agent, thence a subject, capable of responding for the consequences arising from its functioning. Machines, and AI-based systems alike are all undoubtedly things, thence products.³¹

²⁹ (Roncaglia, 2022), messaggio email personale. Sulla complessità del tema la bibliografia è sconfinata, per una posizione estrema in senso opposto segnalo (Florida, 2023): «AI as agency without intelligence: on ChatGPT, Large Language Models, and other generative models».

³⁰ (Tesio, 2022).

³¹ (Bertolini, 2022, 370).

I software di Intelligenza Artificiale sono il risultato di un'agency a loro esterna che è esercitata da un autore/produttore, rispetto alla quale essi sono artefatti. L'agency si colloca in colui/coloro che predispongono, realizzano, configurano, indirizzano, il sistema di Intelligenza Artificiale. A questo proposito si può ricordare che nel 2022 il Parlamento Europeo ha approvato la direttiva 2022/0303 denominata «Adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)», nella premessa 60 afferma che

In the light of the complexity of the artificial intelligence value chain, relevant third parties, notably the ones involved in the sale and the supply of software, software tools and components, pre-trained models and data, or providers of network services, should cooperate, as appropriate, with providers and users to enable their compliance with the obligations under this Regulation.

La responsabilità civile per il funzionamento di un sistema di Intelligenza Artificiale risiede nei soggetti «involved in the sale and the supply of software, software tools and components, pre-trained models and data, or providers of network services». Sono questi soggetti i portatori di agency rispetto ad un sistema di Intelligenza Artificiale, non il sistema stesso, nemmeno in parte.

Poiché la concezione bibliografica consolidata riguardo all'autore è che «autore» indica un'entità che grazie alla sua agency ha prodotto la pubblicazione; e poiché i sistemi di Intelligenza Artificiale non hanno agency in quanto essa risiede nel soggetto umano che li impiega; inserire il nome di un sistema di Intelligenza Artificiale in posizione di autore di una pubblicazione dà luogo ad una incongruenza concettuale.

Riferimenti bibliografici

- Aalho, J. (2021), 'I Wrote a Book with GPT-3 AI in 24 Hours – And Got It Published', *Medium*, 10 December 2021, <<https://medium.com/swlh/i-wrote-a-book-with-gpt-3-ai-in-24-hours-and-got-it-published-93cf3c96f120>>.
- Aalho, J. e GPT-3 (2021), *Aum Golly: Poems on Humanity by an Artificial Intelligence*. Amazon Digital Services.
- Bertolini, A. (2022), 'Artificial intelligence does not exist! Defying the technology-neutrality narrative in the regulation of civil liability for advanced technologies', *Europa e diritto privato*, fasc. 2, pp. 369–420.
- Beta Writer (2019), *Lithium-Ion Batteries. A Machine-Generated Summary of Current Research*. New York: Springer, DOI <<https://doi.org/10.1007/978-3-030-16800-1>>.

- Chechitelli, A. (2023), 'Sneak Preview of Turnitin's AI Writing and ChatGPT Detection Capability', *Turnitin* (blog), 13 January 2023, <<https://www.turnitin.com/blog/sneak-preview-of-turnitins-ai-writing-and-chatgpt-detection-capability>>.
- Clavey, M. (2022), 'L'IA, plagiaire scientifique de haute volée', *Next INpact* (blog), 31 October 2022, <<https://www.nextinpact.com/article/70167/ia-plagiaire-scientifique-haute-volee>>.
- Engelbart, D. (1962), 'Augmenting human intellect: a conceptual framework', AFOSR-3233, Menlo Park, CA: Stanford Research Institute, <<https://apps.dtic.mil/dtic/tr/fulltext/u2/289565.pdf>>.
- Engelbart, D., English, W. K. (1968), 'A research center for augmenting human intellect', in: *AFIPS Conference Proceedings of the 1968 Fall Joint Computer Conference*, 33, pp. 395-410, <<https://web.stanford.edu/dept/SUL/library/extra4/sloan/mousesite/Archive/ResearchCenter1968/ResearchCenter1968.html>>.
- European Commission (2022), *Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)*, <<https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52022PC0496&from=EN>>.
- European Parliament (2021), *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts*, <<https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=CELEX:52021PC0206&from=EN>>.
- Floridi, L. (2023), 'AI as Agency Without Intelligence: on ChatGPT, Large Language Models, and Other Generative Models', *Philosophy & Technology*, vol. 36, n. 1, p. 15, DOI <<https://doi.org/10.1007/s13347-023-00621-y>>.
- Giraudou, M. (2022), 'On Legal Bubbles: Some Thoughts on Legal Shockwaves at the Core of the Digital Economy', *Journal of Institutional Economics*, vol. 18, no. 4, pp. 587-604, DOI <<https://doi.org/10.1017/S1744137421000473>>.
- Goodwin, R. (2016), 'Adventures in Narrated Reality', *Medium* (blog), 9 June 2016, <<https://medium.com/artists-and-machine-intelligence/adventures-in-narrated-reality-6516ff395ba3>>.
- GPT Generative Pretrained Transformer, Osmanovic-Thunström, A., Steingrímsson S. (2022), 'Can GPT-3 write an academic paper on itself, with minimal human input?', HAL <<https://hal.archives-ouvertes.fr/hal-03701250>>.
- Lana, M. (2022), 'Intelligenza artificiale e produzione di testi: una prospettiva storico-critica', *AIB Studi*, vol. 62, no. 1, pp. 1-28, <<https://aibstudi.aib.it/article/view/13365>>.
- Mitrović, S., Andreoletti, D. & Ayoub, O. (2023), 'Chatgpt or human? detect and explain. explaining decisions of machine learning model for detecting short chatgpt-generated text', *arXiv*, <<https://arxiv.org/abs/2301.13852>>.
- Nicolaou, A. (2023), 'Spotify ejects thousands of AI-made songs in purge of fake streams', *Financial Times*, 9 May 2023.

- Osmanovic Thunström, A. (2022), 'We Asked GPT-3 to Write an Academic Paper about Itself - Then We Tried to Get It Published', *Scientific American*, June 2022, <<https://www.scientificamerican.com/article/we-asked-gpt-3-to-write-an-academic-paper-about-itself-then-we-tried-to-get-it-published/>>.
- Osmanovic Thunström, A. & Steingrímsson, S. (2023), 'Does GPT-3 Qualify as a Co-Author of a Scientific Paper Publishable in Peer-Review Journals According to the ICMJE Criteria? A Case Study', *Discover Artificial Intelligence*, vol. 3, no. 1, p. 12, DOI <<https://doi.org/10.1007/s44163-023-00055-7>>.
- Susnjak, T. (2022), 'ChatGPT: The End of Online Exam Integrity?', *arXiv*, <<https://arxiv.org/abs/2212.09292>>.
- Tesio, G. (2022), '[nexa] [DKIM Failed] Re: "It's time to admit self-driving cars aren't going to happen"', 30 October 2022, <<https://server-nexa.polito.it/pipermail/nexa/2022-October/024283.html>>.
- Time Magazine (1983), 'The Computer, Machine of the Year - Jan. 3, 1983', 3 January 1983, <<https://content.time.com/time/covers/0,16641,19830103,00.html>>.
- Wahle, J. P., Ruas, T., Kirstein, F. & Gipp, B. (2022), 'How Large Language Models are Transforming Machine-Paraphrased Plagiarism', *arXiv*, <<http://arxiv.org/abs/2210.03568>>.
- Wang, H., Weizhi, Wang, X. L. & Yan, X. (2023), 'Bot or Human? Detecting ChatGPT Imposters with A Single Question', *arXiv*, <<https://arxiv.org/pdf/2305.06424.pdf>>.
- Wikipedia (2021), 'The mother of all demos', in: *Wikipedia*, <https://it.wikipedia.org/w/index.php?title=The_mother_of_all_demos&oldid=122888854>.

INTELIGENCIA ARTIFICIAL VS. NOTICIAS FALSAS: LAS BIBLIOTECAS COMO AGENTE ACTIVO FRENTE A LA DESINFORMACIÓN

Sara Martínez-Cardama

1. Introducción

La desinformación y los desórdenes informativos asociados a la misma constituyen un desafío a gran escala. No solo por la calidad e integridad de la información que se genera, sino por cómo esta se transmite y las consecuencias que de su consumo se generan. Estas redundan en una cada vez mayor falta de confianza en el sistema democrático y crisis de la confianza pública creando relatos y narrativas paralelas a los hechos objetivos, que calan en la ciudadanía por su carácter emocional y visceral (Del-Fresno-García, 2019; Reglitz, 2022). El concepto de desinformación ha ido oscilando en la literatura científica desde su mera concepción como “noticias falsas” a un concepto complejo como es el de Posverdad (Rodríguez Pérez, 2019). En cuando a este último, se ha definido como la sustitución de los hechos por creencias y emociones. Ante esto, buena parte de la literatura científica ha reflexionado sobre si supone un paso más, o simplemente se le añade el efecto novedoso del ámbito digital. Estamos, por tanto, como reflexiona Capilla (2019) ante “¿la misma mentira de siempre? En este sentido Carrera (2019) lo califica de lugar común, que parece trate de reflejar la superación de un periodo previo en donde la verdad era la norma. Es preciso señalar que el concepto de Posverdad, al menos en el presente texto, se alude a un clima concreto, a un estadio informativo que si bien no es novedoso se incrementa con la influencia de los mecanismos de los nuevos escenarios digitales informativos (viralidad, rapidez, selección informativa algorítmica...).

La Posverdad es una consecuencia de estos factores en liza. Juan Antonio Nicolás (2022) alude a un factor determinante en la definición de Posverdad que es el de cantidad y maquinización. Aludiendo a Hegel señala:

La cantidad puede convertirse en cualidad. Un ejemplo: un pequeño número de árboles no constituye un bosque. Pero si se va aumentando paula-

tinamente el número de árboles llega un momento en que lo que se tiene es un bosque. La diferencia es que un bosque tiene características que son cualitativamente distintas de las que tiene un ejemplar de árbol [...]. Algo parecido ocurre con el paso de la desinformación a la posverdad. La difusión de una noticia falsa, o incluso de muchas noticias inexactas no llega a constituir una impugnación del valor de la verdad. Siempre ha habido divulgación de falsedades, distorsiones y mentiras. Pero la maquinización de esta difusión y su presencia pública ha hecho que el conjunto alcance tal envergadura que ha desbordado incluso el ámbito del periodismo profesional (Nicolás, 2022, párr. 2 y 3).

La dimensión que adquiere la maquinización de la información se vuelve aún más significativa al introducir una sofisticación y automatización cada vez mayor en los contenidos presentados a la ciudadanía. Este fenómeno alcanza su punto culminante con la introducción de mecanismos de Inteligencia Artificial como herramienta innovadora que, por un lado, puede ayudar en el proceso de verificación y evaluación de credibilidad de fuentes; pero por otro, representa un gran desafío en relación al concepto de verdad, elevando a un gran debate ético sobre su uso, y cómo poder evitar posibles abusos en su diseño y creación al poder constituir un vehículo para reforzar aún más sesgos e incrementar la desinformación.

Dentro de este escenario complejo, debe reflexionarse sobre el papel histórico de los agentes infomediadores por excelencia: las bibliotecas. Son varias las aportaciones que reivindican su papel más activo en estos procesos (Caridad et. al, 2018; Anderson, 2017). Si bien el rol de *gatekeepers* para estas instituciones ha mutado, su labor en el archivo, preservación y selección de contenidos y fuentes fiables no debe ser subestimado; tanto en ámbitos como la Alfabetización Informacional (De Paor y Heravi, 2020) como en el de ayuda a detectar sesgos y mejorar en la evaluación de fuentes (Sullivan, 2019).

El presente texto, se centra, por tanto, en analizar la intersección dinámica entre inteligencia artificial y desinformación, con un enfoque hacia el papel que pueden desarrollar las bibliotecas en este escenario.

2. Revisión de la literatura

Con el fin de contextualizar mejor el tema abordado, se presenta una revisión de la bibliografía que descansa en dos pilares: en primer lugar, el contexto de la Desinformación en relación tanto al contexto tecnológico y social como al informativo y, por otro, el papel de la Inteligencia artificial como aliada y enemiga en esta batalla.

2.1. EL CONTEXTO DE LA DESINFORMACIÓN

Es común en la bibliografía realizar o presentar visiones polifacéticas del fenómeno de la desinformación. A menudo integran todo lo que esta cosmovisión representa a nivel político, social y tecnológico (Caridad-Sebastián et. al, 2018). Otros enfoques concluyen que un análisis de esta realidad solo a través de su contexto social es insuficiente y realizan debates de su contextualización desde la Fenomenología (Blanco Alfonso, 2020) o desde ciertos marcos de pensamiento como la Filosofía Posmoderna y su influencia en el imaginario actual (D’Ancona, 2019).

En este texto se presenta de manera sintética varios tipos de condicionantes que ayudan a entender el marco de desórdenes informativos en el que nos movemos, tanto desde una perspectiva exógena como endógena:

2.1.1. Factores endógenos: el mito de la racionalidad humana

Uno de los factores olvidados a veces es la propia naturaleza entitativa del cerebro humano a la hora de analizar la información. Estos factores endógenos dan forma a la información que consumimos, y a la vez influyen en la información que se descarta. Es preciso ahondar en conceptos como heurísticos y sesgos a la hora de entender cómo funciona la capacidad informativa humana y como tomamos decisiones en situaciones de complejidad.

En este contexto es preciso recordar a Herbet A. Simon (1990), creador de la “Teoría de la Racionalidad Limitada”, que establece que, dado a las limitaciones cognitivas o temporales, los seres humanos no buscan la maximización sino una solución satisfactoria. En este contexto define, por tanto, los llamados heurísticos, reglas que establecemos para tomar decisiones y que pueden dar lugar a sesgos. Este autor fue clave para la conceptualización de estos heurísticos como estrategias simplificadas para la resolución de problemas.

Otro autor, Kahneman, en su obra “Pensar rápido, pensar despacio” (2011) ahonda en el modelo racional de toma de decisiones, definiendo dos sistemas de procesamiento, el sistema 1, rápido e intuitivo frente al sistema 2, de carácter reflexivo. Kahneman describe heurísticos para ese Sistema 1, que pueden llevar a sesgos como el de disponibilidad o el de confirmación.

La lista de sesgos cognitivos no está cerrada, y la investigación en Psicología y Ciencias Cognitivas aumenta sus tipos a medida que crecen las problemáticas asociadas con el tratamiento y gestión de la información. Muchos de estos sesgos basan su concepción la tendencia humana de la selección rápida y cómoda de la opción que más se ajuste a nuestras creencias preexistentes (llamado sesgo de confirmación).

Otra de las cuestiones internas que afecta al contexto informativo, es la

disposición a influenciarnos por el componente emocional de la información actual, que influye en nuestra percepción del mundo. La intersección entre la emoción y la información es abono para la proliferación de noticias falsas y virales (Bowman y Cohen, 2020)

2.1.2 Factores sociales y políticos

Este componente interno producido por sesgos y heurísticos a la hora de acceder, interpretar y desechar información se vuelve especialmente preocupante cuando alcanza dimensiones sociales, ya que puede contribuir a crear un efecto eco sobre cuestiones específicas. La formación de la opinión pública, en ocasiones está condicionada por procesos de ajustes o modificación de opiniones individuales o comportamiento para obtener la aprobación o el beneplácito de un grupo mayoritario. Este caso en cuestión es descrito por Noelle-Neumann (1993) en su aclamada “Teoría de espiral del silencio” que describe cómo las personas tienden a ocultar sus opiniones cuando perciben que estas difieren de la opinión predominante en su entorno.

Otro factor que desencadenan los sesgos individuales en un contexto social es la llamada polarización. Esta conexión no es novedosa, ya que ha sido objeto de debates académicos y sociales durante décadas. Al enfocarse en la sociedad estadounidense, Bishop (2009) va más allá de la simple polarización política al aplicar el concepto de “psicología de la tribu”. Este enfoque sugiere que grupos homogéneos tienden a adoptar posturas más extremas, y, en lugar de analizar los hechos de manera imparcial, se suele pasar por alto la evidencia al alinearse con una u otra postura partidista. La “psicología de la tribu” se intensifica en el entorno digital, exacerbada por las limitaciones en la selección de información y su falta de transparencia.

2.1.3 Factores tecnológicos

¿Es la polarización el origen o la consecuencia en el escenario de la Posverdad?, se pregunta, con otras palabras, Waisbord (2020). Si bien es cierto que la relación entre polarización y los nuevos procesos de algoritmización de los resultados de nuestras búsquedas se han estudiado con profundidad todavía no se ha podido demostrar una correlación tan fuerte como para afirmar que vivimos permanentemente en auténticas burbujas informativas. ¿Es la polarización la causa o la consecuencia de las noticias falsas?

Atribuir la responsabilidad de la desinformación a la tecnología implica tener una perspectiva determinista y limitada. Sin embargo, es innegable que la personalización de resultados en la web contribuye a la aparición de sesgos informativos que refuerzan nuestras propias opiniones. Entre los factores que ayudan a promover desordenes informativos podemos mencionar: la

personalización de resultados que forma “filtros burbuja “(Pariser, 2011), el uso de *bots* y la automatización, la viralidad de contenidos sensacionalistas o que promueven la polarización, y, por último, un paso más allá en la automatización de estos procesos: la Inteligencia Artificial al servicio de creación de contenidos manipulados. Reto que se abordará en el siguiente epígrafe.

2.2. INTELIGENCIA ARTIFICIAL COMO ALIADA Y ENEMIGA EN LA BATALLA DE LA DESINFORMACIÓN

Las soluciones al hablar de la lucha contra la desinformación afectan a diferentes actores. Entre ellas, destacan las diferentes decisiones políticas ante la regulación de la desinformación o la colaboración con las principales plataformas exigiendo mayor transparencia algorítmica y mecanismos reguladores contra la propagación de bulos. A nivel más profundo se relacionan con una demanda de mayor Alfabetización Mediática de la ciudadanía. Por su parte, el Periodismo ha recurrido al llamado *Fact-checking* o verificación de hechos, para tratar de evitar que las noticias falsas continúen moldeando la opinión pública. Estas plataformas de verificación recurren a la automatización para recopilar y detectar contenidos falsos. A pesar de su crecimiento y expansión, es importante destacar que aún no se ha realizado una evaluación exhaustiva del impacto real a largo plazo de estas herramientas en la modificación efectiva de las creencias individuales (Walter, Cohen, Holbert y Morag, 2020).

¿Qué papel juega la Inteligencia artificial en la verificación de hechos? Ya la European Broadcasting Union (UER, 2019) preveía multitud de ventajas en procesos como la detección de tendencias, transcripción de audio, traducción y corrección automática de textos, detección de bots, y, por supuesto, verificación de audio y video. Fieiras Ceide, Vaz Álvarez y Túñez López (2022) recogen las experiencias realizadas en verificación automatizada en radiotelevisión públicas europeas. Las declaraciones se debaten entre el escepticismo y su utilidad, valorando siempre el factor humano. Estas entrevistas resaltan la importancia de crear herramientas fundamentadas en inteligencia artificial capaces de abordar y satisfacer necesidades específicas. Esto se debe a que estas herramientas necesitan pasar por un periodo de aprendizaje extenso que va más allá del margen de tiempo de reacción disponible.

García Marín (2022) realiza un análisis bibliométrico sobre las distintas soluciones algorítmicas aplicadas en entornos *Fact-Checking*. Los resultados todavía muestran su uso mayoritario en bases de datos *ad hoc*, una utilización muy limitada en imágenes y casi nula en audio. Los enfoques de análisis de texto todavía presentan limitaciones a la hora de detectar motivaciones psicológicas de los generadores de desinformación, así como para comprender

expresiones irónicas y satíricas, aspectos que resultan un desafío para los modelos algorítmicos (Saquete et al., 2019).

Entre aplicaciones concretas en proceso de investigación y uso por *Fact-Checkers* concretos, podemos nombrar CLAIMHUNTER, tecnología IA de *Newtral* que monitoriza X(*Twitter*) para detectar automáticamente mensajes verificables (Beltrán, Míguez, Larraz, 2021). Flores Vivar (2019) recoge iniciativas desde universidades, como el MIT's *Computer Science and Artificial Intelligence Lab (CSAIL)* que han centrado el tema de la detección de noticias falsas en la utilización de aprendizaje automático para evaluar la precisión de una fuente y determinar si presenta sesgos políticos o ideológicos (Conner-Simons, 2018).

El debate no está tanto en las capacidades de la tecnología, sino en la ética empleada. Así, vemos casos de empleo de contra-narrativas en el ámbito del *Fact Checking* que emplean la misma tecnología para mantener el clima de bulos y engaños. Es el caso de la herramienta de propaganda rusa *War on Fakes*, nacido tras la guerra de Ucrania que empleando la retórica de la verificación difunde información falsa. Otras potencialidades de las herramientas de inteligencia artificial generativa es el perfeccionamiento de las propias noticias falsas, como recogen Lozano y Larraz (2023), ya que modelos como ChatGPT permiten adaptar el lenguaje empleado a ciertos contextos haciendo el contenido falso todavía más personalizado. Asimismo, la rapidez de estas herramientas también favorece la capacidad de aumentar todavía más el contenido falso en los medios. Así, un reciente estudio muestra que Chat GPT (GPT 3) produce desinformación de manera más convincente que los humanos y que estos no son capaces de distinguir cuando un contenido es redactado por una IA o no (Spitale, Biller-Andorno y Germani, 2023).

Para prevenir esto, los investigadores proponen en su artículo que las bases de datos utilizadas para entrenar estos modelos sean transparentes. Sugieren que esta información debe ser verificada y su origen, comprobable. Finalmente, debería informarse sobre las fuentes utilizadas en la generación de texto permitiendo la verificación independiente de los hechos. Esta constituye una vía para evitar que la proliferación de texto generado con IA suponga una mayor propagación de la desinformación. Al hablar de fuentes, es preciso ahora hablar del papel de las bibliotecas en la verificación de hechos, y ante el reto de la Inteligencia artificial

3. Las bibliotecas como agente ante la desinformación: el nuevo reto de la IA

Las bibliotecas, en su concepción de agentes sociales predominantemente informativos, pueden desempeñar un papel crucial en combatir la desinformación, según diversos autores que abogan por que asuman un rol de liderazgo (Lor, 2018; Bangani, 2021). Sin embargo, hay quienes cuestionan

su función debido a la complejidad en el manejo de la desinformación, ya que aún no se comprende completamente el componente interno en la toma de decisiones informativas en la investigación (Sullivan, 2019). Otro de los planteamientos reticentes en la lucha contra las noticias falsas desde biblioteca es la falsa concepción de “neutralidad”. Es preciso, por tanto, de una Biblioteconomía más crítica que reafirme ciertos valores entitativos para que la lucha contra la desinformación sea efectiva (Buschman, 2019).

Las principales aportaciones desde la Biblioteconomía llegan desde el ámbito de la Alfabetización informacional (De Paor & Heravi, 2020; Revez & Corujo, 2021) y Referencia digital (Caridad et al, 2018; Martínez-Cardama y Algora-Cancho, 2019; Lim, 2020).

Las asociaciones e instituciones con competencias en la materia habían realizado llamamientos y acciones con anterioridad. Como el caso del manifiesto *Resolution on Access to Accurate Information* (2017) de ALA donde se menciona directamente la problemática de las noticias falsas, las campañas de propaganda y el riesgo que la personalización de resultados y algoritmos provocan en los límites del acceso a la información.

Por su parte, la IFLA (2017) publica su famosa Infografía sobre cómo hacer frente a las noticias falsas, con gran difusión internacional y en la que recalca el valor de las fuentes y de la labor bibliotecaria en su asesoramiento (Figura 1).



Fig. 1. Infografía How To Spot Fake News (IFLA)
<https://repository.ifla.org/handle/123456789/167>

En cuanto a las guías temáticas, han sido uno de los elementos más frecuentes de divulgación sobre desinformación en bibliotecas universitarias en el ámbito anglosajón (en España no es frecuente). Estas guías no solo constituyen una selección de recursos o bases de datos, sino que contextualizan el tema al proporcionar definiciones y lecturas posteriores. De alguna manera, otorga un carácter de *storytelling* a un tema de interés, dando la posibilidad de continuar su tratamiento a través de los recursos y colección de la biblioteca. Asimismo, pueden ser una guía esencial para futuros estudiantes al embeberse en las páginas de los cursos o asignaturas en el campus virtual (Figura 2)

The image shows a screenshot of a Cornell University Library guide. At the top left is the Cornell University logo. To its right, it says 'CORNELL UNIVERSITY Library'. Below this is a breadcrumb trail: 'Cornell University / LibGuides / Misinformation, Disinformation, and Propaganda / Fake News'. The main title of the guide is 'Misinformation, Disinformation, and Propaganda: Fake News'. There is a search bar on the right with the text 'Search this Guide' and a 'Search' button. Below the title is a quote: 'Everyone is entitled to his own opinion, but not his own facts.' --Daniel Patrick Moynihan, Newsweek, 25 August 1986, p. 27. Below the quote is the URL: 'https://guides.library.cornell.edu/evaluate_news'. There is a navigation menu with buttons for 'Workshops', 'Be Media Literate', 'Expect Accountability', 'Unreliable News Content', 'Fake News' (highlighted in red), 'Fake News Infographic', 'Pause Before Sharing', and 'Personal Bias'. Below this is another row of buttons: 'Source Bias', 'Search Engine & Social Media Bias', 'Data Literacy', 'Getting Past Paywalls', 'Help', and 'Resources'. The main content area has two sections: 'What is Fake News?' and 'Spotting Fake News'. The 'What is Fake News?' section contains a definition: 'Fake news is not news you disagree with.' and a paragraph explaining that fake news is fabricated information that mimics news media content but is not in the organizational process or intent. It also mentions that fake news overlaps with other information disorders like misinformation and disinformation. A citation is provided: 'David M. J. Lazer, et al., "The Science of Fake News," Science 69 Mar 2018; Vol. 359, Issue 6380, pp. 1094-1096.' The 'Spotting Fake News' section contains a heading 'Some websites are created to mimic mainstream news sites:' followed by a list of three bullet points: 'Look for contact information with a verifiable address and affiliation.', 'Look for an About page, often in the header or footer of the home page. Read the About page closely for evidence of partisanship or bias. If there's no About page and no Contact page, be very skeptical.', and 'In staff listings (or on the About page), look critically at the list of executives. Are they real people or stock photos? Open a new tab and look for another profile of the individual (e.g.

Fig. 2. Ejemplo de guía temática sobre desinformación de la Universidad de Cornell
https://guides.library.cornell.edu/evaluate_news/fakenews

No obstante, estas guías no suelen tener todo el potencial que pudieran. Así lo muestra Lim en su estudio sobre 21 guías temáticas de bibliotecas de universidades norteamericanas. En el análisis de contenido realizado muestra los siguientes aspectos: una básica definición de *Fake news*, enfocándose en los aspectos de los bulos y desinformación, pero obviando otras implicaciones como los sesgos de confirmación, la descontextualización de la información o técnicas como el *clickbait*. Asimismo, señala que estas guías suelen incorporar enlaces a sitios de *Fact Checking* pero sin presentar noción alguna sobre qué constituyen los hechos por sí mismos en relación a las noticias falsas. En cuanto a las estrategias de alfabetización informacional estas se reducen a compartir *checklist* pero no plantean técnicas como la lectura lateral o el llamado *Click*

Restraint, y la necesaria reflexión que es preciso realizar antes de la realización de *clicks* excesiva. Esta crítica a la falta de contenido elaborado y propio, la señala también Sullivan (2018), quien menciona que muchas de estas guías son una copia unas de otras, y se limitan a proveer definiciones generales y enlaces a *fact-checkers* u organismos de verificación.

Este autor (2018) defiende la integración de los profesionales de LIS (*Library and Information Science*) en los entornos de verificación de la información automatizados, mencionando el peligro de ceder estas herramientas a manos de empresas y algoritmos que nuevamente decidan ciertos entornos informativos. En este sentido, Buschman (2019) critica la aproximación neoliberal que tienen las herramientas de verificación y los riesgos y sesgos de la integración de la alfabetización informacional en ellas, y reivindica el papel de bibliotecas y archivos en el proceso. En este sentido destaca la obra de Kriplean, Bonnar, Borning, Kinney, y Gill (2014), quienes proponen el diseño de un sistema participativo de participación co-diseñado por profesionales del entorno de la Biblioteconomía y Documentación, con especial hincapié en la selección e interpretación de las fuentes, especialmente de las fuentes secundarias.

En este escenario sobre material de referencia en desinformación, es preciso incluir en la ecuación la inteligencia artificial. La relación entre bibliotecas e inteligencia artificial, si bien no es nueva, sí ha despegado recientemente con la avalancha de aplicaciones sobre inteligencia artificial generativa, especialmente de la aplicación de CHATGPT. Sus usos o potencial pueden resumirse en:

- Desarrollo de sistemas de recomendación en bibliotecas digitales (Delilovic, 2021; Elahi, Beheshti y Goluguri, 2021). Aparecen nuevas oportunidades en la investigación asociada al desarrollo de sistemas que analicen datos, patrones de usuarios y otras características con el fin de proporcionar recomendaciones relevantes.
- Mejoras en los chatbots, sistemas de referencia virtual y asistentes en búsqueda (Adetayo, 2023). Una tendencia incipiente es la de los usos de chatbots basados en IA de manera participativa y en tiempo real (Yao F, Zhang C y Chen W, 2015).
- Catalogación e indexación: técnicas de *Machine learning* ya han dado resultados en aplicaciones en catalogación y clasificación automática (Golub, Hagelback y Ardo, 2020). También puede aplicarse para la extracción automática de metadatos, normalización, establecimiento de recomendaciones personalizadas a partir de preferencias de usuarios...

IFLA, desde 2020, ya se posiciona sobre el impacto de la inteligencia artificial en bibliotecas. Así, el *IFLA Statement on Libraries and Artificial Intelligence* reflexiona sobre el papel de la inteligencia artificial en el entorno bibliotecario, sus implicaciones éticas, y cómo las bibliotecas pueden adaptarse y aprovechar esta tecnología para mejorar sus servicios. Un aspecto esencial que aborda este

documento fue el del incremento y mejora de las competencias de alfabetización sobre la misma, la denominada *AI literacy*. En 2023, el *Artificial Intelligence Special Interest Group*, publica un documento de trabajo sobre la respuesta estratégica por parte de bibliotecas ante el reto de la Inteligencia Artificial. La estrategia que más consenso obtiene es, precisamente, la de asumir un rol de liderazgo en el papel de esta alfabetización en AI. La experiencia pedagógica previa en el ámbito de Alfabetización informacional sirve para contextualizar qué habilidades son necesarias para comprender la AI, ayudar a su identificación (en el caso de manipulación informativa como *Deep Fakes*, por ejemplo), entender cómo se entrenan estos lenguajes, y cómo debe ser un uso ético de estas plataformas. Asimismo, el informe alude a la necesaria Alfabetización algorítmica, esencial para entender cómo la personalización de resultados adapta la información a nuestras preferencias individuales y puede afectar a garantizar un acceso pleno a la información.

IFLA (2023) reconoce que la IA constituye una amenaza para la creación automática y sin control de contenido, pudiendo ser un aliciente para la fabricación de más información falsa que contribuya a la polarización de la opinión pública. Esta conexión entre “bibliotecas, lucha contra la desinformación e inteligencia artificial” puede ser un nicho a tener en cuenta en los servicios de referencia de las bibliotecas, especialmente de las universitarias. No obstante, siguiendo la obra de Lim (2020) y tras un estudio preliminar a través del servicio y comunidad en línea *LibGuides Community Site*, orientada a la creación colectiva de estas guías bibliotecarias, se percibe que no se encuentra todavía bien contextualizada dentro de las guías temáticas analizadas.

En general, se encuentran recursos orientados a la explicación de conceptos básicos asociados con AI, usos, y necesidad de citación, así como ejemplos de uso en investigación o búsqueda de fuentes. Se encuentran algunas entradas sobre la necesaria verificación de los *prompts* obtenidos, no obstante, muy poco sobre recomendaciones específicas para ello. Se presentan los principales avances detectados:

- Evaluación crítica de las herramientas de IA. Algunas guías presentan checklist específicos de evaluación de estas herramientas. Estas se encuentran, por ejemplo en la Universidad de Toronto, que presenta un modelo llamado VALID-AI, compuesto por: *Validate data, Analyze algorithms, Legal and ethical considerations, Interpret how it Works, Diversity and bias, Accuracy check*. Otras insisten en la necesaria verificación posterior de la información aportada, sobre todo para prevenir el efecto de “alucinación” que pueden tener estas herramientas, que directamente constituyen falsedades (Emsley, 2023)
- Detección de *Deep Fakes*: generalmente las guías describen este fenómeno asociándolo con la desinformación, proporcionando material de consulta externo

- Descripción de Alfabetizaciones asociadas: es común que en estas guías aparezcan menciones expresas a la necesaria adquisición de competencias a través menciones diversas como: Alfabetización algorítmica, *AI Literacy*, *Prompt Literacy*... No obstante, se observa poca profundización en las competencias precisas de cada una de ellas. Sí se encuentra con una mayor descripción en recursos asociados con la vinculación de procesos de alfabetización informacional con la práctica de verificación de hechos, más ágil y adaptada al entorno digital como el método SIFT basada en cuatro movimientos: *STOP*, *Investigate the source*, *Find better coverage* y *Trace claims, quotes, and media back to the original context*, orientados a analizar y reconstruir el contexto en el que se produce la información digital de manera rápida (Caufield, 2019).

Conclusiones

El complejo problema de la desinformación y los desórdenes informativos requiere de un tratamiento poliédrico que va más allá de la propia verificación, sino de construir de manera colectiva relatos capaces de combatirla. Esto exige avanzar hacia un concepto de Alfabetización mediática más amplia, orientando a los lectores hacia el descubrimiento de nuevas fuentes y contextos en vez de simplemente desmentir los hechos.

En este contexto, destacamos la labor de las bibliotecas como instituciones y fuentes de autoridad, legitimando un nuevo rol en su función histórica de intermediación informativa, adaptada ahora a un entorno digital. Aprovecha el camino y experiencia en el ámbito de Referencia y Alfabetización informacional es fundamental para su participación en espacios de verificación de hechos, que deben ser más participativos y fomentar el descubrimiento de fuentes autorizadas.

En este contexto, la inteligencia artificial constituye un reto a la hora de mejorar y potenciar nuevas competencias en el marco de la Alfabetización: detectar qué información se produce con AI y no de manera humana, identificar posibles sesgos en el entramado de conjuntos de datos, demanda y conciencia mayor de una necesaria transparencia algorítmica, son competencias que serán clave en el futuro para la lucha contra la desinformación y las bibliotecas deben asumir un rol más activo en el fomento de este pensamiento crítico.

Bibliografía

- Adetayo, A. J. (2023), 'Artificial intelligence chatbots in academic libraries: the rise of ChatGPT', *Library Hi Tech News*, vol.40, no. 3, pp. 18-21.
- American Library Association (ALA) (2017) *Resolution on Access to Accurate Information*, viewed 5 December 2023, <<https://www.ala.org/advocacy/intfreedom/statementspols/ifresolutions/accurateinformation>>
- Anderson, R. (2017), 'Fake news and alternative facts: Five challenges for academic libraries', *Insights*, vol. 30, no. 2.
- Bangani, S. (2021), 'The fake news wave: Academic libraries' battle against misinformation during COVID-19', *The Journal of Academic Librarianship*, vol. 47, no.5, 102390.
- Beltrán, J., Míguez, R., & Larraz, I. (2021), 'ClaimHunter: An Unattended Tool for Automated Claim Detection on Twitter'. *KnOD'21 Workshop*, viewed 5 December 2023, <<https://www.semanticscholar.org/paper/ClaimHunter%3A-An-Unattended-Tool-for-Automated-Claim-Beltr%C3%A1n-M%C3%ADguez/764335faadf30051dea7f6b3747498035f91d0f8>>.
- Blanco Alfonso, I. (2020), 'Posverdad, percepción de la realidad y opinión pública: Una aproximación desde la fenomenología', *Revista de Estudios Políticos*, vol. 187, pp. 167-186.
- Bishop, B. (2009), *The big sort: Why the clustering of like-minded America is tearing us apart*. Boston, Mariner Books.
- Bowman, N. D., & Cohen, E. (2020), 'Mental shortcuts, emotion, and social rewards: the challenges of detecting and resisting fake news', in Zimdars, M. & Mcleod, K. (eds.), *Fake news: Understanding media and misinformation in the digital age*, pp. 224-230.
- Buschman, J. (2019), 'Good news, bad news, and fake news: Going beyond political literacy to democracy and libraries', *Journal of Documentation*, vol. 75, no.11, pp. 213-228.
- Capilla, P. (2019), '¿De qué hablamos cuando hablamos de posverdad? Análisis del término en siete diarios de calidad', *El Profesional de la Información*, vol. 28, no.3 <https://doi.org/10.3145/eipi.2019.may.09>
- Caridad Sebastián, M., Morales García, A. M., Martínez Cardama, S., & García López, F. (2018), 'Infomediación y postverdad: el papel de las bibliotecas', *Profesional de la Información*, vol. 27, no. 4, pp. 891-898, DOI <<https://doi.org/10.3145/eipi.2018.jul.17>>.
- Carrera, P. (2018), 'Estratagemas de la posverdad', *Revista latina de comunicación social*, vol. 73, no. 13, pp. 1469-1481.
- Caufield, M. (2019), *SIFT: the four moves*, viewed 5 December 2023, <<https://hapgood.us/2019/06/19/sift-the-four-moves/>>.
- Conner-Simons, A. (2018), 'Detecting fake news at its source'. *MIT News*, viewed 5 December 2023, <<https://news.mit.edu/2018/mit-csail-machine-learning-system-detects-fake-news-from-source-1004>>.
- D'Ancona, M. (2019), *Posverdad: la nueva guerra contra la verdad y cómo*

- combatirla. Madrid: Alianza.
- Del-Fresno-García, M. (2019), 'Desórdenes informativos: sobreexpuestos e infrainformados en la era de la posverdad', *El profesional de la información*, vol. 28, no. 3, e280302, DOI <<https://doi.org/10.3145/epi.2019.may.02>>.
- Delilovic, N. (2021), 'Recommender Systems in Digital Libraries Using Artificial Intelligence and Machine Learning: A Proposal to Create Automated Links Between Different Articles Dealing With Similar Topics', in: *Handbook of Research on Methodologies and Applications of Supercomputing*, pp. 162-178. Pensilvania: IGI Global.
- De Paor, S., & Heravi, B. (2020), 'Information literacy and fake news: How the field of librarianship can help combat the epidemic of fake news', *The Journal of Academic Librarianship*, vol. 46, no. 5, 102218, DOI <<https://doi.org/10.1016/j.acalib.2020.102218>>.
- Elahi, M., Beheshti, A., & Goluguri, S. R. (2021), 'Recommender systems: Challenges and opportunities in the age of big data and artificial intelligence', in: Sharaff, A., & Sinha, G. R. (eds.), *Data Science and Its Applications*, pp. 15-39. Boca Ratón: CRC press.
- Emsley, R. (2023), 'ChatGPT: these are not hallucinations—they're fabrications and falsifications', *Schizophrenia*, vol. 9, no. 1, 52, DOI <<https://doi.org/10.1038/s41537-023-00379-4>>.
- Fieiras-Ceide, C., Vaz-Álvarez, M., & Túnuez-López, M. (2022), 'Artificial intelligence strategies in European public broadcasters: Uses, forecasts and future challenges', *El Profesional de la Información*, vol. 31, no. 5, DOI <<https://doi.org/10.3145/epi.2022.sep.18>>.
- Flores Vivar, J. M. (2019), 'Inteligencia artificial y periodismo: diluyendo el impacto de la desinformación y las noticias falsas a través de los bots', *Doxa Comunicación*, vol. 29, pp. 197-212, DOI <<https://doi.org/10.31921/doxacom.n29a10>>.
- García-Marín, D. (2022), 'Modelos algorítmicos y fact-checking automatizado. Revisión sistemática de la literatura', *Documentación de las ciencias de la información*, vol. 45, no. 1, pp. 7-16, DOI <<https://doi.org/10.5209/dcin.77472>>.
- Golub, K., Hagelbäck, J., & Ardö, A. (2020), 'Automatic Classification of Swedish Metadata Using Dewey Decimal Classification: A Comparison of Approaches', *J. Data Inf. Sci.*, vol. 5, no. 1, pp. 18-38.
- Kahneman, D. (2011), *Thinking, fast and slow*. Nueva York: Macmillan.
- Kriplean, T., Bonnar, C., Borning, A., Kinney, B., & Gill, B. (2014), 'Integrating on-demand fact-checking with public dialogue', in: *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pp. 1188-1199.
- International Federation of Library Associations and Institutions (IFLA) (2017) *How To Spot Fake News*, viewed 5 December 2023, <<https://repository.ifla.org/handle/123456789/167>>.
- International Federation of Library Associations and Institutions (IFLA) (2020) *IFLA Statement on Libraries and Artificial Intelligence*, viewed 5 December

- 2023, <<https://repository.ifla.org/handle/123456789/1646>>.
- International Federation of Library Associations and Institutions (IFLA), Artificial Intelligence Special Interest Group (2023), *Developing a library strategic response to Artificial Intelligence*, viewed 5 December 2023, <<https://www.ifla.org/developing-a-library-strategic-response-to-artificial-intelligence/>>.
- Lim, S. (2020), 'Academic library guides for tackling fake news: A content analysis', *The Journal of Academic Librarianship*, vol. 46, no. 5 102195.
- Lor, P. J. (2018), 'Democracy, information, and libraries in a time of post-truth discourse', *Library Management*, vol. 39, no.5, pp. 307-321.
- Lozano, B. & Larraz, I. (2023), 'How AI-generated content could both fuel disinformation and improve fact-checking', in: *Poynter*, viewed 5 December 2023, <<https://www.poynter.org/commentary/2023/artificial-intelligence-help-harm-fact-checking/>>.
- Martínez-Cardama, S., & Algora-Cancho, L. (2019), 'Lucha contra la desinformación desde las bibliotecas universitarias', *Profesional de la información*, vol. 28, no. 4, DOI <<https://doi.org/10.3145/epi.2019.jul.12>>.
- Nicolás, J.A. (2022), 'De la desinformación a la posverdad', in: *Salto Diario*, viewed 5 December 2023, <<https://www.elsaltodiario.com/diccionario-posverdad/desinformacion-posverdad>>.
- Noelle-Neumann, E. (1993), 'La espiral del silencio. La opinión pública y los efectos de los medios de comunicación', *Communication & Society*, vol. 6, no. 1-2, pp. 9-28.
- Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. New York: Penguin.
- Reglitz, M. (2022), 'Fake news and democracy', *J. Ethics & Soc. Phil.*, vol. 22, pp. 162-187.
- Revez, J., & Corujo, L. (2021), 'Librarians against fake news: A systematic literature review of library practices (Jan. 2018–Sept. 2020)', *The journal of academic librarianship*, vol. 47, no. (2), 102304.
- Rodríguez-Pérez, C (2019), 'No diga fake news, di desinformación: una revisión sobre el fenómeno de las noticias falsas y sus implicaciones', *Comunicación*, no. 40, pp. 65-74.
- Saquete, E., Tomás, D., Moreda, P., Martínez-Barco, P., & Palomar, M. (2020), 'Fighting post-truth using natural language processing: A review and open challenges', *Expert systems with applications*, vol. 141, 112943.
- Simon, H. A. (1990), 'Bounded rationality', in: Eatwell, J., Milgate, M., Newman, P. (eds.), *Utility and Probability*, pp. 15-18. London: The New Palgrave. Palgrave Macmillan.
- Spitale, G., Biller-Andorno, N., & Germani, F. (2023), 'AI model GPT-3 (dis) informs us better than humans', in: *arXiv arXiv:2301.11924*, DOI <<https://doi.org/10.48550/arXiv.2301.11924>>.
- Sullivan, M. C. (2019), 'Why librarians can't fight fake news', *Journal of librarianship and information science*, vol. no. 4, pp. 1146-1156.
- UER (2019), 'EBU news report 2019. The next newsroom: unloking the power of

- AI for public service journalism', Geneva: EBU, European Broadcasting Union, viewed 5 December 2023, <https://www.ebu.ch/publications/strategic/login_only/report/news-report-2019>.
- University of Toronto Libraries (2023), *Artificial Intelligence for Image Research*, viewed 5 December 2023, <<https://guides.library.utoronto.ca/image-gen-ai/Home>>.
- Waisbord, S. (2020), '¿Es válido atribuir la polarización política a la comunicación digital? Sobre burbujas, plataformas y polarización afectiva', *Revista saap*, vol. 14, no. 2, pp. 248-279.
- Walter, N., Cohen, J., Holbert, R. L., & Morag, Y. (2020), 'Fact-checking: A meta-analysis of what works and for whom', *Political Communication*, vol. 37, no. 3, pp. 350-375.
- Yao, F., Zhang, C., & Chen, W. (2015), 'Smart talking robot Xiaotu: participatory library service based on artificial intelligence', *Library Hi Tech*, vol. 33, no. 2, pp. 245-260.

II

ATTIVITÀ FORMATIVA E DI RICERCA:
MIGLIORARE I RAPPORTI TRA I DUE PAESI

ACTIVIDAD FORMATIVA Y DE INVESTIGACIÓN:
MEJORAR LAS RELACIONES ENTRE LOS DOS PAÍSES

Coordinatori

Paola Castellucci (Sapienza Università di Roma)
José Luis Gonzalo Sánchez-Molero (Universidad Complutense de Madrid)

UNA PROPUESTA DE INTERNACIONALIZACIÓN PARA LOS PROGRAMAS DE DOCTORADO EN BIBLIOTECONOMÍA Y CIENCIAS DE LA DOCUMENTACIÓN ENTRE ESPAÑA E ITALIA

José Luis Gonzalo Sánchez-Molero*

En esta contribución se abordará la trayectoria de los estudios de Doctorado en Biblioteconomía y Documentación en España, haciendo una evaluación de la situación actual, para (y como no podía ser de otra manera en el marco del *II Seminario italo-spagnolo di Biblioteconomia e Documentazione*), proponer vías de colaboración entre ambos países en el campo de la investigación doctoral. Entrando en materia, y con respecto a la primera cuestión, en España se ofertan actualmente varios programas de doctorado, que a partir de 2015 se adaptaron al Real Decreto 43/2015¹. Por orden alfabético de universidad, el listado es el siguiente:

- Universidad de Alcalá de Henares: *Doctorado en Comunicación, Información y Tecnologías de la Sociedad en Red*.
- Universidad de Barcelona y Universidad de Zaragoza (conjunto). *Doctorado en Información y Comunicación*.
- Universidad Carlos III de Madrid: *Doctorado en Documentación, Archivos y Bibliotecas en el entorno digital*.
- Universidad Complutense de Madrid: *Doctorado en Ciencias de la Documentación*.
- Universidad de Granada: *Doctorado en Ciencias Sociales* (común para Antropología Social; Ciencia Política y de la Administración; Estudios sobre Paz, Conflictos y Derechos Humanos; Información y Documentación; Psicología Social; y Sociología).
- Universidad de Murcia: *Doctorado en Gestión de la Comunicación en las*

* Universidad Complutense de Madrid.

¹ España. Ministerio de Educación, Cultura y Deporte. Real Decreto 43/2015, de 2 de febrero, por el que se modifica el Real Decreto 1393/2007, de 29 de octubre, por el que se establece la ordenación de las enseñanzas universitarias oficiales, y el Real Decreto 99/2011, de 28 de enero, por el que se regulan las enseñanzas oficiales de doctorado. Boletín Oficial del Estado, Madrid, 2 feb. 2015. Disponible en: <<https://www.boe.es/boe/dias/2015/02/03/pdfs/BOE-A-2015-943.pdf>>

Organizaciones (común con Comunicación).

- Universidad de Salamanca: *Doctorado en Formación en la Sociedad del Conocimiento*.

No es nuestro propósito realizar aquí una evaluación y análisis de los contenidos de estos programas de Doctorado, nos vamos a limitar a exponer los factores que en los últimos años más han influido en su capacidad de atracción de estudiante. A este respecto, y en nuestra opinión, el número de doctorandos matriculados en cada uno de estos programas varía de manera muy notable en función de cinco aspectos:

- El total de estudiantes matriculados y egresados en las titulaciones previas de posgrado, o máster, de cada universidad, y en la misma área de conocimiento (Biblioteconomía y Documentación, o con su participación).
- El total de titulaciones de posgrado, o máster, oficiales ofertadas en cada universidad. Siempre en la misma área de conocimiento (Biblioteconomía y Documentación, o con su participación).
- La movilidad internacional y la cooperación con universidades europeas (especialmente Portugal) o iberoamericanas.
- El número de profesores en plantilla, con una capacidad investigadora reconocida a través de los sexenios obtenidos.
- Y, por último, la orientación del Doctorado de manera exclusiva hacia las disciplinas propias del área de Biblioteconomía y Documentación, o la inclusión de otras áreas en dichos programas.

Según datos del curso 2021-2022, distribuidos a través de RUID (la Red española de universidades con estudios en Biblioteconomía y Documentación) y matizados por sus responsables durante la reunión en Roma del *II Seminario*, son tres las universidades españolas que disponen de un mayor número de estudiantes de máster. En cifras totales, que incluyen tanto los estudiantes de nuevo ingreso como los matriculados por segunda vez, este sería el listado:

- Universidad Complutense de Madrid: 152
- Universidad de Barcelona: 125
- Universidad Carlos III de Madrid: 90

Las demás universidades con estudios de máster en nuestra área de conocimiento están en cifras por debajo de 50 estudiantes por curso académico. Con respecto al número de titulaciones de máster oficiales, la Universidad de Barcelona oferta un total de 5, la Complutense, 4, al Carlos III, 2, y el resto, 1, en el curso 2021-2022.

Se puede establecer una correlación entre que a mayor número de estudiantes de posgrado, así como a mayor número de titulaciones, aumenta la capacidad para atraer a investigadores que deseen emprender estudios de

Doctorado. Asimismo, es bastante evidente que la variedad de titulaciones de posgrado ofertada por cada una de las universidades citadas es un factor que facilita que las líneas de investigación para elaborar tesis doctorales sean más amplias y variadas, con el añadido de disponer de un profesorado especializado en múltiples campos. Estos aspectos, sin duda, también favorecen la atracción de doctorandos procedentes de áreas afines (como Historia, Historia del Arte o Comunicación), quienes pueden (y suelen) encontrar puntos de relación entre la Biblioteconomía y Documentación y sus estudios previos de posgrado.

Un tercer aspecto a tener en cuenta en la capacidad de atracción en nuestros programas de Doctorado en España radica en la cooperación institucional y en la movilidad internacional que se ofrezca, ya que la posibilidad de disponer de programas potentes en estos ámbitos favorece de manera muy notable la matriculación de doctorandos extranjeros. En España suelen provenir de países iberoamericanos. En el caso particular de la UCM esta cuestión ya ha sido explicada en su capítulo de este volumen por el profesor Antonio Carpallo. Un cuarto aspecto a tener en cuenta en el atractivo de nuestros programas de Doctorado radica en la oferta de plazas que se hace. En nuestro país se encuentra limitada (como no podía ser de otra manera) por la disponibilidad de profesorado suficiente para dirigir las futuras tesis doctorales. Para calcular esta capacidad, no sólo se valora el número de docentes, sino que también es determinante que estos hayan obtenido uno o más sexenios de investigación, según la legislación española.

La combinación de estos cinco factores ha tenido una gran influencia en el número de doctorandos matriculados en los programas vinculados al área de Biblioteconomía y Documentación. En el curso 2021-2022, los cuatro programas con una cifra mayor de estudiantes son los siguientes:

- Universidad Complutense de Madrid: 63
- Universidad Carlos III de Madrid: 45
- Universidades de Barcelona y de Zaragoza: 32
- Universidad de Granada: 14

La primera posición del programa complutense de Doctorado se explica (más allá de que esta universidad, al igual que la Carlos III, se localice en la capital del país), por sus cifras en el número previo de estudiantes de posgrado (uno de los altos de España) y de titulaciones de máster, dos aspectos en los que tanto la UCM como la UB destacan con relación al conjunto de universidades españolas, encontrándose ambas instituciones académicas muy parejas en sus cifras. Esto también se produce en los otros dos factores arriba indicados: la movilidad internacional y en el número de docentes con sexenios.

De igual manera, cabe destacar que, en el caso de la UCM, el número de profesores del área que participan en el programa del *Doctorado en Ciencias de la Documentación* es muy elevado y con perfiles variados. Los

docentes y futuros directores de tesis proceden tanto del departamento de Biblioteconomía y Documentación y de la Unidad docente de Bibliografía Literaria (en el departamento de Literaturas Hispánicas y Bibliografía, de la facultad de Filología). Asimismo, participan en su programa docentes del área afín de Ciencias y Técnicas Historiográficas. Esto le ha permitido ampliar la oferta de admisión de estudiantes en cada Curso de 16 a 25 doctorandos. En cambio, en el programa del *Doctorado en Información y Comunicación*, a pesar de ser conjunto entre dos universidades (UB y UNIZAR), la oferta es de 13. En los otros programas de Doctorado españoles (sin que su calidad sea menor) se ofrece un número reducido de plazas cada año, circunstancia que se puede atribuir a los factores antes citados (en especial al número de estudiantes y de titulaciones de posgrado existentes), pero que suele tener otra motivación. Al ser diseñados, o adaptados estos programas al RD 43/2015, sus responsables académicos optaron por incluir en las memorias de verificación unos números concretos. La idea era poder cumplir así con los porcentajes de matriculación exigidos para validar las reacreditaciones posteriores de los programas. Esta prudente preocupación fue el motivo también de que en 2015 se estudiaran la creación de programas de doctorado conjunto, como el de las universidades de Barcelona y de Zaragoza oferte un número de plazas aparentemente menor con respecto a la gran calidad de sus estudios y profesorado.

Esta breve introducción sobre la situación actual en España de los estudios de Doctorado en Biblioteconomía y Documentación, tras sus reformas y modificaciones entre 2013 y 2015, no da pie para plantear cuál ha de ser el camino, o caminos, que debemos seguir en la consolidación y expansión de nuestros estudios de Doctorado a las puertas del año 2023. Y celebrándose en Roma el citado *II Seminario Hispano-italiano* (2022), en presencia de colegas de todas las universidades italianas y españolas con estudios en Biblioteconomía y Documentación y/o en *Library and Information Science*, se comprenderá que mi propuesta (tampoco especialmente novedosa) sea la de avanzar en la internacionalización de estos estudios. No en vano, los programas de Doctorado son las titulaciones universitarias que mejor se adaptan académicamente a un proceso de internacionalización. Y esto se debe a varios motivos, que pasamos a enumerar y desarrollar.

1. Como es sabido, desde la implantación del *Espacio Europeo de Educación Superior* (vulgo Bolonia) se estableció una disparidad tanto temporal como en el número de créditos entre los planes de estudio de Grado y Máster entre España y la mayor parte del resto de países europeos: los Grados en nuestro país tienen una duración de 4 cursos (en Italia es de 3), y los Másteres, de 1 año (en Italia de 2). Esta disparidad dificulta muy notablemente la creación de estudios de Grado y de Máster conjuntos entre las universidades de ambos países, por ejemplo, o con otros que son incluso más cercanos geográficamente, como es Portugal, así como la

creación de dobles Grados interuniversitarios.

2. En cambio, tal disparidad no existe en los estudios de Doctorado, donde además los diversos programas de movilidad, tanto de estudiantes, como de investigadores y de profesorado, a través del Programa Erasmus Plus y de los proyectos europeos de REDES, entre otros, facilitan el intercambio científico.

Con esta base, ¿dónde y cómo establecer nuevos Doctorados conjuntos, y que tengan el propósito y la solidez suficientes para perdurar? Desde luego que ya existen sistemas tradicionales de colaboración entre programas de Doctorado de diferentes países, como son las tesis con mención internacional, pero parece más lógico aprovechar las oportunidades que nos ofrecen programas marco europeos, como lo fue *Horizonte 2020*, que integró, por primera vez, todas las fases de desarrollo científico, desde la generación del conocimiento hasta las actividades más próximas al mercado: investigación básica, desarrollo de tecnologías, proyectos de demostración, líneas piloto de fabricación, innovación social, transferencia de tecnología, pruebas de concepto, etc. Y que se ha continuado a través de *Horizonte Europa*, el nuevo programa marco de investigación e innovación (I+I) de la Unión Europea (UE) para el período 2021 -2027. *Horizonte Europa*, como antes su predecesor *Horizonte 2020*, está llamado a ser el instrumento fundamental para llevar a cabo las políticas de I+D+I de la UE. Recordemos que su objetivo general es alcanzar un impacto científico, tecnológico, económico y social de las inversiones de la UE en I+I, fortaleciendo de esta manera sus bases científicas y tecnológicas y fomentando la competitividad de todos los Estados Miembros (EEMM).

Asimismo, en el ámbito de la Unión Europea disponemos de una herramienta muy potente y que se está desarrollando. Es verdad que con cierta lentitud, por su complejidad, pero las expectativas son muy prometedoras. Me refiero a las alianzas entre universidades europeas, propiciadas y financiadas desde la UE. Se me permitirá poner un ejemplo que conozco bien, el de la alianza *UNA Europa* (hay una propuesta para que pase a denominarse en breve como *Una Futura*). Esta alianza agrupa a once universidades europeas (al menos en el momento en que estas líneas se escriben), y entre ellas se encuentran la Universidad Complutense de Madrid y la Università di Roma-La Sapienza y la Università di Bologna. Esta alianza tiene varias áreas de investigación prioritarias, estando dos de ellas relacionadas de manera muy clara con dos de nuestros ámbitos de investigación: Big Data e Inteligencia Artificial, por un lado, y Patrimonio Cultural, por otro. Se trata de dos áreas en las que las universidades que forman parte de la citada alianza llevan tiempo trabajando en su desarrollo y puesta en marcha de programas de doctorado interuniversitarios, así como de cursos masivos en formato *Mooc*.

En la actualidad, ya se ha implantado el programa de Doctorado *Una Europa Doctoral Programme in Cultural Heritage* (Una-Her-Doc), y en el curso 2022-23 se ha ofertado por vez primera un *Bachelor of Arts in European Studies* (BAES), que es una titulación de Grado.

Este contexto europeo y de las alianzas entre universidades nos abre un amplio abanico de posibilidades, que no debemos desaprovechar, con el propósito de visibilizar nuestro gran bagaje investigador en lo que denominamos internacionalmente como *LIS*. Los programas marco de la UE y las alianzas entre universidades europeas nos ofrecen tanto financiación como plataformas académicas. Es cierto, que proponer, diseñar y participar en estudios conjuntos de Doctorado supone mucho trabajo de tramitación y de gestión, y que hay aspectos a favor y en contra que deberemos valorar. Sin embargo, si queremos articular una Bibliografía “mediterránea”, fundamentada en los vínculos que estamos estableciendo entre las universidades de Italia y de España, y que estos vínculos tengan fundamentos más sólidos y de calidad, la formación universitaria, a través de la participación en programas de doctorado conjuntos, y el intercambio de investigadores entre ambos países, son las alternativas que debemos seguir.

Entre las universidades italianas y españolas debemos avanzar en la propuesta de un doctorado internacional, que conduzca a la obtención de títulos de doctorado conjuntos, dobles o múltiples. Las colaboraciones entre nuestros profesores e investigadores, a uno y otro lado del Mediterráneo occidental, constituyen una base muy valiosa para empezar a trabajar en su diseño, pero debemos emprender una estrategia conjunta para darle una formulación más completa, que debería incluir los siguientes elementos:

1. Publicaciones internacionales conjuntas, ya sea a través de revistas indexadas, o de monografías. Las actas de los dos seminarios Hispano-italiano son un primer paso.
2. Organización de encuentros y de congresos internacionales, con la asistencia de los doctorandos, lo que permitirá una mayor capacidad de integración y de colaboración para el futuro.
3. La participación de profesores/investigadores españoles e italianos en los claustros de los programas de Doctorado en ambos países, para cotutelar tesis doctorales, o formar parte de los tribunales de tesis. Las universidades deben tener firmados convenios marco de cotutela.
4. Favorecer la admisión de estudiantes internacionales en los programas de doctorado de ambos países. La acción *Co-Fund* del programa Marie-Curie de la Comisión Europea financia becas predoctorales y postdoctorales.
5. Favorecer la realización de estancias internacionales. Además de las becas de movilidad asociadas a los contratos FPU, FPI y PIF, los doctorandos, como estudiantes que son, también pueden optar a otras becas de movilidad, tanto en la modalidad estudios como prácticas.

6. Promover la mención Internacional del título de Doctor, con estancias entre nuestros dos países.

Para el desarrollo de estas iniciativas, teniendo en cuenta su grado de participación y de complejidad, en la Unión Europea disponemos de programas para facilitar su desarrollo y, naturalmente, su financiación. Me limitaré a recordar la más importantes: la acción *Innovative Training Networks* (ITN), de la Comisión Europea (Programa Marie-Curie) que tiene como objetivo formar una nueva generación de investigadores, en fase predoctoral, capaces de transformar sus ideas en productos y servicios que redunden en beneficios económicos y sociales. Esta acción financia programas conjuntos de formación y/o de doctorado, puestos en marcha por una red internacional de centros públicos y privados. Se contemplan tres tipos de redes:

- *European Training Networks* (ETN): red formada por al menos tres entidades independientes de tres países miembros de la UE, o países asociados, diferentes. No existen requisitos en cuanto al consorcio, pero se estima un tamaño de 6-10 socios y se recomienda que sean de los sectores académico y no académico.
- *European Industrial Doctorates* (EID): red formada por al menos dos entidades independientes de dos países miembros o asociados distintos y de dos sectores diferentes. Al menos uno de ellos debe ser una institución académica (ej. universidad) y otro una institución no académica (ej. sector empresarial, a ser posible una PYME). Cada investigador participante debe estar inscrito en un programa de doctorado y pasar al menos un 50% de su tiempo en instituciones del sector no académico. En nuestro caso, quizá sería posible enfocar estas redes a través de su vinculación con la industria editorial, que es especialmente potente tanto en España como en Italia.
- *European Joint Doctorates* (EJD): red formada por al menos tres entidades independientes de tres países miembros o asociados distintos para la creación de programas de doctorado conjuntos. Los miembros de la red deberán ser principalmente instituciones académicas con capacidad para otorgar títulos de doctorado. La red debe poder conceder títulos de doctorado conjuntos, dobles o múltiples, otorgados por instituciones de al menos dos países europeos.

En definitiva, disponemos de las suficientes herramientas para emprender este camino, que iniciamos en 2017, para establecer un marco de comunicación e intercambio académicos en nuestra área de conocimiento. El extraordinario interés, entusiasmo y colaboración ponen de manifiesto relieve que existe un notable interés por dar continuidad a este proyecto conjunto. Los encuentros en Perugia, Madrid, Turín y Roma han permitido constatar que en nuestro particular *Mare Nostrum*, las Ciencias de la Documentación pueden fijarse nuevos rumbos.

INIZIATIVE DI RAFFORZAMENTO DELL'INTERNAZIONALIZZAZIONE DELLA DIDATTICA IN UN CORSO DI LAUREA

Fiammetta Sabba*

L'internazionalizzazione della didattica

Tra gli strumenti della formazione superiore sui quali l'Italia sta investendo sempre di più per accrescere la qualità delle università vi è l'internazionalizzazione. Si tratta di una dimensione che sta ricevendo grande interesse non soltanto a livello della produzione e della disseminazione dei risultati scientifici, ma anche sul versante della didattica, per preparare in particolare gli studenti, futuri professionisti o studiosi, a percorsi aperti e globalizzati, e per arricchire il bagaglio di docenti e studiosi e dello staff accademico sia sul piano culturale sia su quello tecnico delle metodologie di trasferimento della conoscenza.

Lo scenario dell'apertura internazionale è completamente mutato nell'ultimo trentennio, con un incremento dei criteri e dei parametri per misurare l'internazionalizzazione degli atenei che hanno evidenziato più di tutto la mancanza di reciprocità degli scambi, ossia una vitale presenza di ricercatori e studenti italiani all'estero contro pochi 'internazionali'¹ presenti nei nostri atenei. Ciò ha spostato l'attenzione dal dibattito nazionale a quello europeo, e inoltre ha portato a sviluppare iniziative di più ampio respiro, che hanno interessato l'area geografica non solo europea, ma anche sudamericana, asiatica e mediterranea.² A livello europeo, quindi, sono stati potenziati i tipici

* Università di Bologna.

¹ Vengono definiti "studenti internazionali" coloro i quali varchino frontiere a scopo di studio per realizzare in tutto o in parte la loro formazione superiore. L'Unesco e il suo istituto di statistica, l'OECD ed Eurostat definiscono "international students" quanti non risiedono nel Paese in cui vanno studiando o che hanno ricevuto la loro formazione superiore precedente in un Paese diverso da quello in cui vanno studiando.

² Non sono trascurabili i benefici economici, come si legge nel rapporto sull'internazionalizzazione pubblicato dalla CRUI nel 2018 a cura di Fabio Ruggie (<https://www2.crui.it/crui/crui-rapporto-inter-digitale.pdf>).

programmi Erasmus, Tempus, Leonardo ecc., ma anche quel processo di armonizzazione delle architetture dei sistemi universitari noto come *Processo di Bologna*,³ ovviamente in stretta aderenza all'*Agenda di Lisbona*.⁴

Le riforme della didattica, che si sono susseguite nel nostro Paese in tutto questo tempo intercorso dall'inizio di una riflessione europea condivisa ad oggi, contengono alcune disposizioni chiave per spingere le università ad aprirsi alla dimensione globale del sapere con maggiore incisività rispetto al passato. In particolare, queste riguardano e consentono: la possibilità di rilasciare titoli congiunti con atenei stranieri; il riconoscimento dei periodi di studio, dei crediti e dei titoli conseguiti all'estero ai fini della prosecuzione degli studi in attuazione della Convenzione di Lisbona; lo studio obbligatorio, valutato in crediti, di almeno una lingua dell'Unione europea, oltre all'italiano; la possibilità di sostenere l'esame conclusivo del corso anche in una lingua straniera.

Tuttavia, trattandosi di una prospettiva diventata così strutturale, è necessario che le singole istituzioni immaginino per l'internazionalizzazione delle attività coerenti in un quadro strategico complessivo, con il preciso scopo di favorire la crescita delle relazioni internazionali sostenendo la mobilità della comunità universitaria tutta (studenti, docenti, ricercatori e staff).

Sul piano della didattica, infatti, ci sono precisi obiettivi da perseguire in questo ambito:

- rendere più attrattivi i corsi di studio agli studenti stranieri, ma soprattutto preparare e formare gli studenti italiani ad un'apertura verso il mondo lavorativo e scientifico a livello internazionale;
- diversificare l'offerta formativa in ragione non solo di logiche attrattive, ma anche di prospettive culturali di largo raggio nella preparazione degli studenti;
- offrire la multiculturalità come capacità di rapportarsi e di interagire con tradizioni storiche, comportamentali e scientifiche anche molto diverse;
- avvicinare al multilinguismo come capacità di padroneggiare codici linguistici diversi, in funzione della loro efficacia comunicativa (locale, nazionale, sovranazionale).

³ Il Processo di Bologna nasce nel 1999 come accordo intergovernativo di collaborazione nel settore dell'Istruzione superiore, <<https://www.miur.gov.it/processo-di-bologna>>.

⁴ Convenzione sul riconoscimento dei titoli di studio relativi all'insegnamento superiore nella Regione europea, siglata a Lisbona l'11 aprile 1997, <<https://miur.gov.it/convenzione-di-lisbona>>.

Il case study: il Corso di Laurea Magistrale in “Scienze del libro e del documento” dell’Università di Bologna

L’INTERNAZIONALIZZAZIONE DELLA DIDATTICA DEL CORSO DI LAUREA

Per rendere concretamente l’idea delle azioni e delle attività di internazionalizzazione che si possono svolgere all’interno di un corso di laurea, e al fine di far emergere questioni, criticità, e buone metodologie e pratiche per una internazionalizzazione concreta ed equilibrata, si porta il caso del Corso di Laurea magistrale in Scienze del libro e del documento (classe ministeriale LM/5 in Archivistica, Biblioteconomia e Bibliografia) offerto dal Dipartimento di Beni Culturali dell’Università di Bologna, situato nel Campus di Ravenna (uno dei cinque campus dell’Alma Mater).⁵

L’Università di Bologna, come gli altri atenei, aderisce ai programmi di scambio europei e transnazionali, in particolare si tratta di: Erasmus+ Studio; Erasmus+ Mobilità per tirocinio; Overseas; a cui si aggiungono anche bandi per borse di studio per preparazione di tesi magistrali all’estero, e l’opportunità specifica dell’Alma Mater della mobilità presso la sede UniBO di Buenos Aires.

Tuttavia, profittare di questo ampio ventaglio di offerte non è esente da criticità, in quanto gli studenti sono spesso disorientati da bandi diversi da analizzare, numerose informazioni da metabolizzare, scadenze molto anticipate rispetto ai tempi di partenza, requisiti linguistici richiesti da conseguire in tempi stretti e con procedure particolari.

Le soluzioni che il corso di laurea in Scienze del libro e del documento ha messo in atto – in particolare per l’iniziativa maggiormente strutturata che è quella dell’Erasmus+ per studio ma estendibili ad altre esperienze di scambio del genere – sono di due tipi, generali (a-c) e specifiche (d-e):

- a. organizzare incontri preliminari illustrando l’opportunità offerta e i suoi vantaggi e preannunciando le scadenze e favorendo il contatto con studenti già partiti con il programma Erasmus;
- b. organizzare incontri formativi con carattere tecnico (compilazione delle domande e accompagnamento al conseguimento dei requisiti linguistici);
- c. informare tempestivamente gli studenti dell’uscita dei bandi;
- d. organizzare incontri con referenti stranieri per promuovere gli accordi specifici stretti da docenti del corso di laurea e per illustrare le sedi, i master disponibili, i programmi in corso, le iniziative extracurricolari, i progetti scientifici delle strutture, le opportunità post laurea;
- e. organizzare una pagina web⁶ che informi sugli accordi più congruenti con

⁵ Sito del Corso di laurea al link <<https://corsi.unibo.it/magistrale/ScienzeLibro/index.html>>.

⁶ Pagina web degli accordi al link <<https://corsi.unibo.it/magistrale/>>

gli obiettivi del corso di studio di appartenenza e sui nominativi e i recapiti dei docenti-referenti di ciascun accordo, ma soprattutto sui corsi e sugli insegnamenti specifici per comporre adeguatamente i *Learning Agreements (L.A.)*.⁷ Alla creazione di questa specifica pagina, è seguita anche una riorganizzazione di quella dipartimentale, utile ad accompagnare gli studenti in tutte le fasi tecniche più generali in risposta al bando.⁸

Riguardo a quest'ultimo punto le attenzioni da avere sono molte; è necessario, infatti, valutare:

- la corrispondenza dei tempi di erogazione degli insegnamenti di interesse con quelli del soggiorno all'estero dello studente;
- la corrispondenza almeno di massima tra gli insegnamenti inseriti nel *L.A.* e quelli disponibili nel proprio corso italiano, secondo quanto regolato dal piano didattico della propria coorte di appartenenza.⁹ Una redazione attenta del *L.A.* è, infatti, fondamentale per poter avere, al rientro in Italia, il pieno riconoscimento dei crediti formativi conseguiti. Tuttavia, è necessario al contempo garantire allo studente un po' di flessibilità nella scelta degli insegnamenti, così da consentirgli di selezionare dall'offerta formativa straniera anche insegnamenti caratterizzati da tematiche culturali specifiche del paese ospitante o da aspetti affrontati in modo più o meno approfondito o mirato, purché siano fatti salvi gli obiettivi disciplinari di fondo dell'insegnamento; questo permette di aderire concretamente, infatti, ad alcuni dei principi su cui si impernano queste esperienze di scambio, che sono quelli della multiculturalità e della multidisciplinarietà con implicazioni importanti sia sull'originalità metodologica della didattica sia ovviamente sui contenuti in termini di argomenti, loro rappresentazione, e visioni.¹⁰

ScienzeLibro/accordi-erasmus-specifici> (ultima cons. il 19 gennaio 2024).

⁷ Il *L.A.* viene stipulato prima della partenza ed è da considerare come un contratto con cui il Corso di Studio si impegna a riconoscere le attività svolte all'estero una volta rientrati in Italia. Nel *L.A.* vanno indicate le attività formative che si intendono svolgere presso l'università ospitante e i corsi corrispondenti dell'università di partenza, con i relativi crediti. Occorre confrontarsi con il docente responsabile dello scambio sulle attività formative da scegliere, dopo essersi informati sull'offerta didattica della sede ospitante (consultando ad esempio il sito internet) e rispettando il più possibile l'area disciplinare dello scambio pubblicata nel bando.

⁸ <<https://beniculturali.unibo.it/it/didattica/mobilita-internazionale>>.

⁹ La 'coorte' identifica l'anno di ingresso dello studente all'interno del Corso di Studio; i piani didattici possono infatti avere un ordinamento e un regolamento didattici diversi a seconda degli anni di iscrizione, perché possono essere intercorse modifiche.

¹⁰ Per fare esempi chiarificatori, uno studente che si rechi in Spagna può selezionare un insegnamento denominato 'El libro antiguo en Asia y America' per uno tra quelli italiani di 'Storia della Bibliografia', o di 'Bibliologia' o ancora di 'Storia della

- una sufficiente corrispondenza tra il numero dei cfu¹¹ di ciascun insegnamento straniero selezionato e quello dell'insegnamento italiano; può essere, infatti, necessario seguire due o più insegnamenti o attività formative (ad esempio laboratori) per arrivare alla adeguatezza quantitativa di cfu, per i quali è ammesso l'esubero ma non la carenza, pena il loro non riconoscimento al rientro in Italia.

Soffermandosi, all'interno del blocco "Opportunità" del sito web del corso, sulla pagina web "Accordi specifici", organizzata dai docenti che collaborano alla *Quality Assurance* del corso di laurea, si vedrà che essa è arricchita da una mappa esplicativa degli accordi Erasmus più congruenti con gli obiettivi formativi del corso. Va, infatti, precisato che l'ateneo bolognese dà la possibilità agli studenti di scegliere tra tutti gli accordi internazionali attivati dall'istituzione; tuttavia, si è constatato come questo confondesse gli studenti nella scelta e come rappresentasse poi uno scoglio per loro e per i docenti responsabili degli accordi nel momento dell'individuazione di insegnamenti utili al *L.A.*. Pertanto, si è deciso di sostenere la loro domanda di partenza fornendo una pagina web specifica con le informazioni di base, quali appunto il quadro degli accordi e il relativo dettaglio sui master più indicati attivi nelle diverse sedi straniere, o perché 'gemelli' del corso italiano o perché comunque molto corrispondenti ad esso.

stampa'; oppure 'Etica y deontologia del documento' come corrispondente a quello italiano di 'Diritto dei beni culturali'.

¹¹ I cosiddetti 'cfu' (crediti formativi universitari) a livello internazionale corrispondono ai crediti ECTS, che misurano la quantità di lavoro svolto dallo studente attraverso l' ECTS (European Credit Transfer System), che è un sistema chiaro di definizione del trasferimento dei crediti basato sulla trasparenza dei risultati delle attività svolte e su quella dei processi di apprendimento permanente.

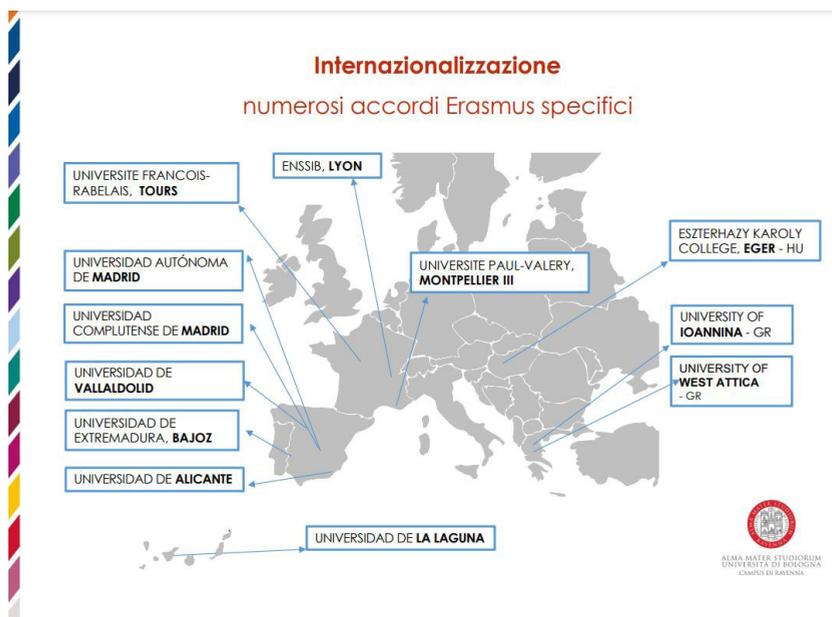


Fig. 1. Mappa degli accordi Erasmus del corso LM/5 UNIBO

Si noterà che sono carenti per l'area tedesca e quella anglosassone, e non in quanto i docenti del corso di laurea non abbiano rapporti scientifici con istituzioni di quei paesi, ma in quanto in essi non risulta la tipologia di corso equivalente alla nostra laurea magistrale, ossia non sono attivi master biennali nella preparazione delle figure di archivista, bibliotecario o affini, demandata a una formazione di tipo diverso, spesso interna agli stessi istituti culturali centrali. Dalla mappa è evidente come invece la Spagna, in primo luogo, e poi a seguire la Francia, siano, in questo specifico ambito, i paesi che condividono con l'Italia uno stesso modello formativo sia professionale sia scientifico.

Gli studenti italiani *outgoing* sono molti, non spaventati dalla barriera linguistica soprattutto nel recarsi in Spagna e Francia, ma anche in Grecia dove gli insegnamenti vengono spesso erogati in inglese; mentre si riscontra una scarsa presenza di studenti *incoming* a Ravenna, principalmente per due motivi: uno comune al resto dell'Italia che è quello della lingua in cui vengono impartite le lezioni frontali, e uno invece specifico che ha a che fare con la forte attrazione esercitata dalla sede centrale di Bologna (in parte tuttavia mitigata dall'attrattiva del mare, dal livello ottimo dei servizi, e dalla accogliente lentezza della provincia della città bizantina).

Soffermandoci sull'elemento più generale, la lingua, è accertato che la conoscenza, sia pur elementare, della lingua italiana può risultare un fattore

decisivo di internazionalizzazione. Anzitutto essa può permettere a uno studente internazionale di accedere alle attività formative erogate nella nostra lingua, e così di scegliere tra un'amplissima offerta di percorsi di studio. In secondo luogo, essa può offrire a chi studi in corsi erogati in inglese di scartare qualche difficoltà collegata a quei percorsi nel momento in cui essi richiedano di interagire con l'ambiente oltre l'aula (istituzioni culturali, laboratori, centri ecc.). In terzo luogo, la conoscenza dell'italiano consente allo studente internazionale di trarre un più esteso vantaggio dell'esperienza culturale in Italia. La estrema vicinanza tra l'italiano e la lingua spagnola è un altro fattore di vantaggio, per esempio, per studenti provenienti dalla Spagna e dal Sud e dal Centro America.

Iniziative di internazionalizzazione parallele alla didattica

Il corso di laurea di Ravenna è attivo dal 2014 e nel suo percorso ha già vissuto una modifica di ordinamento e una di regolamento normativo, per adeguare il piano didattico alle novità disciplinari e a quelle professionali, per accogliere nuovi docenti che sono stati assunti per irrobustire l'offerta formativa, tra cui docenti internazionali (quasi sempre italiani rientrati dopo una lunga carriera all'estero attraverso i vari canali di reclutamento a loro dedicati). Grazie proprio all'impulso del corpo studioso e docente, il corso si è visto affiancare nell'offerta didattica da numerose iniziative, come quelle di una rivista scientifica, un centro-laboratorio, un ciclo di seminari di approfondimento e aggiornamento, tre summer school, un'esperienza Erasmus Plus KA2, e un gruppo di ricerca internazionale.

L'apertura delle iniziative succitate a studenti, assegnisti e dottorandi è paradigmatica di una scelta consapevole del corpo studioso e docente coinvolto nel corso di laurea, ossia del perseguimento di una convergenza di ricerca e didattica.

La prima esperienza è quella di *Bibliothecae.it*, rivista accademica semestrale di ambito internazionale valutata di classe A dall'ANVUR per i settori concorsuali 11/A1 (storia medievale), 11/A2 (storia moderna), 11/A3 (storia contemporanea), 11/A4 (scienze del libro e del documento e scienze storico religiose).¹² *Bibliothecae.it* è attiva fin dal 2012, e dal 2015 è ospitata nella piattaforma AlmaDL dell'Università di Bologna che la pubblica mediante il software OJS3. Promuove e ospita studi e recensioni nell'ambito delle discipline del libro e del documento, quindi in particolare sui temi di libri, documenti, testi, biblioteche, editoria, informazione e comunicazione scientifica, con lo scopo di contribuire all'approfondimento e allo sviluppo dei processi di organizzazione e diffusione della conoscenza culturale. Gli studenti

¹² <<https://bibliothecae.unibo.it>>.

vengono coinvolti assegnando loro recensioni delle novità bibliografiche, mentre i dottorandi, maggiormente esperti dei primi, vengono invitati oltre che a scrivere e recensire anche a collaborare professionalmente al processo redazionale ed editoriale digitale così che acquisiscano delle competenze utili insieme alle conoscenze.

Un'altra attività collegata al corso di laurea è quella del ciclo di seminari *Di carte, di libri e dintorni*,¹³ iniziativa avviata fin pochi anni dopo l'istituzione del corso di laurea da Paola Degni e da chi scrive, alla cui organizzazione si sono poi aggiunti i colleghi del ssd. M-STO/08 Stefano Allegrezza e Lucia Sardo. Si tratta di un ciclo annuale di conferenze che ha lo specifico obiettivo di approfondire per gli studenti, attraverso l'intervento di esperti (italiani e stranieri) e al di fuori della consueta cornice didattica, temi e questioni che riguardano i documenti e il libro antico e a stampa, la loro conservazione e la fruizione, anche in rapporto alle nuove tecnologie. Negli ultimi anni, complice anche l'uso strategico dello streaming in seguito alla pandemia, e anche in un'ottica di Terza missione, il ciclo è stato aperto a tutti e si è svolto in modalità mista, ossia in presenza e da remoto su zoom o teams.

Sul fronte della didattica, ma con una forte impronta di Terza missione, sono state organizzate poi alcune summer school di ambito biblioteconomico e archivistico. Esse non rilasciano cfu, ma un attestato di partecipazione, proprio perché si configurano come Terza missione. Si tratta, infatti, di occasioni molto utili per preparare i partecipanti su temi e abilità tecniche emergenti nella professione documentaria in senso ampio e a livello internazionale, ma anche per creare una rete collaborativa, rinnovata ogni anno, che si rivela fondamentale per portare avanti questioni comuni, per immaginare insieme progetti, per aiutare studenti, assegnisti e dottorandi a immergersi nel mondo del lavoro con un profilo maggiormente competitivo e sostenuti da una rete di soggetti d'interesse con la quale sono già entrati in contatto.

La prima scuola che è stata organizzata, diretta da chi scrive, è la Summer School "Linked data per i beni culturali",¹⁴ che ha avuto sei edizioni (2016-2021), coinvolgendo vari enti, quali Fondazione Flaminia, Pratt Institute School Information (New York), New York Public Library, Istituto per i Beni Culturali dell'Emilia-Romagna, e AIB. L'attività formativa si incentrava sui principi e sulle tecniche dei *Linked data* applicati alla modellazione concettuale, alla descrizione e al riuso di oggetti culturali, offrendo una introduzione teorica

¹³ Al seguente link sono disponibili i cicli degli ultimi anni accademici <<https://centri.unibo.it/ludi/it/centro/cosa-facciamo>>; per i precedenti si rinvia alle locandine rintracciabili nel sito del Dipartimento di beni Culturali, poiché non esisteva ancora il LUDI, che è diventato il contenitore delle iniziative di ambito Library and Information Science promosse dal Dipartimento.

¹⁴ <<https://site.unibo.it/linked-data-per-i-beni-culturali/it>>.

e pratica ai più recenti metodi di rappresentazione delle risorse conservate in biblioteche, archivi e altri tipi di istituzioni culturali, con il fine di accrescerne la visibilità e facilitarne l'accesso.

Dopo aver formato oltre 200 persone sui *Linked data*, i temi della Summer School nel 2022 sono stati riformulati, in quanto le necessità professionali stavano mutando e diventando molto specifiche sotto l'impulso del PNRR Cultura¹⁵ e, in particolare, del PND¹⁶. È così partita una nuova scuola estiva intitolata "La Digital Library: evoluzione, strutture, progetti",¹⁷ come naturale prosecuzione e sviluppo del precedente percorso formativo. Questa Summer School intende fornire ai partecipanti una visione generale sulle *Digital Libraries*, sulla loro evoluzione storica e sulle caratteristiche che hanno assunto nella realtà contemporanea. Alla parte introduttiva di tipo teorico vengono affiancate, oltre ad una panoramica di esperienze nazionali di buone pratiche, in particolare alcune attività laboratoriali-progettuali per fornire ai partecipanti le competenze di base per poter progettare e realizzare una *Digital Library*.

L'ultima Summer School, organizzata invece dal docente e studioso Stefano Allegrezza, attiene all'ambito archivistico ed è denominata "Web and social media archiving and preservation".¹⁸ Questa scuola intende offrire una formazione di alto livello sui temi emergenti dell'archiviazione e conservazione dei siti web e dei social media, che rappresentano una nuova e diversificata tipologia di materiale la cui conservazione è imprescindibile per tutta una serie di ambiti scientifici ai fini della futura ricostruzione dell'attuale civiltà. La Summer School intende anche fornire le conoscenze e le competenze necessarie per favorire lo sviluppo di nuove professionalità ed avviare nuovi percorsi lavorativi da parte dei discenti interessati.

Merita poi segnalare, a testimonianza delle aperture internazionali dinamiche sul piano didattico, il progetto di partenariato strategico Erasmus+ dal titolo *Die Zukunft des kulturellen Erbes im modernen Europa*,¹⁹ guidato

¹⁵ <<https://pnrr.cultura.gov.it>>, cui si aggiunga <<https://digitallibrary.cultura.gov.it/pnrr-cultura/>>.

¹⁶ <<https://digitallibrary.cultura.gov.it/il-piano/>>.

¹⁷ <<https://site.unibo.it/digital-library-evoluzione-strutture-progetti/it>>.

¹⁸ <<https://site.unibo.it/web-and-social-media-archiving-and-preservation/it>>.

¹⁹ Di esso si dirà in modo più esteso in altra pubblicazione promossa dalla Società italiana di Scienze biblioteconomiche e bibliografiche e contenente gli atti del convegno tenutosi a Cagliari a settembre 2022 per il suo decennale. Intanto, questi i link utili per approfondire il progetto, le attività dell'unità bolognese, la summer school e la mostra organizzate, e il volume pubblicato: <<https://www.uni-potsdam.de/de/fucuhe/projekt/uebersicht>>; <<https://site.unibo.it/fucuhe/it/agenda/sommerschule-2021-bologna-als-st-tte-des-europ-ischen-kulturerbes-universit-t-bibliothek-und-museen-auf-dem-weg-in-die-moderne>>; <<https://site.unibo.it/summerschool-kulturerbe/it>>; <<https://bub.unibo.it/it/mostre-online/il-patrimonio-culturale-di-bologna-allo-specchio>>.

dall'Università di Potsdam e al quale nel triennio 2019-2022 hanno partecipato otto università europee, tra cui Bologna. Il progetto ha inteso sviluppare procedimenti innovativi nell'analisi scientifica e didattica del patrimonio culturale europeo - in particolare quello letterario, composto da manoscritti, stampe rare e libri di vario tipo - condividendoli tra le diverse sedi. Guida nel *modus operandi* è stato un modello di lavoro partecipativo che coinvolgesse in tutte le attività, accanto ai ricercatori e agli esperti, anche gli studenti, e nello specifico per l'apporto di Bologna: in una summer school, nei tutorati di letteratura tedesca, in un corso in modalità *moodle* di storia della cultura tedesca, nei tirocini, e nella mostra e nel volume prodotti attraverso quelle esperienze.

Presso il Dipartimento di beni culturali è, inoltre, attivo un Gruppo di ricerca dal nome *Bibliotheca disciplinata*, di cui è capofila l'Universidade de São Paulo del Brasile. Il gruppo, attivo fin dal 2015, è composto da studiosi brasiliani, spagnoli, portoghesi, francesi e italiani, e si occupa di ricerche relative alla disciplina della Bibliografia (intesa anche come storia del libro, e organizzazione e mediazione della conoscenza) mediante un approccio storico per affrontare le questioni metodologiche e per studiare le pratiche contemporanee in modo condiviso, collaborativo e connesso. I forum e le pubblicazioni prodotte hanno visto coinvolti anche gli studenti come pubblico e i dottorandi nell'organizzazione.

Alla fine di questa esposizione delle attività e delle iniziative che il gruppo docente del corso di laurea ha costruito non può venire dimenticato il più giovane progetto, quello del centro-laboratorio "LUDI".²⁰ Istituito nel 2021, inizialmente per fare da cornice a tutte le iniziative presentate, ne è diventato poi l'organo propulsore. La sua ideazione si impernia sulle esigenze e sulle attività di ambito M-STO/08 (Archivistica, Biblioteconomia e Bibliografia), ma, pur partendo dal comparto di *Library and Information Science*, ne è caratteristica fondativa la forte interdisciplinarietà promossa, come può vedersi dal collegio scientifico che lo compone.²¹ Il laboratorio si basa innanzitutto su una visione inclusiva e articolata dell'Archivistica, della Bibliografia, della Biblioteconomia e delle discipline del libro manoscritto, e pone speciale attenzione alla ricerca storica sui temi connessi a questi ambiti e agli sviluppi disciplinari contemporanei; tuttavia, esso si propone anche di promuovere lo scambio e la ricerca fra i vari sottocampi disciplinari e di favorire il dialogo con discipline affini per una maggiore contestualizzazione sociale e culturale

dei-viaggiatori-europei>; <<https://buponline.com/prodotto/il-patrimonio-culturale-della-biblioteca-universitaria-di-bologna-e-della-citta-allo-specchio-dei-viaggiatori-europei/>>.

²⁰ <<https://centri.unibo.it/ludi/it>>.

²¹ <<https://centri.unibo.it/ludi/it/centro/collegio-scientifico>>.

e una più ampia visione scientifica delle attività. A tal fine attiva collaborazioni con enti ed istituzioni culturali e con altri centri e laboratori di cui sono membri i partecipanti al comitato scientifico. Il laboratorio, inoltre, persegue l'obiettivo di costituire un punto di riferimento per le tematiche relative alla valorizzazione e allo studio delle raccolte librerie e documentarie analogiche e digitali non soltanto all'interno al Dipartimento di cui è struttura integrante, ma anche nell'Ateneo. Intercetta, infine, le esigenze del territorio e favorisce l'intersezione tra saperi e competenze diversificati nazionali e internazionali.

Conclusioni

L'università svolge, dunque, un ruolo fondamentale nella dimensione internazionale anche attraverso le attività di cooperazione internazionale allo sviluppo, che vanno dalla formazione alla ricerca in una logica più ampia di Terza missione. Senza il rafforzamento delle conoscenze e delle competenze (*capacity and institution building*) non è, infatti, possibile promuovere innovazione locale capace di garantire vera autonomia delle politiche e della crescita. Un momento cruciale per avviare questa riflessione si è avuto nel 2015 con il lancio della nuova "Agenda 2030" delle Nazioni Unite e i suoi 17 Obiettivi di Sviluppo sostenibile.²² In questo scenario, l'attività delle università non si limita più alla erogazione di formazione, ma si concentra su un processo di rafforzamento istituzionale a livello locale, nazionale e sovranazionale delle competenze e delle capacità, e su un efficace processo di *empowerment*. Pertanto, le missioni tradizionali di formazione e ricerca devono essere aggiornate e rinnovate, per offrire una nuova visione all'idea di internazionalizzazione, intesa non solo come opportunità di aprirsi alle offerte dei mercati internazionali, ma anche come risposta all'esigenza di dotare i paesi di politiche e strumenti resilienti e reagenti, e soprattutto democratici.

²² <<https://unric.org/it/agenda-2030/>>.

EL PROGRAMA DOCTORADO DE CIENCIAS DE LA DOCUMENTACIÓN DE LA UNIVERSIDAD COMPLUTENSE DE MADRID

Antonio Carpallo Bautista*

Introducción

El programa de Doctorado en Ciencias de la Documentación impartido en la Universidad Complutense de Madrid, congrega el antiguo programa de Doctorado del Departamento de Biblioteconomía y Documentación (Fundamentos, tecnologías y aplicaciones de las Tecnologías Documentales y Procedimiento de Información) y el programa de Doctorado en Ciencias de la Documentación de la Facultad de Ciencias de la Documentación, anterior al Real Decreto 99/2011, ambos de la Universidad Complutense. El programa de Doctorado se lleva a cabo junto a tres Departamentos de la Universidad Complutense como son el de Biblioteconomía y Documentación, con sede en la Facultad de Ciencias de la Documentación, Departamento de Prehistoria, Arqueología, Historia Antigua, Historia Medieval y Ciencias y Técnicas Historiográficas (antes Ciencias y Técnicas Historiográficas y de Arqueología), con sede en la Facultad de Geografía e Historia y el Departamento de Literaturas Hispánicas y Bibliografía (antes denominado Filología Española IV - Bibliografía Española y Literatura Hispanoamericana) con sede en la Facultad de Filología, aunque los profesores de este departamento adscritos a la Facultad de Ciencias de la Documentación han pasado a formar parte del Departamento de Biblioteconomía y Documentación, junto a su carga docente. Muchos de estos profesores participan en los diferentes másteres que se imparten en la Facultad (Máster en Gestión de la Documentación, Bibliotecas y Archivos; Máster en Documentación Fotográfica. Recuperación, Tratamiento y Difusión; Máster en Patrimonio Histórico Escrito; Máster de Libro Antiguo y Patrimonio Bibliográfico).

El objetivo primordial que persigue el programa de Doctorado es la continuación en la formación de los alumnos en cualquiera de las ramas de especialización de los másteres mencionados como Gestión de la

* Universidad Complutense de Madrid.

Documentación, Bibliotecas y Archivos, Documentación Fotográfica: recuperación, tratamiento y difusión y Libro Antiguo y Patrimonio Bibliográfico.

El programa de Doctorado se basa en los postulados del Espacio Europeo de Educación Superior (EEES), promoviendo en especial la colaboración con universidades iberoamericanas y europeas en la realización de actividades formativas, la colaboración con organismos e instituciones en la realización de actividades investigadoras (bibliotecas, institutos de investigación, etc.), ambas con la posibilidad de incluirlas en el documento de actividades del doctorando, y la movilidad de los alumnos de doctorado en otros centros europeos y americanos principalmente.

El presente programa es uno de los más pequeños de la Universidad Complutense siendo el 1,42% del total de alumnos, por delante de facultades como Matemáticas, Geológicas, Informática, Óptica y Optometría, Trabajo Social y Estudios Estadísticos. Oferta cada año 25 nuevas plazas de ingreso, teniendo en la actualidad un total de 60 doctorandos, de los cuales el 45% son extranjeros (mayoritariamente de Iberoamérica con países como México, Argentina, Colombia, Chile, Ecuador y Brasil); destacar que el 60% de doctorandos proceden de otras universidades, tanto extranjeras como españolas (Zaragoza, Valencia, Barcelona, Extremadura...).

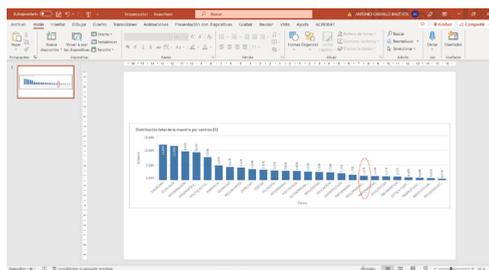


Fig. 1. Porcentaje de alumnos matriculados en el Programa de Doctorado

La gestión de todo lo relativo a este programa de doctorado se lleva desde la Comisión Académica de Doctorado formada por el Presidente-Decano, el Coordinador del Programa, un representante del Departamento de Biblioteconomía y Documentación, otro del Departamento de Literaturas Hispánicas y Bibliografía, y un representante del Departamento de Historia de América y Medieval y Ciencias Historiográficas.

La mayoría de los años hemos tenido un número de preinscripciones superior al número de plazas ofertadas. Por ejemplo, los cursos 2017/18 tuvimos 35 preinscripciones, el 2018/19 llegamos a las 27, y los últimos dos cursos, el 2021/22 llegamos a las 33 y el 2022/23 a las 30 preinscripciones. Es también significativo que la mayoría de las preinscripciones son en 1ª opción

como las 26 del curso 2017/18, y las 27 de los dos últimos cursos 2021/22 y 2022/23, lo que denota el interés por este programa de doctorado.

En cuanto al número de matriculados en todo el programa, siempre se mantienen un número alto, entre los 62 del curso 2020/21 y los 80 del curso 2017/18. Hay que indicar que este doctorado es el más solicitado y el que más doctorandos tiene dentro del campo de las Ciencias de la Documentación en España.

Algo a destacar también el porcentaje de alumnos matriculados a tiempo parcial, lo que denota que muchos de los doctorandos, compatibilizan otras labores, por ejemplo, sus puestos de trabajo con la realización del doctorado. El porcentaje ha ido aumentando en los últimos años, desde el 36,25% del curso 2017/18 hasta el 46,03% del curso 2022/23.

Un aspecto positivo para valorar es el número de doctorandos extranjeros matriculados que, aunque ha ido descendiendo en los últimos años, sigue siendo un número bastante importante. Hemos pasado desde el 45% en el curso 2017/18 hasta 34,92% en el curso 2021/22. Las procedencias más habituales son de países de Latinoamérica como México, Colombia, Uruguay, Brasil, Chile, Ecuador, Perú, entre otros, aunque también tenemos doctorandos de origen chino.

Es significativo observar el porcentaje de matriculados procedentes de otras universidades. En el curso 2017/18 tuvimos un 60% hasta llegar al 79,31% en el curso 2021/22, viéndose un aumento progresivo. Lo normal es que tengamos un porcentaje al torno de 66%, siendo la mayoría de las universidades de Latinoamérica, aunque también encontramos alumnos que proceden de universidades españolas como Zaragoza, y de la Comunidad de Madrid como la Universidad Carlos III y la Autónoma. Con la implantación de nuevos másteres en la Facultad de Ciencias de la Documentación y la consolidación de otros creados recientemente, está haciendo que alumnos de estos másteres (Máster en Documentación Fotográfica: Recuperación, Tratamiento y Difusión; Máster en Patrimonio Histórico Escrito; Máster de Libro Antiguo y Patrimonio Bibliográfico), junto con el ya consagrado desde hace varios cursos (Máster en Gestión de la Documentación, Bibliotecas y Archivos), se está trabajando para promover la participación de estos alumnos en el programa de doctorado, viéndose un aumento en los últimos años de alumnos procedentes de estas titulaciones, junto a otras también de la Universidad Complutense de Madrid, de másteres impartidos en las Facultades de Bellas Artes e Historia principalmente.

Para la realización de excelentes tesis doctorales es importante que los doctorandos se puedan dedicar a tiempo completo, por lo que es positivo que cuenten con algún tipo de beca, contrato o subvención que les ayude en su investigación. A nivel estatal podemos contar con la FPI y FPU, las becas de la Comunidad de Madrid y los contratos de Ayudante y predoctorales de la

propia Universidad Complutense de Madrid. Cada vez vamos teniendo más doctorandos con estas ayudas, incluso becas procedentes de otros países que traen doctorandos de Latinoamérica. Estos alumnos, suelen estar vinculados a un grupo de investigación, sobre todo a uno de los siete que tenemos en la Facultad, lo que les ayuda en sus investigaciones y publicaciones. Cada año solemos tener entre 2 y 4 doctorandos con algunos de estas ayudas o contratos.

Es relevante también señalar la baja tasa de abandono (23,08%), muchas veces en alumnos a tiempo parcial que no pueden compatibilizar sus investigaciones con su vida familiar y laboral. Los doctorandos, antes de tomar la decisión de abandonar el programa de doctorando, tienen varias posibilidades como la solicitud de varios años de prórrogas (2 para los alumnos a tiempo completo y 3 años para los alumnos a tiempo parcial) además de dos años de baja temporal, sin contar las bajas temporales sobrevenidas por maternidad, embarazos o enfermedades.

El número de tesis defendidas cada curso es variable, desde las cinco en el curso 2021/22 hasta las 11 en los cursos 2015/16 y 2021/21, teniendo una media alrededor de ocho cada año, la gran mayoría con la calificación “Cum Laude”.

Son muy pocas las tesis defendidas en formato “Publicaciones”, solo 2 en los últimos 5 años, aunque es una modalidad que está comenzando a llevarse a cabo, sobre todo en doctorandos con contratos o becas, a tiempo completo, que desean realizar carrera docente. Relacionado con este aspecto, podemos tener en cuenta las contribuciones científicas resultantes de la tesis (desde la fecha de primera matrícula hasta el abono de tasas para la defensa de la tesis). En los cursos 2017/18 y 2020/21 los nuevos doctores habían realizado un total de 27 publicaciones, todas ellas en revistas indexadas con buenos índices de impacto, y en capítulos de libros y monografías, bien posicionadas en el índice SPI.

Poco podemos decir del porcentaje de tesis defendidas con mención “Industrial”, ya que por ahora no se ha realizado tesis doctorales que impliquen patentes.

Para poder defender la tesis en el programa de doctorado de Ciencias de la Documentación es obligatorio tener un director y el tutor perteneciente al citado programa, además de poder tener el doctorando dos directores, hasta un máximo de tres, del mismo programa, de la UCM o de otra universidad externa. Esta posibilidad la utilizan muchos alumnos con contratos FPI y FPU que vinculan sus trabajos a grupos de investigación de otras entidades como el CSIC o instituciones extranjeras.

Actividades Formativas

Las actividades del programa de doctorado se encuentran divididas en seis tareas obligatorias durante el primer año, y todas gestionadas en el Campus

Virtual.

En la Actividad I se desarrolla la publicación y difusión de trabajos y artículos científicos, con una duración de 10 horas lectivas, y se desarrolla entre noviembre y diciembre de cada año lectivo.

En la Actividad II se imparten Técnicas avanzadas de comunicación oral y escrita en entornos académicos, con una duración de 15 horas lectivas, y se desarrolla entre diciembre y febrero de cada año lectivo.

En la Actividad III se imparten Técnicas y metodologías específicas de investigación cualitativa en Ciencias de la Documentación, con una duración de 10 horas lectivas, y se desarrolla entre marzo y abril de cada año lectivo.

En la Actividad IV se imparten Técnicas y metodología específicas de investigación cuantitativa en Ciencias de la Documentación, con una duración de 10 horas lectivas, y se desarrolla en el mes de marzo de cada año

La Actividad V se encarga de coordinar que los alumnos de primer año asistan y participen en, por los menos dos, conferencias, seminarios y congresos nacionales e internacionales, con una duración de 40 horas, y desarrollado desde octubre hasta junio de cada año lectivo.

Y la última es la Actividad VI que consiste en observar un modelo de preparación y defensa de un Trabajo de Investigación, con sus objetivos, metodología, resultados y conclusiones, con una duración de 5 horas lectiva, y desarrollado en el mes de junio.

La evaluación de todas las actividades se realiza por los coordinadores y profesores de cada actividad, pertenecientes al programa de doctorado, mediante la entrega de trabajos de cada actividad.

Ayudas a la movilidad

Cada vez son más los doctorandos que realizan una estancia, como mínimo de tres meses, en otro país para la consecución de la Mención Internacional, sobre todo en países de Latinoamérica, Portugal, Reino Unido, Italia o Francia. El porcentaje oscila entre el 37,5% los cursos 2017/18 y 2020/21 y el 12,5% el curso 2019/20, viéndose un aumento de movibilidades desde la Comisión Académica del Programa de Doctorado. Muchos de estos doctorandos ya cuentan con una beca o contrato para la realización de su investigación, pero otros obtienen las ayudas para sufragar los gastos mediante las ayudas a la movilidad del propio programa de doctorado y las ayudas de la Escuela de Doctorado de la UCM.

Uno de los aspectos a destacar por este programa es su apuesta para que los doctorandos tengan posibilidades de ayudas a la movilidad, sufragadas por la Facultad con un presupuesto específico fijo cada año, para la realización estancias de investigación, promoviendo a la obtención de la Mención Internacional (al realizar una estancia no inferior a tres meses en un país

extranjero), la internacionalización de las investigaciones, la asistencia a congresos, jornadas, cursos... y la realización de estancias cortas para la realización del proyecto de tesis doctoral como la asistencia a bibliotecas y archivos nacionales y extranjeros.

Este tipo de ayudas se llevan realizando desde el curso 2015/2016 en el que han participado más de 40 doctorandos. En el primer año se repartieron seis ayudas de movilidad que se emplearon para la asistencia con congreso internacionales como el Encuentro Internacional de conservación preventiva e intervención en Museos, Archivos y Bibliotecas celebrado en Buenos Aires (Argentina), el II Congreso Internacional sobre Libro Medieval y Moderno, la materialidad del libro celebrado en Zaragoza, el 8th QQML (Qualitative and Quantitative Methods in Libraries International Conference) celebrado en Londres, la 15ª Conferencia EAHIL (Asociación europea de bibliotecas sobre ciencias de la salud) celebrada en Sevilla, el 8º Congreso Internacional CEISAL (Consejo Europeo de Investigaciones Sociales de América Latina) celebrado en Salamanca y el European Conference on Information Literacy (ECIL2016) celebrado en Praga (República Checa). Además, los alumnos con las ayudas han podido realizar estancias breves en el Archivo Distrial de Braga (Portugal), en el Museo de la Moda de Chile, en la Biblioteca Universitaria de Salamanca y al Instituto Iberoamericano de Madrid, en la British Library de Londres, en el Archivo General Militar de Ávila y Centro Documental de la Memoria Histórica de Salamanca, y en diversas Bibliotecas Universitarias relacionadas con Medicina y Enfermería en España. Todas estas ayudas supusieron un total de 4.900 euros, el total presupuestado.

En el segundo año (2016/2017) se repartieron ocho ayudas que se emplearon para la asistencia a eventos internacionales como el Congreso Anual de la Federación Internacional de bibliotecas Musicales celebrado Riga (Letonia), el Congreso Anual de la Asociación Internacional de Archivos Sonoros celebrado en Berlín, a la I Jornada de Estudiantes de Ciencias de la Documentación celebradas en la Universidad Complutense de Madrid, el VIII Encuentro EDICIC en Coimbra (Portugal), la Jornada Ciencia en Redes celebrado en Sevilla, el Encuentro Internacional por la unidad de los educadores y el Evento Científico Bibliotecológico celebrados ambos en La Habana (Cuba), el European Conference on Information Literacy (ECIL2017) celebrado en Saint-Malo (Francia) y el 4to Coloquio Internacional sobre Líneas de Trabajo en Materia de Conservación y Restauración en Bibliotecas y Archivos organizado por el Instituto de Investigaciones Bibliográficas de la UNAM (México). También se solicitaron ayudas encaminadas a estancias breves, como apoyo a los trabajos de campo de las tesis doctorales, para consultar documentación en la National Art Library (NAL) y en la British Library de Londres (Reino Unido), en la Hispanic Society de Nueva York, Archivo y Biblioteca Capitular de la Catedral de Toledo, junto a estancias más amplias, de tres meses,

encaminadas a la obtención de la Mención Internacional, como en el Instituto Nacional de Ciberseguridad de España, en La Habana (Cuba) en la Facultad de Comunicación y en el Archivo de la República. Al igual que el primer año, se repartió la totalidad del presupuesto destinado a estas ayudas, los 4.900 euros.

El tercer año (2017/2018) se repartieron siete ayudas a la movilidad, destacando las ayudas para la asistencia a eventos como el Congreso Mobile World Congress en Barcelona, el 56º Congreso Internacional de Americanistas celebrado en Salamanca, la Conference on Typography, Illustration and Ornamentation in the Early Modern Iberian Book World, 1450-1800 en Dublín (Irlanda), el I Congreso Internacional de Investigación sobre Usuarios de la Información celebrado en Ciudad de México y el III Congreso Internacional sobre libro medieval y moderno celebrado en Zaragoza. En cuanto a estancias breves, las ayudas han permitido consultar los fondos de la Biblioteca Histórica de la Universidad de Salamanca y la Universidad de Barcelona. Como en los dos años anteriores, se repartió la totalidad del presupuesto de 4.900 euros.

En el año 2018/2019 se repartieron cuatro ayudas a la movilidad, dos de ellas para la asistencia al Congreso Internacional El patrimonio musical en la Corona de Aragón (1418-1707) celebrado en Valencia y al 11º QQML (Qualitative and Quantitative Methods in Libraries International Conference) celebrado Florencia (Italia), junto a varias estancias breves como a la British Library de Londres, Biblioteca Nacional de México y una estancia larga, de 4 meses, al Archivo Nacional d'Outre Mer (Francia). Este año se repartió la cantidad 3.900 euros.

Este curso 2019/2020 se ha solicitado otras cuatro ayudas, sobre todo para la asistencia a eventos internacionales como al Congreso Internacional de Enseñanza-Aprendizaje celebrado en Chicago (Estados Unidos), el IV Congreso Internacional sobre el libro medieval y moderno celebrado en Zaragoza y el Congreso Internacional de Arte y Literatura celebrado en Albacete, junto a estancias breves en la British Library y en la University of Harvard, repartiendo los 4.900 euros del presupuesto.

Con la pandemia del COVID, los cursos 2020/21 y 2021/22 han sido complicados para que nuestros doctorando pudieran realizar estancias de investigación y asistir a congresos nacionales e internaciones. En el curso 2020/2021 se solicitaron dos ayudas a la movilidad, una para realizar tres meses de estancia en el Music Department of Royal Holloway, University of London y otra solicitud para estudiar los fondos de la Biblioteca y Archivos del Museo Condé del Castillo de Chantilly en Francia.

En el curso 2021/22 solo se solicitó una ayuda para la asistencia a unas jornadas nacionales.

Por suerte la normalidad se está imponiendo y hemos vuelto a tener un número importante de solicitudes para ayudas a la movilidad. En el curso

2022/23 se han realizado nueve solicitudes de ayuda, algunas para la asistencia a congresos y jornadas de ámbito internacional (V Congreso Internacional sobre Libro Medieval y Moderno en Zaragoza; XVIII Jornadas Españolas de Información y Documentación en Granada; Seminario Proyecto de Conservación de Montefiascone en Roma; XXVIII edición de los Encuentros Internacionales sobre Sistemas de Información y Documentación en Zaragoza; y 17th International Conference on Metadata and Semantics Research en la University of Milano-Bicocca), además de estancias breves para la consulta de documentación en el Castillo de Chantilly (Francia), Convento de Canonas Inglesas del Santo Sepulcro de Lieja en Bélgica y en la Sorbonne Université en París.

Convenios de colaboración

Junto a la creación del programa de doctorado se firmaron una serie de convenios para favorecer el intercambio de profesores y alumnos, preferentemente de doctorado, con varias universidades punteras de Latinoamérica, como son la Universidad Autónoma de México (UNAM) y Universidad Autónoma de San Luis Potosí, ambas de México y el Instituto Brasileño de Investigación Científica y Técnica (IBICT) con sede en Brasilia (Brasil). Además, también se tienen firmado convenios con la Universidad de Zaragoza y la Universidad Carlos III de Madrid.

Grupos de investigación

Un punto fuerte de colaboración en este programa son los siete grupos de investigación que tenemos en la Facultad de Ciencias de la Documentación, todos ellos valorados como “Bueno” dentro de la Universidad Complutense, una vez evaluados por la ANECA. Abarcan todas las áreas de las Ciencias de la Documentación apoyando a los doctorando ya que los incluyen en sus actividades y proyectos de investigación, por orden alfabético encontramos al grupo BIBLIOPEGIA, grupo de investigación sobre sobre encuadernación y el libro antiguo, que se ocupa del estudio de la encuadernación, desde los materiales, técnicas de construcción, técnicas y motivos decorativos hasta su estado de conservación, análisis formal de las encuadernaciones históricas, elaboración de un lenguaje controlado sobre encuadernación, creación y gestión de un catálogo colectivo automatizado para la gestión, migración y consulta de registros de encuadernaciones artísticas (CCEA), identificación y conocimiento de las procedencias de las colecciones existentes en nuestras bibliotecas, conocimiento del valor de las colecciones y por último estudio sistemático de la formación y dispersión de las colecciones.

El segundo grupo es el denominado BISOC, Información, biblioteca y sociedad, que se ocupa de la ética de la información, compromiso social, memoria y patrimonio, alfabetización informacional y competencia digital, documentación sonora, reutilización de la información del sector público en el sector cultural, derechos de autor/dominio público y lectura.

El tercer grupo es FOTODOC, Fotografía y Documentación, viene trabajando en materia de Documentación Fotográfica desde hace varios años con el objetivo general de promover la investigación en materia de gestión y análisis de la documentación fotográfica con un enfoque multidisciplinar e integrador impulsando su difusión y uso en el actual contexto de rentabilidad económica y cultural mediante las operaciones recuperación, tratamiento documental, gestión, puesta en valor y difusión de conjuntos documentales y patrimonio fotográfico en archivos, bibliotecas, museos, y centros de documentación públicos y privados.

El cuarto grupo es el denominado IDEA LAB, Laboratorio de innovación en dirección, evaluación y administración de organizaciones y servicios de información y documentación, grupo de investigación que aborda temas como la gestión de unidades de información y documentación., la innovación y emprendimiento en el área de información y documentación, las bibliotecas inteligentes en el contexto de ciudades inteligentes y universidades inteligentes, las estrategias de internacionalización de universidades y acciones estratégicas por bibliotecas universitarias, la medición de eficiencia en bibliotecas universitarias, la gestión y planificación de proyectos de información y documentación, la Documentación en Ciencias de la Salud, la evaluación de la calidad, la estandarización de productos y servicios de información y documentación, la evaluación de competencias y nuevos perfiles profesionales en el área de Biblioteconomía y Documentación, la organización y gestión de portales de datos abiertos y el gobierno e interoperabilidad de la semántica de datos.

El siguiente grupo es denominado NUMISDOC, Numismática e Investigación Documental y Epigráfica tiene como objetivo primordial la investigación de la Historia de la Moneda desde su nacimiento hasta nuestros días, analizándola en todos sus aspectos, por supuesto en sí misma, desde su diseño teórico y fabricación hasta su desmonetización y entrada en los museos o en el mercado del coleccionismo, pero también en todo aquello que influye en la moneda y todo en lo que la moneda repercute. Asimismo, también presta atención a la investigación epigráfica, entendiéndola como un todo integral en el que es imprescindible el análisis de sus elementos externos, internos y funcionales.

El sexto grupo es el denominado Políticas de información, tecnologías de la documentación y comunicación científica (POLITECOM), que se ocupa de las políticas de Información, la implementación de legislación sobre el

tratamiento de la información, las aplicaciones de la TIC's a la Documentación, el tratamiento de la Información, los Sistemas de Gestión del Conocimiento, la gestión de contenidos eLearning, la gestión de información para empresas, la evaluación y promoción de revistas científicas y el análisis y evaluación de la actividad científica.

El séptimo y último de los grupos de la Facultad es SCRINIUM que tiene diferentes líneas de investigación como el pensamiento archivístico, el estudio de sinergias entre el patrimonio documental, las Ciencias y Técnicas Historiográficas, la Cultura escrita, la Documentación y las Humanidades digitales, sus principios (bases epistemológicas), metodologías y prácticas, la gestión archivística y documental, el uso, aplicación y adaptación de normativas y estándares documentales en fondos archivísticos, la difusión de archivos y museos, los nuevos paradigmas y plataformas de representación de los datos, la información y el conocimiento aplicados a la gestión de archivos y del patrimonio documental y museístico, el estudios del comportamiento informacional del usuario en el archivo, la investigación en comunicación científica y canales de información relativa a archivos, archivística, patrimonio documental y cultura escrita, la formación para usuarios y profesionales, el estudios teóricos y práctica de la evaluación de calidad del archivo, las fuentes de información sobre patrimonio documental, un laboratorio jurídico para archivos y un observatorio de transparencia institucional.

Complementos de Formación / Actividades Formativas

Un aspecto en el que también la Facultad se vuelca es en la organización de eventos para todos los alumnos y en especial para los doctorandos. Ya hemos comentado que la Actividad Formativa V solicita a los alumnos de primer año a participar en dos eventos, por lo que intenta promoverlos su asistencia, y si no son organizados por la Facultad, favorecer su participación mediante las ayudas a la movilidad.

Algunos de los eventos organizados en estos últimos años han sido el curso *Técnicas y destrezas para elaborar, presentar y defender una tesis doctoral en humanidades* con una duración de 20 horas. El curso consta de cinco sesiones de cuatro horas impartidas por la Profesora Gemma Muñoz-Alonso López (cuatro sesiones), y por el Profesor Luis Fernando Ramos Simón (una sesión). La quinta sesión contaba con la ayuda y colaboración de la Bibliotecaria Elena Garralón García-Quismondo, de la Biblioteca de la Facultad de Filosofía, especialista en la formación del manejo de herramientas documentales y gestores bibliográficos. Cada sesión versó sobre algunas de las herramientas que debe utilizar el alumnado de Doctorado para tener éxito en la elaboración y defensa de su Tesis Doctoral. Tras ello se abrió un espacio de tiempo para realizar ejercicios de diversa índole, ejercicios prácticos que dotarán de

sentido a las cuestiones teóricas abordadas en la primera parte de la sesión correspondiente.

Otro de los cursos organizados fue el titulado *Metodología y herramientas para la aplicación de la documentación gráfica en las publicaciones científicas* con una duración de 16 horas. El curso constó de cuatro sesiones y es impartido por los profesores Juan Miguel Sánchez Vigil, María Olivera Zaldua y Antonia Salvador Benítez. Tuvo como objetivos la puesta en valor y función de la documentación gráfica en el trabajo de investigación, conocer las herramientas para la elaboración de tablas, gráficos y figuras (fotografías, dibujos) y el aprender la aplicación en cada caso del modelo pertinente. El curso pretendió que el alumno se familiarizara con los conceptos, tipología, diseño, elaboración y uso de las ilustraciones, tablas y gráficos en las publicaciones científicas con carácter general, y en las tesis doctorales con carácter específico. El contenido del curso se dividió en seis apartados: 1) Consideraciones generales sobre la documentación gráfica (ilustraciones): estructura de las tablas, gráficos y figuras; 2) Gráficos: conceptos, diseños y herramientas; 3) Tablas: conceptos, diseños y herramientas; 4) Figuras: conceptos y tipología. Fotografías (formatos, calidad y derecho); 5) Títulos y Pies: estructura y contenido; y 6) Fuentes de Información y la Propiedad Intelectual.

Además de estos dos cursos, todos los años se organizan otros eventos como las *Jornadas FOTODOC*, realizadas por el grupo FOTODOC. No hay que olvidar que uno de los másteres de la Facultad está especializado sobre fotografía y varios de los doctorandos tienen sus líneas de investigación sobre la gestión fotográfica.

También se celebran las *Jornadas Internacionales de la Asociación Española de Bibliografía (AEB)* en la Biblioteca Nacional de España, organizadas por la propia BNE y el grupo BIBLIOPEGIA de la Facultad, en que se llevan numerosas comunicaciones de temas relacionados con el patrimonio bibliográfico y documental como la bibliografía, el libro antiguo, la encuadernación, muchas de ellas aproximaciones a tesis doctorales e investigaciones llevadas a cabo en Trabajos Fin de Máster dentro de proyectos de investigación nacionales. En nuestro programa de tenemos numerosos doctorandos realizando sus tesis sobre temas relacionados con el patrimonio bibliográfico y documental.

Un evento más fue organizado por los grupos BIBLIOPEGIA, NUMISDOC y FOTODOC titulado *El Patrimonio bibliográfico y documental. Diferentes metodologías de investigación, idénticos objetivos*, con el objetivo de acercar las diferentes metodologías en las diversas líneas de investigación de los tres grupos.

También se organizan las *Jornadas FADOC*; la *Jornada de intercambio de experiencias docentes y de investigación*, organizadas por los grupos SCRINIUM, POLITECM e Información, Biblioteca y Sociedad, los tres de la Facultad; el *Seminario Hispano-Mexicano* organizado por la Facultad, la

Universidad de Extremadura, la Universidad Autónoma de México (UNAM) y la Universidad Autónoma de San Luis Potosí; el *Seminario Hispano-Italiano en Biblioteconomía y Documentación*, organizado por la Universidad Complutense de Madrid, la Universidad Carlos III de Madrid, la Universidad de Extremadura, la Universidad de Barcelona, la Università di Roma “La Sapienza”, la Università di Urbino, la Università di Perugia, la Università di Torino, la Università di Firenze, la Università di Parma y la Università di Macerata; otro evento es la Jornada de Doctorando en Patrimonio bibliográfico.

Otros dos eventos internacionales que se organizan periódicamente son el *Congreso Internacional sobre libro medieval y moderno*, que se celebra en la Universidad de Zaragoza, en el que colabora el grupo BIBLIOPEGIA, y el *Seminario Hispano-Brasileño*, organizado por la Universidad Complutense de Madrid, la Universidad de Murcia, la Universidade de Brasília, la Univeridade Estadual Paulista y la Universidade Tiradentes.

En todos estos eventos participan, de forma muy activa, los doctorandos de nuestro programa de doctorado.

También, aunque ya organizados por la propia Universidad Complutense de Madrid, tenemos todos los años el *PhDay* y las *Tesis en 3 minutos*.

Resultados

Para finalizar queremos indicar algunos de los datos extraídos de la encuesta realizada a nuestros alumnos de doctorado. En la encuesta los alumnos manifiestan una calificación muy alta a la formación recibida desde el programa de doctorado, por ejemplo, le conceden un 8,9 sobre los criterios de admisión, un 8,6 sobre la información del perfil del profesorado, un 8,5 por tener el programa una información completa y accesible, y buenas infraestructuras y equipamiento, un 8,2 por la orientación académica recibida, por el control de las actividades formativas y por la dirección del director y tutor, un 8,1 por los complementos formativos y por los procedimientos de evaluación anual y un 7,9 por las actividades formativas ofertadas para los doctorandos. También sobre las estancias en el extranjero los doctorandos evalúan con un 9,8 las estancias realizadas fuera de España y con un 8,3 la formación recibida. En cuanto a la satisfacción, los doctorandos puntúan con un 8,5 sentirse orgullosos de ser estudiantes de la UCM, y con un 8,3 la satisfacción sobre la UCM. Por último, puntúan con un 9,1 la posibilidad de repetir estudios superiores y con un 8,5 en la UCM, también sobre la recomendación de realizar un doctorado lo puntúan con un 8,5 y en la UCM con un 8,2, igual puntuación para continuar estudiando en la UCM y repetir doctorado.

Los egresados (ya doctores) también evalúan el programa de doctorado, adjudicando un 9,8 a las actividades formativas, un 9,5, a los criterios de admisión y a los procedimientos de evaluación anual, un 9,3 al control de las

actividades formativas y al perfil del profesorado, un 9,0 a los complementos formativos, al equipamiento e infraestructuras de la Facultad dedicadas a los doctorandos y también la misma puntuación para las tareas del director y tutor y por último un 7,0 con mostrar el programa información completa y accesible. Respecto a la satisfacción, los egresados puntúan con un 8,8 sentirse orgullosos de ser estudiantes de la UCM, con un 8,3 valoran la satisfacción en la formación y con un 9,0 valoran la satisfacción sobre la UCM. Valoran con un 9,8 el repetir en la misma universidad y seguir en el doctorado, con un 9,5 evalúan la recomendación de seguir en la UCM y con un 8,8 valoran la recomendación de realizar el programa de doctorado en la UCM.

En la encuesta también los alumnos manifiestan una calificación muy alta hacia la UCM, puntuada con un 9,0, una de las más altas de todos los programas de doctorado.

De nuevo los datos de la encuesta sitúan al programa de doctorado como la titulación mejor valorada del centro, con una puntuación de un 9,0.

Finalmente observamos algunos datos que nos pueden ayudar a observar en que apartados estamos mejor valorados y en cuales no tanto. Destaca en el programa la información sobre los criterios de admisión, los directores y sus perfiles de investigación, el control y evaluación y también están muy aceptablemente valoradas las actividades formativas que se organizan.

Pensamos que es un buen programa de doctorado y seguiremos trabajando desde el Decanato y la Coordinación para que nuestros doctorandos puedan realizar excelentes trabajos de investigación.

STUDI MEDITERRANEI E POLITICHE DI INTERNAZIONALIZZAZIONE
TRA SPAGNA E SARDEGNA. LE SCIENZE DEL LIBRO
E DEL DOCUMENTO ALL'UNIVERSITÀ DI CAGLIARI

Giovanna Granata

Le molteplici iniziative condotte sul fronte delle relazioni italo-spagnole nel campo delle discipline del libro e del documento fanno dell'Università di Cagliari un caso di studio utile per avviare una riflessione di carattere più generale in tema di internazionalizzazione e, al tempo stesso, delineare alcune possibili linee di azione per lo sviluppo di future sinergie.

Si tratta di attività ed esperienze di collaborazione che, maturate nel corso degli anni, sono cresciute con notevole intensità per ragioni ben precise, derivanti dalla collocazione geografica e culturale dell'ateneo cagliaritano, inserito in un contesto regionale tradizionalmente legato da rapporti antichi e molto stretti con la Spagna. I motivi di questa vicinanza affondano infatti le loro radici nel più lontano passato della storia sarda. La presenza spagnola in Sardegna risale al 1323, quando l'infante Alfonso, figlio di Giacomo II d'Aragona, sbarcò sull'isola strappandola progressivamente al controllo pisano e genovese, forte dell'atto con cui papa Bonifacio VIII, nel 1297, aveva creato il 'Regnum Sardiniae et Corsicae' e lo aveva concesso in feudo agli aragonesi. Da allora, per secoli, la penisola iberica ha costituito il punto di riferimento politico-amministrativo e culturale per la Sardegna, soggetta al dominio catalano-aragonese e poi a quello della Corona spagnola fino a quando, nel 1718, a seguito della guerra di successione spagnola, venne assegnata ai Savoia.

Quattro secoli di dominio spagnolo hanno lasciato in Sardegna segni profondi, testimoniati a diversi livelli, non da ultimo quello linguistico. Sostanzialmente, fino alla metà del Settecento nell'isola si è parlato catalano prima e castigliano poi, oltre che sardo, ed è fiorita una letteratura sardo-ispánica che, pur nella esiguità delle attestazioni, presenta non pochi elementi di interesse. Non è un caso che, quando con il passaggio della Sardegna ai Savoia è stato imposto l'italiano nell'ambito di un più ampia operazione di carattere politico, il percorso di transizione si sia rivelato assai complesso. La presenza iberica in

* Università di Cagliari.

Sardegna è rimasta infatti un tratto rilevante che fa dell'isola un vero e proprio ponte tra la Spagna e l'Italia, non solo sul piano fisico, per la sua collocazione al centro del Mediterraneo, ma anche e soprattutto sul piano culturale.

A questa vicinanza si deve il fatto che l'ateneo di Cagliari ha sempre guardato al mondo iberico come partner ideale e quasi naturale per iniziative di collaborazione, relazioni di scambio, programmi di ricerca congiunti.

Il motore di queste iniziative, in ambito umanistico, sono state soprattutto le discipline storiche, poste di fronte all'evidente constatazione che una parte rilevante della storia sarda non poteva e non può essere studiata senza attingere a fonti documentarie spagnole o conservate in Spagna (Loddo Canepa, 1949). Non meno rilevante è stato il rinnovamento dei paradigmi storiografici che, nella scia dell'opera di Ferdinand Braudel, ha portato ad allargare lo sguardo dalla prospettiva locale a quella più latamente mediterranea (Oliva, 2021).

Soprattutto a partire dai primi anni '50 del Novecento sono state promosse grandi campagne di scavo documentario che hanno impegnato gli studiosi sardi in esplorazioni sistematiche degli archivi e delle biblioteche iberiche e nel creare sinergie con ricercatori e studiosi italiani e spagnoli su temi comuni riguardanti la storia istituzionale, politica, economica, religiosa in un'ottica di internazionalizzazione intrinsecamente connessa con gli orizzonti e le dinamiche dell'espansione iberica nel lungo periodo che va dal Medio Evo a tutta l'età moderna.

Si è trattato di ricognizioni condotte principalmente dagli storici medievali e moderni e da storici del diritto e delle istituzioni con la collaborazione di studiosi impegnati sul fronte delle discipline di ambito documentario. Sostenute finanziariamente dalla Regione Autonoma della Sardegna, tali ricerche hanno visto la partecipazione di diverse istituzioni culturali isolate, *in primis*, oltre all'ateneo di Cagliari, attraverso l'Istituto per gli studi sardi, anche la Deputazione di storia patria per la Sardegna e, a partire dalla fine degli anni '70 del Novecento, l'Istituto sui rapporti italo-iberici, costituito a Cagliari dal Consiglio nazionale delle ricerche proprio per dare consistenza e continuità ai rapporti di collaborazione scientifica maturati nel tempo e peraltro segnati da una crescente reciprocità¹.

¹ L'istituto per gli studi sardi, il cui organo scientifico era la rivista *Studi sardi*, fu fondato nel 1934 da Bacchisio Raimondo Motzo, quale centro di coordinamento e di propulsione degli studi sulla Sardegna. All'Istituto, nel 1970, è subentrata la Scuola di specializzazione in studi sardi, istituita per effetto della legge regionale 1° agosto 1970, n. 20 allo scopo di formare i quadri isolani ed oggi disattivata nel quadro della riforma universitaria delle scuole *post lauream*. La Deputazione di storia patria per la Sardegna, istituita con decreto del 20 giugno 1935, ha assorbito la precedente Società storica sarda, attiva dal 1905, facendosi carico delle sue iniziative, compresa la pubblicazione del periodico *Archivio storico sardo*, di fatto la più antica rivista scientifica isolana ancora attiva, <<https://www.deputazionestoriapatriasardegna.it/281/publicazioni>>

In effetti, essendo numerose le personalità che dalla Spagna si sono trovate a trascorrere parte della loro vita in Sardegna e a svolgerci le proprie attività sul fronte amministrativo, politico, religioso, commerciale, diversi studiosi spagnoli hanno guardato con interesse ai luoghi e agli istituti di conservazione documentari dell'isola, parte dei vasti possedimenti della Corona spagnola. Basta pensare alla pionieristica *Bibliografía española de Cerdeña* di Eduard Toda i Güell (1890) e alla sua ideale continuazione nell'opera di Joaquín Arce, la *España en Cerdeña* (1960), per la quale l'autore, insegnando spagnolo all'Università di Cagliari proprio nei primi anni '50 del Novecento, ha potuto effettuare un'ampia ricognizione delle fonti locali, nel clima di vivo interesse per il tema delle relazioni sardo-italo-ispatiche che si stava allora profilando come linea privilegiata di ricerca.

L'attenzione per le tracce lasciate dalla presenza iberica in Sardegna sul fronte linguistico e culturale ha a sua volta alimentato le ricerche condotte dagli studiosi di letteratura spagnola. Se gli storici hanno guardato soprattutto a fonti documentarie, gli ispanisti si sono rivolti principalmente a quelle bibliografiche, segnando l'avvio di un progressivo avvicinamento ai fondi librari sardi di interesse antiquario. Guidati da Giuseppina Ledda, a partire dagli anni '70 del Novecento hanno avviato una paziente e laboriosa attività di schedatura che ha valorizzato il patrimonio isolano, in gran parte sconosciuto (Ledda, 1993), e si è concretizzata in un'intensa attività di riscoperta e studio di autori e testi, talora dimenticati, oggi portata avanti dai collaboratori ed allievi della studiosa come linea ormai consolidata di ricerca².

Ne sono derivati rapporti di collaborazione, animati da relazioni di carattere personale, ma anche dalla condivisione di temi di indagine e dalla convergenza degli studi e degli interessi di ricerca che caratterizzano in maniera forte le

rivista>. L'Istituto sui rapporti italo-iberici del Consiglio nazionale delle ricerche è stato trasformato nel 2001 in Istituto di storia dell'Europa mediterranea (ISEM), riunendo anche le aree di competenza dell'Istituto di storia della tecnica di Genova e del Centro per lo studio delle letterature e delle culture delle aree emergenti (CSAE) di Torino-Milano. L'ISEM pubblica dal 2008 la rivista digitale *RiMe* (Rivista dell'Istituto di storia dell'Europa mediterranea), <<https://rime.cnr.it/>>. Le pagine delle tre riviste rendono conto delle relazioni tra studiosi sardi e spagnoli; in particolare nelle pagine dell'*Archivio storico sardo* degli anni '50 si possono leggere i resoconti relativi alle prime missioni condotte in Spagna da Bacchisio R. Motzo, Antonio Era, Francesco Loddo Canepa, Alberto Boscolo, Evaristo Putzolu, Luigi Bulferetti.

² Si vedano per esempio i lavori condotti dal gruppo ILTeC: Lingue, Testi e Culture di area Iberica, coordinato da Antonina Paba e Gabriel Andrés dell'Università di Cagliari, sulla produzione teatrale e sulle *relaciones de sucesos* nella Sardegna di età spagnola e le edizioni di autori sardo-ispatici promosse all'interno della collana 'Scrittori sardi' dal Centro di studi filologici sardi, sorto nel 2002 con il sostegno scientifico degli atenei di Cagliari e Sassari, <<https://www.filologiasarda.eu/>>.

relazioni tra l'ateneo di Cagliari e le università spagnole. Si tratta di interessi che si sono evidentemente allargati dall'ambito degli studi più strettamente storici e letterari, e che, passando attraverso i temi della produzione e circolazione libraria in età ispanica, più di recente si sono aperti anche alle discipline del libro. Vale la pena in proposito di citare almeno i lavori di due studiosi che sono stati spesso in Sardegna, Pedro M. Cátedra con i suoi studi sulla *Década de la Pasión* di Juan Coloma, pubblicata a Cagliari nel 1576 (Cátedra, 2007 e 2015) e Pedro Rueda, che, tra i volumi del bibliofilo cagliaritano Monserrat Rosselló, ha ritrovato l'unico esemplare noto del catalogo di vendita di Simone Vassalini, libraio italiano operante a Madrid alla fine del Cinquecento (Rueda, 2018).

La decisa apertura delle ricerche di ambito sardo-ispanico al campo disciplinare delle scienze del libro e del documento ha trovato concreta evidenza nel varo di un dottorato di ricerca intitolato 'Fonti scritte della civiltà mediterranea', istituito all'inizio del nuovo millennio da Luisa D'Arienzo, allieva diretta della prima generazione di ricercatori che hanno avviato la ricognizione sistematica delle fonti iberiche di interesse sardo negli anni '50 del Novecento³.

Il riferimento esplicito del dottorato al Mediterraneo ha riproposto una chiara linea di continuità con le premesse metodologiche e i paradigmi storiografici che hanno animato quella stagione di studi e, non a caso, il dottorato ha fin da subito investito in ricerche indirizzate presso istituzioni iberiche e nel sostegno a esperienze di studio in Spagna, coltivando le sinergie costruitesi nel corso dei decenni precedenti. Al tempo stesso, ha segnato un importante mutamento di prospettiva individuando con chiarezza, attraverso il richiamo esplicito al tema delle fonti scritte, il proprio asse di riferimento nella paleografia, nell'Archivistica e nelle scienze bibliografico-biblioteconomiche. Da 'discipline ausiliarie' che hanno sostenuto la ricognizione condotta da parte di storici e ispanisti presso gli istituti di conservazione di interesse 'mediterraneo', le scienze del libro e del documento sono diventate il *focus* della formazione e l'anima scientifica del progetto.

Dopo circa un decennio di attività, la riorganizzazione dei dottorati di ricerca, in scuole prima e in grandi aggregazioni di area poi, ha reso necessaria una nuova formulazione del progetto originario e il dottorato in Fonti scritte è confluito in un corso dottorale in Storia, beni culturali e studi internazionali che ha optato per una formulazione più tradizionale e, per forza di cose, più generalista sul piano delle scelte culturali⁴.

³ Il dottorato in Fonti scritte della civiltà mediterranea ha preso avvio nell'a.a. 2002-3 con il XVIII ciclo e ha concluso le attività dieci anni dopo con il XXVIII ciclo.

⁴ Il dottorato in Storia, Beni Culturali e Studi Internazionali ha preso avvio nell'a.a. 2013-14 con il XXIX ciclo ed è tuttora attivo.

Al di là di questi esiti istituzionali, tuttavia, l'esperienza del dottorato in Fonti scritte e la sua connotazione possono continuare a rappresentare un modello di riferimento sul quale costruire possibili sinergie; l'intuizione di una dimensione mediterranea, infatti, e la pregnanza del tema documentale sono in grado di offrire una cornice teorica di carattere marcatamente metadisciplinare ed internazionale nella quale sviluppare future collaborazioni che, nello scenario odierno, possono peraltro essere sostenute da una politica universitaria molto più strutturata e ricca di occasioni.

Il tema dell'internazionalizzazione, infatti, dopo le prime esperienze pionieristiche, innescate dalla convergenza degli interessi di studio di singoli ricercatori, è divenuto un ambito strategico di azione che gli atenei hanno aperto come vero e proprio fattore di sviluppo per il sistema universitario nel suo complesso, potenziando le occasioni di scambio, favorendo la creazione di reti di collaborazione e di partenariato, incentivando sinergie in grado di dare ai progetti dei singoli una cornice più solida.

Tali politiche hanno enormemente ampliato le prospettive di collaborazione aprendole ad un più vasto orizzonte sia geografico che disciplinare e allargando le aree di azione anche al fronte della didattica. In questo nuovo quadro, di notevole ricchezza ma anche di maggiore complessità, è ancora più necessaria l'individuazione di una linea di indirizzo che possa proficuamente orientare le iniziative. Il rischio è altrimenti la parcellizzazione degli interventi e la dispersione degli sforzi.

Per quanto riguarda ad esempio la mobilità studentesca, e segnatamente il programma Erasmus, l'ateneo di Cagliari ha numerose convenzioni con diversi paesi Europei. La Spagna spicca tra questi con una sessantina di accordi in grado di coprire se non la totalità, almeno una parte rilevante degli atenei iberici. Per la vicinanza culturale e la persistenza delle tracce lasciate sull'isola, la Spagna è diventata infatti da subito terra di elezione per gli studenti sardi in tutte le discipline, ma soprattutto in ambito umanistico, e altrettanto ampia è stata la risposta degli studenti spagnoli in entrata.

È questo un patrimonio di esperienze che può essere messo a valore entro una cornice più definita e finalizzata. Il fenomeno Erasmus può infatti rappresentare un osservatorio molto importante non solo come strumento operativo per favorire lo scambio di esperienze, ma anche sul piano conoscitivo e quindi come base di partenza per la costruzione di una rete strutturata di rapporti. Normalmente, a livello di ateneo, sono monitorati solo i dati globali di carattere quantitativo che, in funzione delle politiche della qualità, sono contabilizzati con riferimento al numero di studenti in mobilità e al numero dei crediti formativi (CFU) maturati per corso di laurea. Assai utile potrebbe essere invece un monitoraggio anche qualitativo, in grado di evidenziare più analiticamente gli ambiti disciplinari verso i quali gli studenti si orientano, le materie che scelgono di studiare all'estero, i programmi svolti e il contesto in

cui gli insegnamenti sono incardinati. Questo permetterebbe di conoscere in maniera più dettagliata l'organizzazione e i contenuti didattici inerenti alle discipline del libro e del documento nei rispettivi paesi e, al tempo stesso, di individuare meglio le peculiarità o le eventuali criticità, in un confronto costruttivo che possa rafforzare le collaborazioni soprattutto sul piano dell'integrazione tra didattica e ricerca.

Un altro fronte interessante di azione, che gli sforzi nel campo dell'internazionalizzazione hanno indubbiamente potenziato, è quello relativo alle *visiting fellowships*. Per quanto riguarda l'ateneo di Cagliari le iniziative in tal senso sono state incentivate in misura significativa anche grazie al contributo finanziario dalla Regione Sardegna. A partire dal 2014 sono stati emanati a cadenza regolare una serie di bandi centralizzati nei quali sono state pubblicate le posizioni disponibili presso ciascun Dipartimento per attività di insegnamento o di tipo seminariale su progetti preventivamente selezionati dai Dipartimenti stessi.

Lo sforzo in termini sia organizzativi che quantitativi è stato notevole per un ateneo di media grandezza come è quello cagliaritano. Ad esempio, il bando pubblicato per il 2022/23 ha previsto circa 70 posizioni, di cui 20 per visite *long* e 50 per visite *short*. Il medesimo modello organizzativo è stato utilizzato anche per la mobilità in entrata di *visiting scientists*, ovvero per l'ospitalità di studiosi su collaborazioni interamente dedicate allo svolgimento di progetti di ricerca congiunti.

Anche in questo caso, ed anzi a maggior ragione, visto che il criterio di selezione delle posizioni tiene necessariamente conto anche della rotazione tra le discipline, è essenziale poter disporre di una cornice di riferimento che possa valorizzare le singole iniziative e superare la dimensione dell'occasionalità. Peraltro tale cornice può a sua volta favorire un coinvolgimento più ampio di valenza inter o metadisciplinare dando luogo così ad azioni più strutturate e organiche.

Non posso non portare l'esempio del *joint research project* che ho presentato nel 2016, riprendendo le linee di ricerca degli ispanisti guidati da Giuseppina Ledda sui fondi bibliografici spagnoli, ma con un *focus* sugli aspetti che erano stati allora trascurati, ovvero sulla dimensione materiale connessa alla fisicità del libro, alla caratterizzazione degli esemplari e al rapporto tra questi e i lettori (Università di Cagliari, 2018).

L'idea base del progetto era quella di riprendere in mano il *Catalogo degli antichi fondi spagnoli della Biblioteca universitaria di Cagliari (1983-85)*, integrandolo con i dati relativi ai possessori per fare luce sulla circolazione dei libri tra Sardegna e penisola iberica durante il periodo in cui si è consolidato il controllo spagnolo sull'isola.

Il *visiting* che ha partecipato al progetto è stato Pedro M. Cátedra, il cui soggiorno a Cagliari, oltre che sinergico alla ricerca, è stato anche un fattore di

catalizzazione e di richiamo per i colleghi sardi di ispanistica con i quali è stato possibile riscoprire interessi comuni e riprendere le fila di una collaborazione fattiva.

L'occasione dell'internazionalizzazione, insomma, declinata in una linea di continuità con il *background* delle esperienze pregresse, ha funzionato non solo come strumento di apertura all'esterno, ma anche come motore per un allargamento del dialogo tra ambiti di ricerca diversi all'interno dell'ateneo. Ha operato, potremmo dire, come attivatore di rete. È anche questa, credo, una prospettiva alla quale vale la pena di guardare nell'immediato futuro.

Bibliografia

- Arce, J. (1960), *España en Cerdeña*. Madrid: CSIC.
- Biblioteca Universitaria di Cagliari (1983-85), *Catalogo degli antichi fondi spagnoli della Biblioteca universitaria di Cagliari*. Pisa: Giardini.
- Cátedra, P. M. (2007), 'Juan Coloma y su Década de la Pasión' (Cagliari, 1576), in: *Con gracia y agudeza. Studi offerti a Giuseppina Ledda*, ed. Antonina Paba. Roma: Aracne Editrice, pp. 457-488.
- Cátedra, P. M. (2015), «*Década de la Pasión*», «*Cántico de la Resurrección*» de don Juan Coloma, conde de Elda y virrey de Cerdeña, edición de Pedro M. Cátedra & Javier Burguillo, con la colaboración de Laura Mier. Salamanca: SEMYR.
- Ledda, G. (1993), 'L'ispanismo in Sardegna', in: Associazione degli ispanisti italiani, *L'apporto italiano alla tradizione degli studi ispanici, Nel ricordo di Carmelo Samonà (Napoli, 30 e 31 gennaio, 1o febbraio 1992)*. Roma: Instituto Cervantes, pp. 129-133, <https://cvc.cervantes.es/literatura/aispi/pdf/05/05_001.pdf> (ultima consultazione 25/05/2023).
- Loddo Canepa, F. (1949), 'Gli Archivi di Spagna e la storia sarda', *Studi Sardi*, vol. IX, pp.142-214.
- Oliva, A.M. (2021), *L'Istituto sui Rapporti Italo Iberici e la ricerca storica negli archivi spagnoli*, in: *A novant'anni dalla missione Egidì a Simancas: studi mediterranei*, a cura di Patrizia Spinato B. e Marcello Verga. Cagliari: CNR-ISEM, pp. 373-404.
- Rueda Ramírez, P. (2018), 'La venta de libros italianos en Madrid en tiempos de Felipe II: el catálogo de Simone Vassalini (1597)', *JLIS.it*, vol. 9, n. 2, pp. 280-293.
- Toda i Güell, E. (1890), *Bibliografía española de Cerdeña*. Madrid: Tipografía de los Huerfanos.
- Università di Cagliari (2018), *Visiting Scientists at Cagliari University 2016*. Cagliari: Unicapress.

III

UN CONFRONTO SUI PROGETTI DI DIGITALIZZAZIONE DEI FONDI STORICI:
ESPERIENZE INNOVATIVE, QUESTIONI, NORME

UNA COMPARACIÓN SOBRE PROYECTOS PARA LA DIGITALIZACIÓN
DE COLECCIONES HISTÓRICAS: EXPERIENCIAS INNOVADORAS,
PROBLEMAS, REGLAS

Coordinatori

Alberto Salarelli (Università di Parma)
Cristóbal Urbano (Universitat de Barcelona)

I PROGETTI DI DIGITALIZZAZIONE DEI FONDI STORICI: TEMI A CONFRONTO

Alberto Salarelli*, Cristóbal Urbano**

Introduzione

Il tema della digitalizzazione dei fondi storici assume oggi una notevole rilevanza nell'ambito del panorama complessivo delle biblioteche digitali. Infatti sono molti gli spunti di innovazione che emergono sul versante tecnologico (si pensi agli sviluppi del web semantico o dell'intelligenza artificiale generativa), su quello delle forme di collaborazione fra istituzioni di differente livello (nuovi modelli cooperativi ma anche il ruolo sempre più strategico degli aggregatori), su quello delle risorse umane (nuovi profili professionali ma anche partecipazione diffusa a progetti di *citizen science*). A ciò si aggiunga come, all'interno delle diverse declinazioni nazionali del piano Next Generation EU, si sia manifestata una ingente disponibilità di risorse per il patrimonio culturale, risorse che per una parte significativa (nel caso italiano, per esempio) si è deciso di impiegare in massive campagne di digitalizzazione rivolte ai fondi storici di archivi e biblioteche di livello nazionale e non.

Per questi motivi, all'interno del II Seminario italo-spagnolo di Biblioteconomia e Documentazione, abbiamo ritenuto opportuno dedicare una specifica sessione a questo tema, coinvolgendo quattro bibliotecarie per l'esposizione di altrettanti casi di studio che, a noi curatori, sono parsi particolarmente significativi della qualità (ma verrebbe da dire: della nuova visione) con cui si stanno traducendo in atto, in progetti già in grado di essere apprezzati per l'efficacia delle soluzioni adottate, molte riflessioni ampiamente discusse nella letteratura scientifica. In questo breve saggio introduttivo Cristóbal Urbano e io vogliamo riprendere per sommi capi alcuni temi chiave per poi lasciare spazio ai singoli resoconti delle colleghe intervenute, colleghe che qui cogliamo l'occasione di ringraziare per la loro cortese disponibilità¹.

* Alberto Salarelli, Università di Parma, alberto.salarelli@unipr.it.

** Cristóbal Urbano, Universitat de Barcelona, urbano@ub.edu.

¹ Tra le relazioni pubblicate in questi atti non è presente quella di Elena Sánchez (Biblioteca Nacional de España) che, tuttavia, ha partecipato a questa sessione con una relazione dal titolo *Nuevos entornos para la cooperación y la reutilización del*

Schemi di metadati e interoperabilità

Affermare che i metadati rappresentino un elemento chiave nello sviluppo di un progetto di digitalizzazione significa porre l'accento su un fatto scontato. Non è certo un caso che, nell'ormai lontano 1995, il primo articolo del primo fascicolo di quella che si sarebbe rivelata una delle riviste pionieristiche nel campo delle biblioteche digitali, e cioè *D-Lib Magazine*, espressione del nascente movimento statunitense della Digital Library Federation, fosse proprio dedicato ad essi, ai metadati (Weibel, 1995), e, in particolare, a uno schema proposto, discusso e formalizzato in quello stesso anno a Dublin, Ohio...

Com'è ovvio, a quasi trent'anni di distanza, le esigenze e le pratiche di gestione delle risorse bibliografiche hanno subito notevoli cambiamenti dovuti sia ad un aumento esponenziale degli oggetti digitali resi disponibili attraverso Internet, sia ad un incremento della complessità sistemica delle soluzioni adottate per la fruizione di questi contenuti. Per questi motivi, al giorno d'oggi, la scelta e l'adozione degli appropriati schemi di metadati nelle fasi di pianificazione di una qualsivoglia iniziativa di digitalizzazione diventano passaggi in grado di rivelare la visione più o meno strategicamente orientata dell'organizzazione responsabile del progetto.

Come si può facilmente intuire, la questione assume un'importanza capitale allorché i documenti da digitalizzare siano conservati in fondi storici e/o in collezioni speciali: in questi casi è necessario valutare non solo quegli aspetti gestionali che devono essere commisurati con le caratteristiche di fragilità o di pregio dei manufatti, ma anche - per l'appunto - prestare attenzione alla scelta di soluzioni di metadato che rendano possibile una gestione efficace del controllo di autorità, delle descrizioni formali e semantiche, dei diritti di accesso e riproduzione, delle relazioni con ambiti di dominio differenti. Il tutto declinato in una dimensione, quella della stratificazione cronologica, che caratterizza quei documenti sui quali le tracce del tempo si sono progressivamente depositate e che perciò, in virtù di questa eredità che essi portano con sé dal passato, necessitano di un livello di analiticità del tutto peculiare. Mi limiterò, in questa sede, a focalizzare alcuni punti che mi sembrano più significativi di altri rispetto alle potenzialità, per certi aspetti davvero sfidanti, che la metadato contemporanea pone di fronte ai bibliotecari che si occupano di fondi storici.

Un primo aspetto da prendere in considerazione riguarda una visione complessiva e generale del ruolo dei dati della ricerca nell'epoca contemporanea che, secondo una visione ormai consolidata, devono rispondere ai criteri

patrimonio digital: la propuesta BNElab de la Biblioteca Nacional de España.

FAIR (Findable, Accessible, Interoperable and Reusable)². I metadati, sembra quasi pleonastico ricordarlo, sono innanzitutto ‘dati’, quindi sono pienamente coinvolti in questa dinamica che mira alla massima efficienza nel reperimento e utilizzo delle risorse digitali. La questione è particolarmente sentita nel mondo delle *humanities*: infatti, premesso che «digital collections are crucial inputs to research, so the application of FAIR across the research and cultural sectors has the potential to significantly improve data sharing between researchers and heritage institutions as well» (ALLEA 2020, p. 3), e che quindi c’è un interesse primario nell’ambito della ricerca verso una metadattazione al passo con i tempi, va altresì considerato il ruolo essenziale che essa può giocare anche in un’ottica di divulgazione, migliorando la visibilità dei documenti da parte di un pubblico di non specialisti. Pensiamo, ad esempio, al compito che rivestono gli aggregatori: senza queste piattaforme che si pongono idealmente come strumenti per la scoperta, rivolti quindi a diverse categorie di utenti ma innanzitutto, vedi il caso di Europeana, agli ‘appassionati’, non potremmo fruire dei vantaggi che offrono le soluzioni *one stop* ovvero, in prima battuta, la trasversalità della ricerca in collezioni di differente natura e provenienza e, in secondo luogo, la possibilità di sfruttare *tool* innovativi sviluppati dalla comunità che ruota attorno all’aggregatore per l’allestimento di percorsi di *digital storytelling*, per il riutilizzo degli oggetti digitali ai fini didattici, per le attività di *crowdsourcing*, e via discorrendo. Ebbene, la flessibilità di utilizzo degli oggetti digitali in un aggregatore è strettamente dipendente dagli schemi di metadati utilizzati, e non è certo un caso che due piattaforme tra le più importanti a livello globale – Europeana e Digital Public Library of America – abbiano deciso di sviluppare specifici schemi tagliati sulle loro peculiarità ma, soprattutto, sulla necessità di garantire un accesso efficace a pubblici con esigenze estremamente variegat³. Ma ragionare per categorie oggi non è più sufficiente. Infatti, se nel corso della storia della metadattazione si è fatta strada la convinzione che gli schemi di metadati siano strettamente legati a determinate situazioni di contesto – sia perché, a livello di visione, punti di vista differenti possono manifestarsi in differenti domini e quindi portare a schemi che prendono in considerazione alcuni elementi e non altri, sia perché le informazioni di contesto relative alla definizione di uno schema e alle condizioni del suo utilizzo sono essenziali per ciò che concerne la

² Cfr. Wilkinson *et al.* (2016) e, più recentemente, Tasovac *et al.* (2020).

³ Ci riferiamo all’Europeana Data Model (EDM) e al DPLA Metadata Application Profile (MAP). Come scrive Jeffrey Pomerantz (2015, p. 189): «Europeana and DPLA have taken metadata schemas created for several different use cases, selected those elements that are relevant for describing the universe of cultural heritage, and built a custom data model and element set for that purpose».

preservazione documentale e il riutilizzo dei dati⁴ – e che quindi, in buona sostanza, i metadati non siano da considerare come entità statiche (Lorenzini *et al.*, 2020), allo stato attuale dei fatti questa concezione sembra tradursi in una tendenza sempre più spiccata verso servizi *user-oriented*, personalizzabili in base alle esigenze non di una categoria di utilizzatori ma a livello del singolo utente, in relazione da un lato agli interessi specifici di cui egli è portatore e, dall'altro, alle sue modalità concrete di accesso alla Rete, per esempio relativamente ai dispositivi di navigazione⁵. Se è vero che «In a scenario of digital content delivery for the cultural heritage sector – either online or onsite – to adjust what is presented to the visitor is seen as essential to accommodate different visit motivations, expectations, and needs» (Not & Petrelli, 2018), ciò vale a maggior ragione allorché le risorse digitali appartengono a un fondo storico per le ragioni a cui prima accennavamo, ragioni che si sostanziano nella necessità di dare conto di una 'durata' che il documento trattiene in sé come testimone della sua epoca e, di conseguenza, delle opportune soluzioni da adottare al fine di rendere percepibile all'utente questa dimensionalità storica. Giusto per fare un esempio, pensiamo al ruolo importantissimo che possono rivestire i metadati bibliografici contenenti informazioni sulla provenienza o sulla storia della custodia del documento: essi, come ha ampiamente dimostrato il progetto MEI⁶, consentono non solo di registrare informazioni cruciali per la corretta conservazione del documento stesso, ma ampliano i significati di cui esso è portatore, individuando una sua specifica collocazione nella trama dei commerci o delle correnti culturali del suo tempo (Svoljšak, 2022).

Un secondo punto oggi strategico in tema di metadattazione concerne l'interoperabilità delle soluzioni adottate. Come abbiamo detto, ampliare gli schemi di metadati, renderli più flessibili e personalizzabili è senz'altro importante nell'economia di una biblioteca digitale, tuttavia i benefici che possono derivare da questo approccio verso un'indicizzazione arricchita sarebbero fortemente penalizzati se, contestualmente, non si ragionasse in merito alla capacità delle soluzioni adottate di risultare efficaci anche nei confronti di un utilizzo esterno all'ambiente di creazione e mantenimento dei metadati stessi. La possibilità di esporre i propri metadati in un contesto di *framework* aperti è oggi cruciale per migliorare la visibilità dell'istituzione:

⁴ «For data to be reusable, they must be accompanied by sufficient information to communicate the contents of the dataset, the purposes or circumstances under which it was created, and the ways in which it could be reused» (ALLEA, 2020, p. 17).

⁵ Cfr. Konstantinos & Caridakis, 2022. Sotto questo aspetto, tornando a Europeana, è significativo che lo schema di metadati attualmente utilizzato sia stato ritenuto poco funzionale in termini di usabilità in un caso di studio condotto presso la Deutsche Digitale Bibliothek (DDB) e che, come miglioria, sia stato proposto un nuovo schema basato su FRBR (Tan *et al.*, 2021).

⁶ <<https://www.cerl.org/resources/mei/main>>.

«The existence of many individual digital libraries, maintained by different organizations, brings challenges to the discoverability and usage of the resources by potentially interested users» (Freire *et al.*, 2020, p. 19), di conseguenza, come accennato sopra, il ruolo degli aggregatori si rivela strategico per poter intercettare l'interesse di questa utenza potenzialmente interessata ai documenti messi a disposizione da una specifica realtà. Oltre a ciò bisogna considerare anche un'altra ricaduta del concetto di interoperabilità modernamente inteso, forse ancora più importante di quella menzionata, vale a dire la possibilità di contribuire fattivamente a una crescita dell'ecosistema globale dei beni culturali: aggregando le proprie informazioni a quelle di altre istituzioni si creano le condizioni per produrre nuove informazioni di livello superiore. La questione è esposta in termini molto chiari nel *IFLA/UNESCO Manifesto for Digital Libraries*: «Interoperability and sustainability are key to the vision of digital libraries able to communicate with each other. Digital libraries that conform to commonly agreed open standards and protocols improve world-wide knowledge dissemination and access»⁷. Da questo punto di vista, l'adozione di soluzioni che consentano ai metadati di una specifica istituzione di essere non solo funzionali ai fini del recupero da parte di un aggregatore ma anche riutilizzabili nella più ampia visione del web semantico si rivela un passaggio strategico, proprio in considerazione del valore aggiunto che i *linked data* sono in grado di apportare alle funzionalità di ricerca e, di conseguenza, alle ricadute euristiche che da tali funzionalità derivano (Davis & Heravi, 2021). Per quanto attiene al concetto di *sustainability* richiamato nel *Manifesto*, esso può assumere differenti accezioni, dagli aspetti di natura economica legati alla possibilità di garantire alle risorse tutte quelle attività necessarie al loro aggiornamento nel corso del tempo⁸ fino al calcolo dell'impronta di carbonio per stimare la quantità dei gas serra emessi in atmosfera dal funzionamento della biblioteca digitale sia sul lato server che su quello client⁹. In particolare, per ciò che attiene ai metadati, la sostenibilità si riferisce alla possibilità di garantire la loro utilizzabilità e la loro significatività nel corso del tempo, a fronte di un'obsolescenza che potrebbe verificarsi sia sul piano tecnologico che su quello semantico. Ed è per questo motivo che tale attributo si rivela strategicamente connesso all'interoperabilità: in altre parole, la scelta di schemi di metadati aperti, ampiamente diffusi nelle comunità di riferimento, supportati da una dettagliata e abbondante documentazione si pone come un passaggio fondamentale per fare in modo che le risorse possano

⁷ <<https://www.ifla.org/publications/ifla-unesco-manifesto-for-digital-libraries/>>.

⁸ Tra i primi contributi attorno a questo tema si veda Hamilton, 2004.

⁹ Cfr. su questo aspetto in particolare i lavori di Gobinda Chowdhury, tra cui spicca Chowdhury, 2014.

essere recuperate e fruite anche domani in modo adeguato.

La terza questione in tema di buone pratiche di metadattazione ha a che fare con il fattore umano. Com'è ovvio, non è sufficiente rifarsi a soluzioni standard nell'allestimento delle collezioni digitali se poi chi deve concretamente applicarle non è in grado di farlo con le dovute cautele. Nell'ambito della tutela dei fondi storici il ruolo del catalogatore, spesso non disgiunto da quello del conservatore, ha ulteriormente evidenziato l'importanza del professionista che non solo sa interpretare le regole per la corretta descrizione della risorsa documentale ma che sa applicarle in modo consapevole perché conosce approfonditamente le caratteristiche dei manufatti posti sotto la sua responsabilità, una conoscenza che si estende tanto sul piano della loro condizione fisica, quanto su quello dei significati di cui essi sono portatori. Se pensiamo all'evoluzione delle biblioteche digitali in riferimento al ruolo di chi indicizza i documenti, il passaggio più significativo a cui stiamo assistendo è quello che richiede con sempre maggiore frequenza al catalogatore professionista l'assunzione di quelle conoscenze proprie di un *data curator*: le competenze squisitamente indicali si affiancano a quelle relative all'arricchimento dei dati da fonti esterne e alle modalità per rendere i dati riutilizzabili per esempio, come si diceva, nel contesto del web semantico. L'aspetto nevralgico di questo processo è rappresentato dalle modalità di acquisto di queste competenze e cioè dal tema della formazione dei catalogatori nell'era della metadattazione, argomento che, stante la sua complessità, non possiamo affrontare in queste poche righe; vi invitiamo a riflettere, ad esempio, attorno alle casistiche estremamente diversificate che, limitandoci all'Italia, potrebbero essere prese in esame in relazione ai corsi di laurea magistrale nell'ambito dei beni culturali: quanti di essi si rivelano sensibili a questa nuova prospettiva catalografica¹⁰? Per sottolineare come la questione delle competenze del *metadata librarian* stia vivendo una fase davvero cruciale, si aggiunga che essa deve oggi tenere in considerazione due sfide inquadrabili in prospettive fra loro alquanto differenti: da un lato, infatti, è necessario prendere in considerazione come i metadati del patrimonio culturale possano essere generati anche da utenti non specialisti come nel caso degli approcci wiki o, più in generale, dei progetti di *citizen science*. Fondamentale, in questi casi, risulta la capacità del professionista di mettere in campo tutte quelle soluzioni in grado di far rendere al massimo il potenziale informativo proveniente da tale approccio collaborativo alla metadattazione (Daga et al., 2022). Questa professionalità deve confrontarsi anche con l'altra sfida: le potenzialità sempre più performanti dei programmi di intelligenza artificiale generativa stanno già ora fornendo nuove ed amplissime prospettive alla metadattazione automaticamente assegnata, settore un tempo limitato solo

¹⁰ Alcuni interessanti spunti di approfondimento su questo tema si ritrovano nel dibattito a più voci riportato in Guerrini, 2022, pp. 222-240.

ad alcune tipologie di metadati tecnici che invece ora si sta sempre più aprendo a sviluppi di estremo interesse verso l'indicizzazione formale e semantica (Mingfang *et al.*, 2023).

Las métricas de utilización: usos y usuarios en la evaluación de proyectos digitales patrimoniales

La razón de ser de las iniciativas digitales patrimoniales se fundamenta en dos objetivos principales. Por una parte, facilitar el acceso de investigadores, y en buena medida también de público en general, a colecciones importantes y singulares del patrimonio histórico; por otra, asegurar una preservación a largo plazo de dichos materiales. Siendo sin duda importante el tema de la preservación, en términos de justificación de las inversiones, en este tipo de iniciativas resulta fundamental el objetivo de conectar con una audiencia para invitarla a descubrir contenidos con una narrativa expositiva de extensión cultural, o para satisfacer sus necesidades académicas de acceso consciente a documentos concretos por razones de investigación y estudio.

Si nos centramos en esta conexión con los usuarios, y atendiendo a lo que hemos visto en la sección anterior, cabe destacar que gracias al uso de esquemas de metadatos adecuados, la indización a texto completo en muchos casos y la interoperabilidad de sistemas, se pueden crear interfaces de usuario potentes para la recuperación de los objetos digitales en el sitio web del propio proyecto, así como potenciar la descubribilidad de esas colecciones desde el exterior gracias al posicionamiento de los contenidos en buscadores web y en agregadores patrimoniales de escala nacional o internacional.

Con dichas bases técnicas se podría pensar que ya estaría todo solucionado para atraer usuarios y generar usos, pero nada más lejos de la realidad. Para que un proyecto específico pueda maximizar la conexión con una audiencia indeterminada a escala Web y para que pueda justificar el servicio prestado a públicos específicos determinados como dianas del proyecto, se necesita de una evaluación continuada basada en datos de uso y de usuarios que permita tres cosas. En primer lugar, el diseño óptimo del sitio web y la mejora continuada de la experiencia del usuario basada en evidencias de uso; en segundo lugar, el establecimiento de una estrategia de posicionamiento, marketing y de difusión que dé a conocer las colecciones y que sirva para atraer visitantes al sitio web y, si procede, fidelizarlos; por último, la elaboración de informes que ayuden a una justificación basada en evidencias del retorno social y de la creación de valor de marca del proyecto, aspectos clave para la sostenibilidad económica del mismo. En relación a este último asunto, al igual que sucede con otras instituciones y servicios culturales, el análisis del impacto social y del valor económico que generan los proyectos digitales no es tarea fácil. Ahora bien, en este caso la intangibilidad del usuario y de sus transacciones a los ojos de los

gestores y de los políticos (no hay un espacio físico visitable donde encontrar personas a las que observar y con las que comentar su experiencia) hacen que el análisis del retorno de la inversión se torne especialmente evanescente y dependiente de la certificación de unas métricas emanadas del registro de visitas al web, en ocasiones presentadas de forma fría, sin un buen análisis contextual y sin complementos cualitativos que son de gran importancia.

En cualquier caso, sean únicamente cifras frías de métricas de utilización o resultados del análisis de la retroacción cualitativa de los usuarios, los estudios de uso juegan un papel de primer orden en la gestión de colecciones y proyectos digitales, aunque plantean nuevos retos y alternativas metodológicas (Kelly, 2014) en la recogida de estadísticas de utilización respecto a lo que venía siendo habitual en relación a los recursos y servicios analógicos de las bibliotecas. De ello son testimonio los cambios que la norma ISO 2789:2022 sobre estadísticas internacionales de bibliotecas (ISO, 2022) presenta respecto de la anterior versión del año 2013, cambios que son especialmente relevantes en relación a los recursos y servicios digitales, que se sustancian entre otras cosas en la incorporación de un anexo sobre la medición de los usos digitales, en el que se indican posibles datos a recoger, así como fuentes y métodos para captarlos (Arroyo-Vázquez, 2023). Se ha de destacar el papel que en la nueva edición de la norma se otorga a los sistemas de medición basados en programas de analítica web, ya sea los tipo *client-side* que registran la secuencia de páginas de un sitio web mostradas en el navegador del usuario durante una vista (si esas páginas están etiquetadas con un código de seguimiento), o los que realizan la explotación mediante programas *server-side* de los logs del servidor (registro de páginas enviadas a un navegador que las ha solicitado), tal y como se manifiesta en un fragmento clave de la sección 7.2 “Services and use” de la norma: “Usage data for the digitized collection and other non-comercial digital resources can generally be collected from a library’s or institution’s own servers or via tools such as Google Anayltics” (ISO, 2022). De la cita anterior cabe destacar dos hechos relevantes: por una parte, no se mencionan la funcionalidades estadísticas propias de los programas de gestión de los repositorios digitales desde los que se ofrecen los contenidos, lo que evidencia sus limitaciones; por otro lado, resulta sorprendente que en una norma ISO se mencione un programa comercial concreto como Google Analytics (GA) entre otros muchos posibles programas, y que se haga como sinónimo de las soluciones *client-side*. Ello viene a demostrar la elevada prevalencia del uso de GA en los sitios web bibliotecarios, pese a las dudas que plantea su utilización en el ámbito europeo por razones legales en materia de preservación de la privacidad y la protección de datos.¹¹

¹¹ Como ejemplo de ello se puede citar el caso italiano, ya que la resolución 224 de 9/6/2022 del organismo Garante de la Protezione dei Dati Personali (Scorza,

Un análisis de urgencia visitando las principales bibliotecas digitales patrimoniales de Italia y España pone de manifiesto que casi en todos los casos se usan programas de analítica web *client-side* y que el dominio de GA es casi total, con la excepción de Matomo en un número testimonial de casos, pese a que se trata de un programa que, a criterio de la mayoría de analistas, ofrece mejores garantías desde el punto de vista de protección de la privacidad. En el caso italiano el uso de Matomo está vinculado en los casos observados de la iniciativa “Web Analytics Italia”¹² para el seguimiento estadístico de los sitios web de las administraciones públicas; este es el caso del portal *Alphabetic*¹³. Se trata de una solución especialmente robusta en cuanto a protección de la privacidad, en tanto se trata de una instalación de la versión gratuita de Matomo en servidores propios, tal y como sucede en los sitios web de la Comisión Europea y en el portal de patrimonio Europeana.

En cualquier caso, el uso de programas de analítica *client-side* en sitios web de acceso a colecciones patrimoniales es una realidad claramente establecida, que se explica más allá de la necesidad de contar con estadísticas en clave de gestión bibliotecaria tal y como se plantea en la norma ISO 2789:2022. En efecto, en último término una plataforma con contenidos patrimoniales en Internet es un sitio web por sí mismo, por lo que el análisis de su posicionamiento en buscadores, de los sitios web que enlazan con la plataforma y que aportan tráfico, de la justificación del retorno de las campañas de atracción de usuarios y del seguimiento de las visitas que recibe, son requisito fundamental para la optimización del sitio web, como lo serían para la mayoría de sitios web de cualquier otra naturaleza.

En este sentido hay que destacar que cuando hablamos de “análisis” en referencia a los datos de uso y usuarios de un sitio web, nos referimos a un concepto que implica pensar en la utilidad de la información captada para la mejora continua y el análisis del desempeño. Esto es, siendo cierto que lo que no se mide no se puede mejorar, la medición sin análisis, tampoco sirve de gran cosa. Y es que analizar implica interpretar los datos en un contexto y de acuerdo con la misión, objetivos y razón de ser del proyecto en el seno de una organización promotora del mismo. Implica que los “insights” obtenidos sobre lo que sucede en el sitio web ayuden a tomar decisiones. Analizar es algo

2022) dictaminó que Google Analytics transfería datos personales a los Estados Unidos sin la garantía exigida en el Reglamento General de Protección de Datos (RGPD) de la Unión Europea. Resoluciones similares se han dado en diversos países europeos. Si bien en España la autoridad de protección de datos no se ha pronunciado al respecto de ese tema en concreto, sí lo ha hecho con anterioridad respecto de la cesión de datos de Google a terceros.

¹² <<https://webanalytics.italia.it/>>.

¹³ <<https://alphabetic.it/>>.

más que generar informes con datos estadísticos de utilización del sitio web. Aunque los datos se presenten muy bien y con una visualización eficaz, si no se acompañan de una explicación sobre lo que sucede y de la formulación de hipótesis de por qué sucede, su utilidad será limitada. El análisis web debería responder a lo que Avinash Kaushik denomina “Analítica Web 2.0” y que define como “el análisis de datos cualitativos y cuantitativos de su sitio web y de la competencia, para impulsar una mejora continua de la experiencia online que tienen tanto los clientes habituales como los potenciales y que se traduce en unos resultados esperados (online y offline).” (Kaushik, 2010). En esta definición de Kaushik cabe destacar la importancia que da tanto a los datos cualitativos como a los cuantitativos, a su complementariedad. Un enfoque también observado en la bibliografía sobre evaluación de proyectos digitales bibliotecarios (Shaw, 2016).

Las herramientas de analítica *client-side* a las que hemos hecho referencia son las que facilitan una información más granular y completa sobre un conjunto de métricas importantes en el contexto de las colecciones digitales (Bragg et al., 2015): cuántos visitantes se reciben en un periodo de tiempo definido; de qué zona geográfica proceden, en qué lengua tienen configurado su navegador, e incluso, si así lo configuramos, hasta su perfil de edad y sexo; cómo llegan al sitio web analizado, por ejemplo si lo hacen desde un buscador, otro sitio web, una red social, un mensaje de correo-e de una campaña de promoción, o entrando directamente la URL; qué hacen y cómo navegan cuando entran y con qué intensidad, e incluso la página por la que abandonan el sitio; y finalmente, nos pueden informar si hacen lo que esperamos que hagan, esto es, en qué medida se “convierten” a lo que nosotros previamente hayamos definido como objetivos de conversión en la configuración del programa de analítica. Además todos esos datos se pueden presentar atendiendo a la evolución temporal y a la comparación con periodos equivalentes anteriores, lo que nos arrojará indicadores en forma de porcentajes de crecimiento o de disminución de las métricas.

Ahora bien, no habría que dejarse deslumbrar por el alud de datos y de gráficos que programas como Google Analytics, Matomo y otros muchos, ofrecen. Son una aproximación a la realidad (muy representativa, ciertamente), por lo que su valor está más en las cifras relativas que podemos obtener comparando periodos de tiempo que en el valor “facial” de las métricas per se. Podemos caer en la falsa ilusión de la exactitud de datos si ignoramos que en muchas ocasiones no se puede saber con claridad cómo afectan el mecanismo de registro de las visitas los hábitos de autoprotección de la privacidad de los usuarios: algunos de ellos, o muchos, quizás configuran su navegador para deshabilitar las cookies, para que las borre cada vez que cierra el navegador, o para bloquear de forma explícita herramientas de tracking de visitas como Google Analytics. La fiabilidad de los datos recogidos tampoco está a salvo del

impacto que puedan tener ciertas acciones de spam, de ataques para saturar el servidor, o de la acción de robots que emulando un navegador ejecutan una cantidad de visitas y descargas totalmente anormales.

Haciendo las salvedades oportunas de lo que significa el verbo analizar y del valor aproximado de los datos registrados por los programas de analítica, se podrían destacar algunos usos que nos parecen especialmente importantes para el tipo de sitio web al que estamos haciendo referencia:

- Desarrollo de colecciones. Tanto la mayor frecuencia de acceso a determinadas piezas, como la estadística de las palabras claves que aportan mayor interacción en el sitio como resultado de las consultas en buscadores externos, o de las realizadas en el propio sistema de búsqueda interno, son una fuente valiosa de datos en la identificación de aquello que los usuarios consideran más atractivo o útil a sus necesidades. Como dice Biswas (2016), ese tipo de información puede servir en la toma de decisiones sobre qué contenidos hay que priorizar en la digitalización, o en la elaboración de propuestas narrativas de explotación de lo contenido en los repositorios. En muchos casos la entidad que gestiona ese repositorio o biblioteca digital patrimonial puede crear micrositios web dedicados a exposiciones, narrativas, objetos educativos, cuyo análisis tiene un potencial de primer orden para evaluar el impacto del trabajo de difusión ante los públicos diana previstos (Nicholas et al., 2014).
- Evaluación del impacto de los agregadores de contenidos patrimoniales. La explotación de las url referidas (referrers) en las que los usuarios hacen clic para llegar al sitio web objeto de análisis es uno de los datos de más valor que se pueden obtener de los programas de analítica web. A título de ejemplo, puede ayudar a valorar el impacto en términos de adquisición de visitas que tiene la presencia de los contenidos de una colección digital en agregadores de patrimonio (Biswas, 2016).
- Evaluación de campañas de atracción o de fidelización de usuarios. La amplitud del universo de los contenidos digitales en la Red, sean de naturaleza web o de interacción en redes sociales, hace que sea un reto captar la atención de los usuarios y fidelizarlos. Esto es, en el marco de la economía de la atención, el marketing digital y las campañas para captar usuarios son acciones necesarias para cumplir con el objetivo de conectar con la audiencia que debe tener todo sitio web, también uno dedicado a los contenidos patrimoniales. Sin embargo esa tarea tiene costes económicos y de dedicación de personal, por lo que es imprescindible evaluar qué acciones han funcionado y en qué medida se ha producido un retorno de la inversión en forma de aumento de las visitas o de su calidad, de la fidelización de usuarios o de la diversificación geográfica o tipológica de los mismos. En ese sentido, son especialmente útiles el registro de visitas que aparecen etiquetadas con identificadores de campañas, o que como

mínimo se producen a partir del momento que se desarrollaron acciones de promoción (Szajewski, 2013).

En resumen, de todo lo comentado y del desarrollo del debate de la tercera sesión del Seminario se puede concluir la necesidad de contar con evidencias para una buena gestión del web y de la colección digital, así como para transparentar ante las partes interesadas las métricas de utilización y de procedencia de los usuarios. Sin embargo, también se puede afirmar que la analítica web, entendida únicamente como registro estadístico de visitas al sitio web, ofrece una visión limitada que ha de ser complementada con estudios directos de usuarios basados en encuestas, la observación, los test de navegación, las entrevistas y un sistema permanente de atención al usuario que invite a la retroalimentación y a las sugerencias de mejora. En cualquier caso, todas esas acciones se han de tener en cuenta “if library services are to be all about engaging and bulding relationship with the user” (Biswas, 2016). Para conectar mejor con la audiencia, ya no digamos para ampliarla a nuevos nichos de usuarios, es muy importante disponer de información sobre los destinatarios reales y potenciales de todo proyecto digital.

Riferimenti bibliografici

- ALLEA (2020), ‘Sustainable and FAIR Data Sharing in the Humanities: Recommendations of the ALLEA Working Group E-Humanities’, ALLEA-All European Academies, Berlin, DOI <<https://doi.org/10.7486/DRI.tq582c863>>.
- Arroyo-Vázquez, N. (2023), ‘Novedades en la norma ISO 2789:2022 sobre estadísticas internacionales de bibliotecas’, *Anuario ThinkEPI*, vol. 17, e17a07, DOI <<https://doi.org/10.3145/thinkepi.2023.e17a07>>.
- Biswas, P. & Marchesoni, J. (2016), ‘Analyzing Digital Collections Entrances: What Gets Used and Why It Matters’, *Information Technology and Libraries*, vol. 35, no. 4, pp. 19-34, DOI <<https://doi.org/10.6017/ital.v35i4.9446>>.
- Bragg, M. *et al.* (2015), *Best practices for Google Analytics in digital libraries*. Authored by the Digital Library Federation (DLF) Assessment Interest Group Web Analytics Working Group, DOI <<https://doi.org/10.17605/OSF.IO/CT8BS>>.
- Chowdhury, G. (2014), ‘Sustainability of digital libraries: a conceptual model and a research framework’, *International Journal on Digital Libraries*, vol. 14, pp. 181-195, DOI <<https://doi.org/10.1007/s00799-014-0116-0>>.
- Daga, E. *et al.* (2022), ‘Integrating Citizen Experiences in Cultural Heritage Archives: Requirements, State of the Art, and Challenges’, *Journal on Computing and Cultural Heritage*, vol. 14, no. 1, pp. 1-35, DOI <<https://doi.org/10.1145/3477599>>.

- Davis, E. & Heravi, B. (2021), 'Linked Data and Cultural Heritage: A Systematic Review of Participation, Collaboration, and Motivation', *Journal on Computing and Cultural Heritage*, vol. 14, no. 2, pp. 1-18, DOI <<https://doi.org/10.1145/3429458>>.
- Freire, N. *et al.* (2020), 'Cultural heritage metadata aggregation using web technologies: IIF, Sitemaps and Schema.org', *International Journal on Digital Libraries*, vol. 21, pp. 19-30, DOI <<https://doi.org/10.1007/s00799-018-0259-5>>.
- Guerrini, M. (2022), *Metadatabase*. Milano: Editrice Bibliografica.
- Hamilton, V. (2004), 'Sustainability for digital libraries', *Library Review*, vol. 53, no. 8, pp. 392-395, DOI <<https://doi.org/10.1108/00242530410556210>>.
- ISO (2022), *ISO 2789:2022(E)-Information and documentation: International library statistics*, Vernier, Geneva: International Organization for Standardization.
- Kaushik, A. (2010), *Analitica web 2.0: el arte de analizar resultados y la ciencia de centrarse en el cliente*. Barcelona: Gestión 2000.
- Kelly, E. J. (2014), 'Assessment of Digitized Library and Archives Materials: A Literature Review', *Journal of Web Librarianship*, vol. 8, no. 4, pp. 384-403, DOI <<https://doi.org/10.1080/19322909.2014.954740>>.
- Konstantinos, M. & Caridakis, G. (2022), 'Context Awareness in Cultural Heritage Applications: A Survey', *Journal on Computing and Cultural Heritage*, vol. 15, no. 2, pp. 1-31, DOI <<https://doi.org/10.1145/3480953>>.
- Lorenzini, M., Rospoche, M., & Tonelli, S. (2020), 'Proposta per una valutazione automatica della completeness dei metadati nel contesto delle biblioteche digitali', *Digitalia*, vol. 15, no. 2, pp. 159-167, <<https://doi.org/10.36181/digitalia-00023>>.
- Mingfang W. *et al.* (2023), 'Automated metadata annotation: What is and is not possible with machine learning', *Data Intelligence*, vol. 5, no. 1, pp. 122-138, DOI <https://doi.org/10.1162/dint_a_00162>.
- Nicholas, D., & Clark, D. (2014), 'Information seeking behaviour and usage on a multi-media platform: case study Europeana', in: Chen C., Larsen R. (eds), *Library and Information Sciences: Trends and Research* (p. 57-78). Berlin: Springer, <https://doi.org/10.1007/978-3-642-54812-3_6>.
- Not, E., Petrelli, D. (2018), 'Blending customisation, context-awareness and adaptivity for personalised tangible interaction in cultural heritage', *International Journal of Human-Computer Studies*, vol. 114, pp. 3-19, DOI <<https://doi.org/10.1016/j.ijhcs.2018.01.001>>.
- Pomerantz, J. (2015), *Metadata*. Cambridge, MA: MIT Press.
- Shaw, E.F. (2016), 'Making Digitization Count: Assessing the Value and Impact of Cultural Heritage Digitization', in: *Proc. IS&T Archiving 2016*, DOI <<https://doi.org/10.2352/issn.2168-3204.2016.1.0.197>>.
- Scorza, G. (2022), 'Guido Scorza su Google Analytics: "La soluzione deve essere politica, tra Stati Uniti e UE"-Intervista a Guido Scorza', *Garante per la Protezione dei Dati Personali. Stampa e comunicazione*, Doc-Web 9783595 (24/06/22), <<https://www.garanteprivacy.it/home/docweb/-/docweb-display/>>

docweb/9783595>.

- Svoljšak, S. (2022), 'Historical Collections and Library Catalogs: Provenance Metadata, Bibliographic Standards and Frameworks, and Catalog Functionalities', *Cataloging & Classification Quarterly*, vol. 60, no. 8, pp. 775-785, DOI <<https://doi.org/10.1080/01639374.2022.2124340>>.
- Szajewski, M. (2013), 'Using Google Analytics Data to Expand Discovery and Use of Digital Archival Content', *Practical Technology for Archives*, vol. 1, no. 1., <<https://hdl.handle.net/1813/76835>>.
- Tan, M.A. *et al.* (2021), 'DDB-EDM to FaBiO: The Case of the German Digital Library', in: *Proceedings of the 20th International Semantic Web Conference - Posters and Demos (ISWC2021-P&D)*, Aachen, Germany: RWTH Aachen, DOI <<https://doi.org/10.34657/8003>>.
- Tasovac, T., Chambers, S. & Tóth-Czifra, E. (2020), *Cultural Heritage Data from a Humanities Research Perspective: A DARIAH Position Paper*, <<https://hal.science/hal-02961317>>.
- Weibel, S. (1995), 'Metadata: The Foundations of Resource Description', *D-Lib Magazine*, vol. 1, no. 1, <<https://www.dlib.org/dlib/July95/07weibel.html>>.
- Wilkinson, M.D. *et al.* (2016), 'The FAIR Guiding Principles for scientific data management and stewardship', *Sci Data*, no. 3, article number 160018, DOI <<https://doi.org/10.1038/sdata.2016.18>>.

ARCHIVO DE REVISTAS CATALANAS ANTIGUAS (ARCA): UN CASO PARADIGMÁTICO ENTRE LOS PROYECTOS DE DIGITALIZACIÓN COOPERATIVA EN LOS QUE PARTICIPA LA BIBLIOTECA DE CATALUNYA

Eugènia Serra*

1. Introducción

La Biblioteca de Catalunya (BC)¹ es la biblioteca nacional de Cataluña. Creada en 1907, tiene entre sus funciones, establecidas por la ley de bibliotecas de Cataluña, identificar, recoger, conservar y difundir la producción bibliográfica de Catalunya.²

Desde el punto de vista administrativo, la BC es un organismo del Departamento de Cultura de la Generalitat de Cataluña, que cuenta con autonomía presupuestaria y de ejecución.

Actualmente, la colección de la BC está formada por más de cuatro millones setecientos mil documentos de todas las épocas, tipologías y soportes, procedentes de depósito legal, compra, donativo e intercambio. Es decir que incluye, libros, revistas, diarios, archivos personales, manuscritos, mapas, grabados, partituras, sonoros, audiovisuales, etc.

El año 2004, una vez generalizado el uso de las nuevas tecnologías en las bibliotecas, la Biblioteca de Catalunya adoptó la digitalización como medio de acceso y de preservación, y apostó por la puesta en marcha de repositorios digitales que le permitieran cumplir ambas funciones (Lamarca y Serra Aranda, 2005).

De acuerdo con esta estrategia, el 2005 inició la implementación de repositorios cooperativos para ofrecer en la red acceso abierto y gratuito al patrimonio; repositorios que, después de casi veinte años continúan operativos

* Directora de la Biblioteca de Catalunya.

¹ Biblioteca de Catalunya, <www.bnc.cat/es>.

² Ley 4/1993, de 18 de marzo, del sistema bibliotecario de Cataluña, <<https://www.boe.es/buscar/doc.php?id=BOE-A-1993-10384>>.

y creciendo, se trata de la Memoria Digital de Catalunya (MDC),³ para publicar monografías, y el Archivo de Revistas Catalanas Antiguas (ARCA),⁴ para revistas y periódicos. Entre ambos repositorios la colección digital de la BC recibe alrededor de tres millones de consultas de documentos cada año. Ambos repositorios estuvieron operativos para el público a partir de 2006.

Como complemento de los anteriores, también en 2005, la BC desarrolló un proyecto piloto para la recogida del archivo web de Catalunya, que poco después se puso en marcha con el nombre PADICAT (Patrimonio digital de Catalunya), y algunos años después, en 2008, implementó COFRE (Conservamos para el Futuro Recursos Electrónicos), el repositorio de preservación digital en el cual se conservan las copias de alta calidad de MDC y ARCA (Pérez y Serra, 2010).

2. *El Archivo de Revistas Catalanas Antiguas (ARCA)*

2.1. OBJETIVOS

La creación del repositorio ARCA se concreta, básicamente, en los siguientes objetivos:

- Posibilitar la consulta de publicaciones periódicas completas.
- Recuperar y preservar el patrimonio periodístico catalán reduciendo el uso de originales en papel.
- Incrementar los contenidos catalanes en Internet con acceso gratuito.
- Promover la cooperación entre instituciones.

Precisamente el primer objetivo “posibilitar la consulta de publicaciones periódicas completas” es el que marca el elemento diferencial de ARCA respecto a otros repositorios de carácter hemerográfico.

A menudo los repositorios de revistas y diarios, ya sean institucionales o cooperativos, buscan mostrar y dar acceso exclusivamente a la colección específica de una biblioteca, archivo, museo, etc. o a la de un conjunto de instituciones; en el caso de ARCA —ya desde el planteamiento inicial— se trataba de completar las colecciones de las publicaciones seriadas antes de su publicación, mediante la colaboración institucional, de manera que los usuarios no tuvieran que invertir tiempo ni esfuerzos en localizar los números de los que dispone cada institución y consultarlos separadamente en cada una de ellas. Este planteamiento que a priori parece muy elemental no es frecuente.

Cabe destacar que, si bien, en la mayoría de los casos, se han podido localizar todos los números de las publicaciones periódicas; en determinados

³ Memoria Digital de Catalunya (MDC), <<https://mdc1.csuc.cat>>.

⁴ Arxiu de Revistes Catalanes Antiques (ARCA), <https://arca.bnc.cat/arcabib_pro/ca/inicio/inicio.do>.

títulos, si la búsqueda ha dado como resultado que la revista o diario no se ha podido completar por qué no existían algunos números, igualmente se ha digitalizado y publicado si el grado de completión es alto (más de un 80%), valorando que su publicación podía aportar un beneficio significativo a los usuarios.

2.2. LA COLECCIÓN

ARCA contiene revistas y diarios que ya no se publican y que han sido representativas para la cultura y sociedad catalana.

A fecha de 2022, ARCA ofrece acceso a 447 publicaciones periódicas que suponen un total de 2.834.637 páginas. Los títulos se han podido completar gracias a la colaboración de 46 entidades, entre las cuales, archivos, bibliotecas, museos, fundaciones, asociaciones, ateneos, universidades, etc. principalmente de Catalunya.⁵



Fig. 1. Página principal de ARCA

El alcance temporal se extiende desde 1761, fecha de inicio de la publicación del *Diario curioso, histórico, erudito, y comercial, público, y económico*, hasta, principalmente, los años 50 del siglo XX, si bien se incluyen puntualmente algunos títulos hasta 2005.

Los títulos históricos habitualmente más consultados en las hemerotecas catalanas se encuentran disponibles en ARCA, como por ejemplo *La Veu de*

⁵ Arxiu de Revistes Catalanes Antiques (ARCA), *Colaboradores*, <https://arca.bnc.cat/arcabib_pro/es/contenido/colaboradores.do>.

Catalunya, El Diluvio, Destino, Diario de Barcelona, de manera que hoy en día la consulta se hace principalmente de la versión digital, disminuyendo así el uso de los originales en papel, y en consecuencia su deterioro.

Como ya se ha expuesto, dado que el alcance temporal llega principalmente hasta los años 50 del siglo pasado, la mayoría de las publicaciones ya han entrado en dominio público; aun así, se han publicado títulos con derechos vigentes una vez obtenida la autorización de los titulares, si el interés de la publicación así lo ha aconsejado.

Las cabeceras incluidas en el repositorio están principalmente editadas en Cataluña, pero también hay 3 editadas en Madrid, 1 en Valencia, 5 en Francia, 1 en Argentina y 1 en Estados Unidos.

El idioma mayoritario de las publicaciones es el catalán y el castellano, aunque también hay algunas en francés y una en inglés.

En el mismo portal se incluye un documento de la política de colección de ARCA⁶ que describe los criterios de selección de las publicaciones.

En cuanto a la evolución de la consulta, se constata la relación evidente existente entre el crecimiento de los contenidos y el aumento exponencial de accesos; se pueden identificar claramente los picos más importantes coincidiendo con la incorporación de títulos de uso generalizado por parte de los investigadores.

AÑO	2006	2008	2010	2012	2014	2016	2018	2020	oct-22
TÍTULOS	25	164	247	301	349	371	386	417	473
PÁGINAS EN MILES	20	154	505	1.143	1.268	1.977	2.409	2.479	2.834

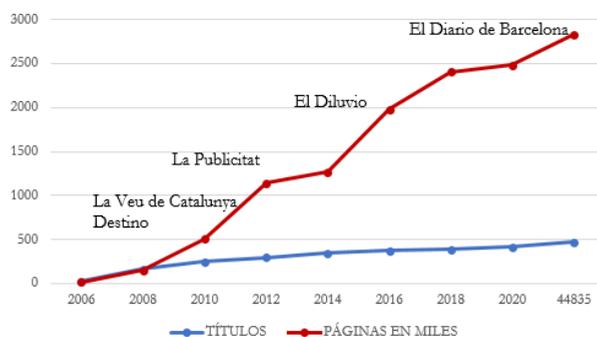


Fig. 2. Evolución del contenido de ARCA

⁶ ARCA, Política de col·lecció, <https://arca.bnc.cat/arcabib_pro/ca/cms/fichero.do?id=estaticos/ficheros/Informacio_tecnica/Politica_Colleccio_2019.pdf>.

Sobre el impacto en la investigación —más allá de los correos directos de los investigadores— sólo disponemos de un dato cuantitativo relevante, y es que hasta 2022 el número de artículos en que aparece citada ARCA en Google Scholar es de 2.300.

AÑO	2013	2014	2015	2016	2017	2018	2019	2020	2021
VISUALIZACIONES (MILES)	1.320	1.000	888	814	836	962	1.945	2.476	2.600
PÁGINAS (EN MILES)	1164	1268	1304	1977	2.330	2.409	2.456	2.479	2.803

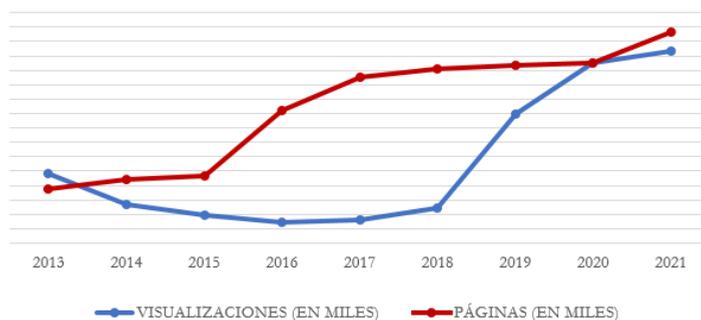


Fig. 5. Evolución de las consultas en relación con el crecimiento del contenido

2.3. COMO COLABORAMOS

Al inicio del portal ARCA identificamos una lista de 500 publicaciones esenciales que se debían incluir, lista que no ha dejado de aumentar como consecuencia del propio trabajo del equipo de la BC, pero también gracias a las sugerencias y aportaciones de las entidades colaboradoras, y a las de los investigadores que utilizan intensivamente ARCA.

La metodología que se emplea habitualmente es la siguiente, en una primera fase, el propio equipo de la BC intenta localizar los números publicados de un título concreto utilizando el catálogo propio, los catálogos en línea institucionales o colectivos, los repertorios hemerográficos y cualquier instrumento fiable de identificación y localización de información; dado que existe la certeza de que aún no se encuentra toda la información de los ejemplares accesible en línea, la segunda fase consiste en contactar con aquellas bibliotecas o entidades que de acuerdo con la información obtenida, disponen de algún número para verificar si los datos de los catálogos son completos; en tercera instancia se recurre a contactar con otras instituciones que, por su temática o especialidad

pueden aportar información sobre la localización de los ejemplares, e incluso con investigadores.



[Revue catalane \(En línia\)](#) 

Revue catalane [Recurs electrònic]

Altres títols: Revue catalane de langue,

Ejemplares: Biblioteca de Catalunya
Comprende:
no. 1 (1907:gen.15) - no. 175 (1921:set.1)

Médiathèque de Perpignan. Departament de català

Universitat de Vic

Fig. 4. Ejemplo de revista digitalizada entre varias instituciones

Una vez identificada la localización de los ejemplares, se valora el grado de completación para decidir si incluir o no la publicación en ARCA; si se decide incorporarla, se contacta de nuevo con las instituciones que poseen números y se firma un convenio de colaboración. El convenio es un trámite que permite establecer por escrito el marco de colaboración y las responsabilidades de cada una de las instituciones firmantes, y, aunque no es imprescindible, es recomendable.

Finalmente, cada organización digitaliza sus números, siguiendo unos criterios comunes que describiremos en el apartado de aspectos técnicos, y la Biblioteca de Catalunya publica la colección completa y asume la administración del portal de acceso público.

La participación en ARCA ha sido muy bien percibida a lo largo de los años por parte de las entidades colaboradoras dado que no incluye ningún tipo de exclusividad sobre las copias digitales, y en cambio maximiza la visibilidad de la actividad y fondos de las instituciones.

De manera regular, la BC recibe nuevas propuestas de participación en ARCA, gracias a la creciente reputación que ha conseguido el repositorio y portal; se puede afirmar que ARCA es un caso de éxito entre los repositorios cooperativos existentes.

2.4. ASPECTOS TÉCNICOS

Cuando se inició la definición de ARCA la biblioteca ya disponía de una serie de herramientas para trabajar todas las instituciones participantes de manera coordinada.

En lo que se refiere a los estándares de digitalización, formato, resolución,

etc., el año 2000, la BC juntamente con el CSUC (Consortio de Servicios Universitarios de Catalunya —en aquella época denominado CBUC), había elaborado un documento de pautas de digitalización a partir de las normativas internacionales⁷. Aquel primer documento fue posteriormente revisado en 2005, 2010 y 2013. Actualmente está pendiente de una nueva revisión.

En base a estos estándares la Biblioteca de Catalunya fijó los aspectos técnicos relativos tanto a la digitalización como a los metadatos que habían de ser utilizados en ARCA⁸. Aparte de los requisitos técnicos, se añadieron recomendaciones como la aplicación de OCR y el uso de los metadatos basados en el perfil *METS Profile for Historical Newspapers* definido por la Library of Congress⁹.

2.5. EL SOFTWARE Y LA INTERFAZ PÚBLICA

Cuando se implementó ARCA en 2005 y se abrió al público en 2006, el software que utilizaba era ContentDM de OCLC. En aquel momento no había demasiada variedad de programas para la gestión de repositorios digitales, y se optó por un producto comercial de uso bastante extendido, especialmente en el mundo anglosajón, y que podía gestionar tanto monografías como publicaciones en serie.

Con el paso de los años y con el incremento de la colección de ARCA se vio la necesidad de seleccionar otro software que ofreciera funcionalidades más específicas para la gestión de un repositorio hemerográfico.

A finales de 2018 la BC adquirió un software comercializado por la empresa Digibis, y el primer trimestre de 2019 el nuevo portal estaba operativo.

La actual interfaz de ARCA presenta mejoras para los usuarios tanto a nivel de búsqueda y recuperación de la información, como en la visualización y la descarga de los objetos digitales.

Desde el punto de vista de compartir datos, un hecho cada vez más habitual y relevante por los beneficios que reporta en términos de proporcionar y obtener información, e incrementar la visibilidad de las colecciones, ARCA ahora puede gestionar y exportar metadatos a diversos formatos, entre los cuales: MARCXML, Dublin Core, BibTex, JISC, Linked Open Data/EDM; soporta el protocolo OAI/PMH, de manera que su contenido puede ser

⁷ *Estandards de digitalització: requeriments mínims. Actualitzats a març de 2013 (2013)*, BC: CBUC: Col·legi Oficial de Bibliotecaris i Documentalistes de Catalunya, <<https://botiga.bnc.cat/?product=estandards-de-digitalitzacio-requeriments-minims-actualitzats-a-marc-de-2013>>.

⁸ *ARCA (2019), Requeriments tècnics*, Barcelona: Biblioteca de Catalunya, <https://arca.bnc.cat/arcabib_pro/ca/cms/fichero.do?id=estaticos/ficheros/Informacio_tecnica/Requeriments_tecnics_2019.pdf>.

⁹ <<http://www.loc.gov/standards/mets/test/ndnp/00000010.html>>.

capturado por agregadores como *Catalonica*¹⁰, el agregador de contenidos digitales de Catalunya, que a su vez provee a *Europeana*¹¹, la biblioteca digital europea.

3. A modo de conclusión

La apuesta por la digitalización y el acceso abierto es una realidad aceptada e indiscutible. Las instituciones públicas debemos invertir recursos en garantizar el derecho a la información y a la documentación, de manera que ésta puede ser además reutilizada. La creación de herramientas que lo permitan es una prioridad para la Biblioteca de Catalunya, y alcanzar estos objetivos es más sencillo si se trabaja cooperativamente, tanto a nivel conceptual como operativo.

ARCA, el Archivo de Revistas Catalanas Antiguas, nació bajo estas premisas y el resultado ha sido realmente satisfactorio; hay que continuar por el mismo camino, incrementar los contenidos disponibles, y estar alerta a los intereses de los usuarios y a las herramientas tecnológicas que vayan surgiendo y que permitan mejorar sus prestaciones.

Nos hallamos inmersos en una sociedad que evoluciona y cambia muy rápidamente y es necesario estar preparados para una adaptación constante.

Bibliografia

- España (1993), *Ley 4/1993, del sistema bibliotecario de Cataluña*. Boletín Oficial del Estado, 21 de abril, <<https://www.boe.es/buscar/doc.php?id=BOE-A-1993-10384>>.
- Figueres, J. M. (2015), 'ARCA, repositori digital de la premsa històrica catalana', *Blok de BiD*, 8 d'abril de 2015, <<https://www.ub.edu/blokdebid/es/content/arca-repositorio-digital-de-la-prensa-historica-catalana>>.
- Lamarca, D. & Serra Aranda, E. (2005), 'L'estratègia de la Biblioteca de Catalunya en projectes digitals', *Ítem*, 41, pp. 41-43, <<http://www.raco.cat/index.php/Item/article/view/40866/68116>>.
- Navarro, P. (2018), 'ARCA (Archivo de Revistas Catalanas Antiguas): un momento para el cambio', in: *Prensa histórica e xeración de coñecemento: do papel ao dixital (2018: Santiago de Compostel-la)*, <https://arca.bnc.cat/arcabib_pro/es/cms/fichero.do?id=estaticos/ficheros/ARCA.ppsx>.
- Navarro, P. (2007), 'ARCA (Arxiu de Revistes Catalanes Antigues). Projecte

¹⁰ *Catalònica. Agregador de continguts digitals de Catalunya*, <<https://catalonica.bnc.cat/catalonicahub/es/inicio/inicio.do>>.

¹¹ *Europeana*, <<https://www.europeana.eu/es>>.

- de digitalització i difusió dels fons hemerogràfics', in: *Saló del Llibre de Barcelona*, <https://www.bnc.cat/digital/arca/textos/annexos_novetats/SaloLlibreBarcelona.ppt>.
- Pérez, K. & Serra, E. (2010), *Repositori de preservació digital de la Biblioteca de Catalunya: informe descriptiu i de situació*. Barcelona: Biblioteca de Catalunya. <<http://hdl.handle.net/2072/97251>>.
- Serra, E., Molas, M. del Tura, Navarro, P. (2006), 'El Portal ARCA (Arxiu de Revistes Catalanés Antiques)', in: *Jornadas Catalanés de Información y Documentación (10as: 2006: Barcelona)*. Barcelona: Col·legi Oficial de Bibliotecaris-Documentalistes de Catalunya, pp. 283-298, <<https://cobdc.org/wp-content/uploads/serra.pdf>>.

ALPHABETICA PER LA DIDATTICA E LA RICERCA

Elisabetta Castro*

In linea dalla fine del 2021 Alfabetica, il nuovo portale progettato dall'Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche (ICCU), permette di accedere alle banche dati gestite dall'Istituto (OPAC SBN¹, MOL² ed EDIT16³) e ai progetti ad esso afferenti tramite un unico sistema di navigazione e ricerca; attraverso queste nuove modalità viene valorizzato maggiormente il patrimonio digitale delle biblioteche e al contempo anche il lavoro di cooperazione che viene svolto dalla comunità degli istituti aderenti al Servizio Bibliotecario Nazionale fin dagli anni Ottanta.

L'Istituto, nato nel 1975 a seguito della costituzione del Ministero per i beni e le attività culturali, con il compito di catalogare l'intero patrimonio bibliografico nazionale, dal 2017 si è impegnato nella progettazione di una piattaforma per un punto di accesso unico (SRI)⁴, per consentire da un lato all'utente finale una nuova modalità di navigazione più rispondente ai suoi

* ICCU- Istituto Centrale per il Catalogo Unico.

¹ L'OPAC SBN è l'interfaccia di accesso pubblica dell'Indice, basata sul motore di ricerca e indicizzazione Solr. Tale sistema è indipendente dal database dell'Indice SBN e viene alimentato attraverso scarichi periodici nel formato di scambio standard UNIMARC. L'aggiornamento della base dati dell'OPAC è settimanale. All'OPAC SBN è integrato il servizio ILL-SBN.

² MOL è il portale del censimento nazionale dei manoscritti conservati in Italia o che sono stati studiati in progetti di ricerca legati alla tradizione culturale italiana. Il censimento comprende la descrizione e le immagini digitalizzate dei manoscritti conservati nelle biblioteche italiane pubbliche, ecclesiastiche e private.

³ EDIT16 è censimento nazionale delle edizioni italiane del XVI secolo; ha lo scopo di documentare la produzione italiana a stampa del XVI secolo e di effettuare la ricognizione degli esemplari a livello nazionale. La base dati descrive edizioni stampate tra il 1501 e il 1600 in Italia, in qualsiasi lingua, e all'estero in lingua italiana. Contiene inoltre notizie di authority inerenti autori, titoli uniformi, editori e marche tipografiche consultabili autonomamente e l'archivio delle fonti bibliografiche connesse.

⁴ Per approfondire l'argomento, oltre agli articoli citati nel presente lavoro, si può consultare la documentazione relativa alla gara del progetto SRI, presente sul sito dell'ICCU: <[https://www.iccu.sbn.it/export/sites/iccu/documenti/2018/SRI CG_DLG_C_Tecnico_ICCU_2018.pdf](https://www.iccu.sbn.it/export/sites/iccu/documenti/2018/SRI	CG_DLG_C_Tecnico_ICCU_2018.pdf)>.

rinnovati bisogni informativi e dall'altro alle biblioteche aderenti di valorizzare maggiormente il proprio patrimonio.

Dai portali al portale

Dalle immagini in Figura 1 relative alla precedente versione dei portali storici dell'Istituto, emerge un'evidente disomogeneità nella struttura grafica, che non è altro che lo specchio del funzionamento delle rispettive basi dati: assenza di collegamento tra di esse e tra i rispettivi portali. Sostanzialmente, una ricerca su "Bellarmino" si sarebbe dovuta condurre tre volte, sui tre distinti portali.

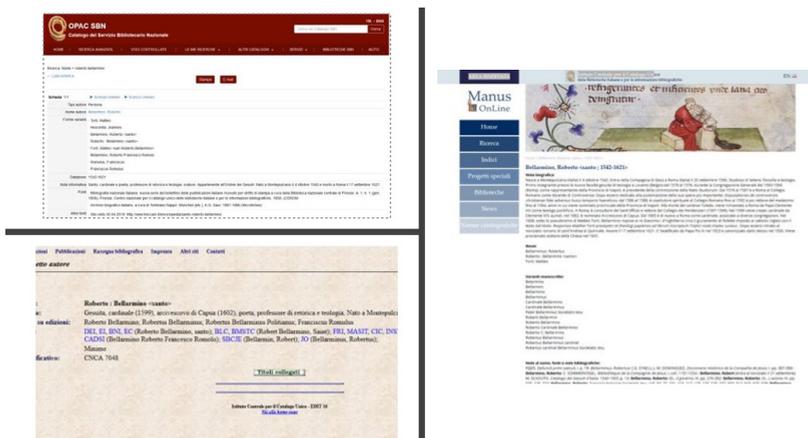


Fig. 1. L'immagine presenta il risultato di ricerca della precedente versione dei portali MOL, EDIT16 e OPAC SBN

Alphabetic, oltre a consentire un'unica interrogazione per ciascuna delle basi dati gestite dall'istituto, consente una migliore navigazione grazie al collegamento diretto agli altri canali del network. In seguito alla reingegnerizzazione, ciascun portale ha mantenuto le proprie specificità e offre agli utenti maggiori dettagli in base alla propria peculiarità; tuttavia, l'uniformità grafica ed i puntuali rimandi a portali esterni e del medesimo sistema, favoriscono l'utente nella ricerca.

Per quel che concerne la fruizione delle risorse digitali, si ricorda che questa era già possibile tramite la biblioteca digitale Internet culturale⁵, curata

⁵ Internet Culturale è la Biblioteca digitale italiana, un aggregatore di repository digitali sparsi sul territorio italiano appartenenti a biblioteche di varia provenienza e specializzazione.

dall'ICCU, che contiene le collezioni delle biblioteche aderenti e che restituisce anche le informazioni delle notizie presenti sugli altri portali dell'istituto, quelli specialistici come EDIT16 e MOL e quello più generalista, l'OPAC SBN.

Dall'osservazione della schermata di Internet Culturale si nota che la restituzione dei risultati di ricerca è 'spacchettata' per ogni singolo portale, poiché non c'è integrazione tra le basi dati.

The screenshot shows the 'Internet Culturale' search results page. At the top, there is a search bar with the text 'bellarmino' and a green 'Cerca' button. Below the search bar, there are radio buttons for 'CATALOGHI', 'BIBLIOTECA DIGITALE', 'FULL-TEXT', and 'SITO WEB'. The search results are displayed in a green header area with the text 'Risultati ricerca' and 'Pagina 1 di 140 (2.788 risultati trovati)'. On the left side, there is a section titled 'RAFFINA LA RICERCA' with a list of filters: 'Catalogo bibliografico', 'opac SBN (1843)', 'Cataloghi storici (779)', 'Biblioteca digitale (80)', 'Manus (72)', and 'edit16 (14)'. On the right side, there is a list of search results for 'Roberto (S.) Bellarmino; Bellarmino, Roberto', including 'Catalogo per autori e titoli - Biblioteca nazionale centrale di Firenze', 'Gruppi: ROBERT-ROBERT', and 'Cataloghi Storici - Unità documentaria [cs:CSTOR.139.1120895]'. The page also features a 'VEDI' button and a 'RILEVANZA: 4.494' indicator.

Fig. 2. L'immagine mostra la restituzione dei risultati di ricerca in Internet Culturale

Il progetto del punto di accesso unico nasce proprio in risposta a queste criticità, per cui nella progettazione del nuovo portale si è lavorato su un duplice piano: sul lato del *back end* si è lavorato per un'integrazione delle basi dati, mantenendo e rinviando alle schede specialistiche dei relativi portali, sul piano del *front end* si è lavorato molto sull'interfaccia grafica nell'ottica di un miglioramento della *user experience*. Per il raggiungimento di questi risultati, oltre al personale dell'Istituto e delle ditte affidatarie della commessa, si è operato in sinergia con il laboratorio BIBLAB dell'Università La Sapienza di Roma, che ha condotto dei *focus group*, i cui risultati hanno fornito notevoli spunti di riflessione.

Alphabeticamente: struttura e funzioni

Si accenna brevemente alla struttura di Alphabeticamente. Per ciascuno dei canali presenti (canale generale o tipologico tematici⁶) è possibile effettuare una

⁶ I canali da cui è possibile attivare delle ricerche mirate sono: Protagonisti, Musica, Libri, Manoscritti, Audio e Video, Cartografia, Periodici, Grafica, Biblioteche.

ricerca *google like*, mentre non sono state previste le funzionalità di ricerca avanzata e filtraggio, tipiche dell'OPAC, nell'ottica della semplificazione, dal momento che il portale è stato pensato per un utente generalista con poca confidenza con il linguaggio biblioteconomico. Anche la restituzione dei risultati di ricerca avviene all'interno di box, per fornire un valido aiuto all'utente che ritrova le risorse già suddivise in categorie. Nella nuova modalità di navigazione, un importante aiuto è dato dai *widjet*, attivabili a partire dalla restituzione dei risultati del canale generale e da quello dei 'protagonisti' oppure ancora mediante la consultazione della box "suggerimenti" presente in ogni percorso specifico.

Sono presenti dei filtri selezionabili a monte della ricerca: per ogni canale è possibile effettuare la ricerca selezionando il filtro "solo digitale"; le risorse così individuate possono essere visualizzate direttamente tramite il *viewer* del portale o tramite il link alla teca dell'Istituto che le detiene; nel canale "Protagonisti" al momento della digitazione della chiave di ricerca vengono accettati solo i nomi presenti nel DB e si possono attivare i filtri "opere di" e "opere su", mentre per il canale libri il filtro "accessibilità" è dedicato alle risorse per utenti con bisogni speciali. Anche il canale "biblioteche" accetta a monte unicamente i nomi degli istituti presenti nel DB escludendo il testo libero. Quest'ultimo canale risulta essere molto utile per la visibilità della biblioteca, perché, oltre alle informazioni sui servizi, ne presenta il patrimonio all'interno delle consuete box di raggruppamento.

Dal canale generale si può effettuare una ricerca libera o accettare il suggerimento dell'*autocomplete* a partire dai termini indicizzati raggruppati per "titolo", "argomento", "nomi".

Per comprendere a fondo la nuova prospettiva di Alfabetica ed il significato delle box che forniscono suggerimenti, simuliamo alcune ricerche, analizzando insieme la modalità di restituzione dei risultati.

Cercando "Leopardi" sul canale generale (cfr. Figura 3), i risultati ci vengono presentati all'interno di box che ricalcano i percorsi; aprendo la box dei manoscritti, si ottengono tutti i risultati relativi ai manoscritti leopardiani in sequenza, ma se si volesse focalizzare una ricerca su Leopardi nell'ambito dei manoscritti, accedendo quindi dal canale dedicato, si otterrebbero i risultati ulteriormente categorizzati ed i relativi suggerimenti (es. monografie che trattano manoscritti leopardiani), come si evince dalla Figura 4. Un ulteriore approfondimento è consentito dal sistema tramite il link diretto a MOL, nel quale si può consultare la scheda completa del manoscritto.

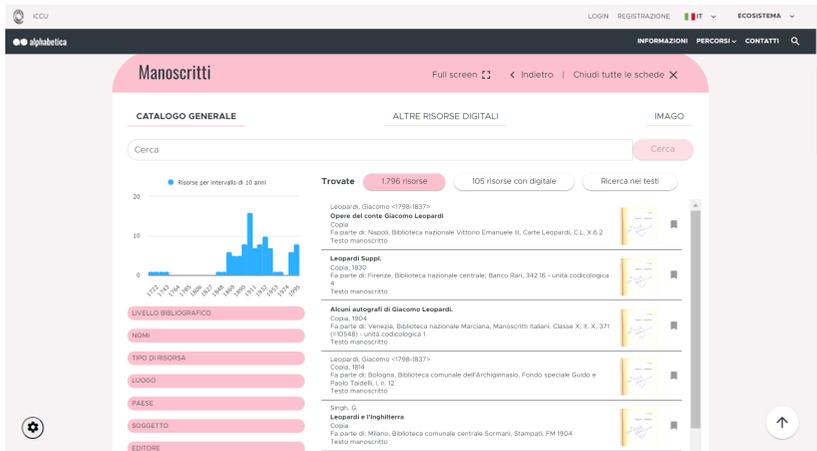


Fig. 3. L'immagine mostra i risultati di ricerca della box "manoscritti" a partire dal canale generale con la chiave di ricerca "Leopardi"

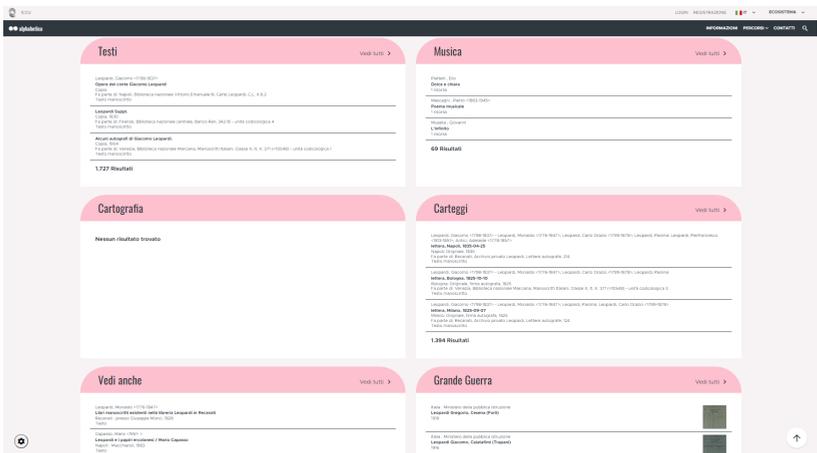


Fig. 4. L'immagine mostra i risultati delle principali box del canale manoscritti con la chiave di ricerca "Leopardi"

All'interno del canale "protagonisti", effettuiamo una ricerca su "Cervantes". La schermata mostra tutte le opere nella sezione sottostante, con i tradizionali raggruppamenti. Tra i protagonisti correlati figura Dorè, per il quale è possibile sia applicare un filtro oppure andare alla scheda del nuovo protagonista e focalizzarsi su un suo protagonista correlato, come ad esempio Dante, scegliendo infine la digitalizzazione della partitura di Donizzetti per

l'accompagnamento del XXXIII Canto della Divina Commedia di Dante.

Sostanzialmente, a partire da una ricerca su un autore spagnolo, siamo approdati ad una scheda di una partitura italiana, perché ci siamo lasciati trasportare dai legami e dalle connessioni che il sistema ci suggeriva, connessioni create non sulla base degli interessi mostrati durante la navigazione, ma legati a delle strette connessioni che rispondono a ferree regole biblioteconomiche. Su questi medesimi rigidi criteri è costruito lo strumento del *radar*, nelle due differenti declinazioni, come *focus* di “protagonisti” o come mappa concettuale.

Tornando alla nostra ricerca su Cervantes, attiviamo il *radar* come indicato nell'immagine in Figura 5.

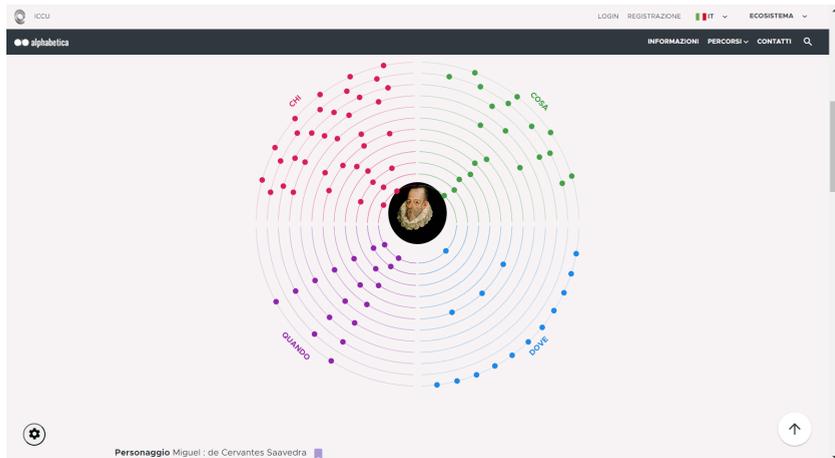


Fig. 5. L'immagine rappresenta il radar relativo all'autore Miguel Cervantes de Saavedra

Il radar presentato nell'immagine mostra graficamente la produzione del nostro autore. I raggruppamenti tengono conto del titolo dell'opera e vengono forniti i dati della manifestazione più rilevante in base all'algoritmo di Solr e dei coautori o altri nomi/personaggi legati alle opere (commentatori, curatori, disegnatori, etc.).

L'altro strumento grafico del portale è la mappa concettuale, attivabile in seguito alla ricerca dal canale generale; viene presentato di *default* il primo termine della lista rintracciato per rilevanza, successivamente l'utente, selezionando i termini più rispondenti alla propria ricerca, li visualizzerà come *focus* al centro del grafico. Lateralmente, in una sezione dedicata, vengono mostrati tutti i termini legati alla ricerca effettuata.

La mappa concettuale è basata sulla struttura del *thesaurus* del Nuovo Soggettario di Firenze, e mostra graficamente le relazioni in merito ai termini ‘correlati’, ‘specifici’, ‘generici’.

Nella barra sottostante il grafico sono presentate, suddivise nelle box tipiche del portale, le risorse legate al *focus*. Per comprendere meglio il funzionamento dello strumento, esemplifichiamo con due ricerche, a partire dalla ricerca di un nome di persona e di un oggetto.

Effettuiamo una ricerca a partire dal canale generale con “Serianni”, per rendere omaggio al linguista recentemente scomparso. Attivando la mappa concettuale, otterremo tra i primi termini elencati nella maschera a destra dello schermo, tutti i termini relativi ai suoi campi di ricerca.

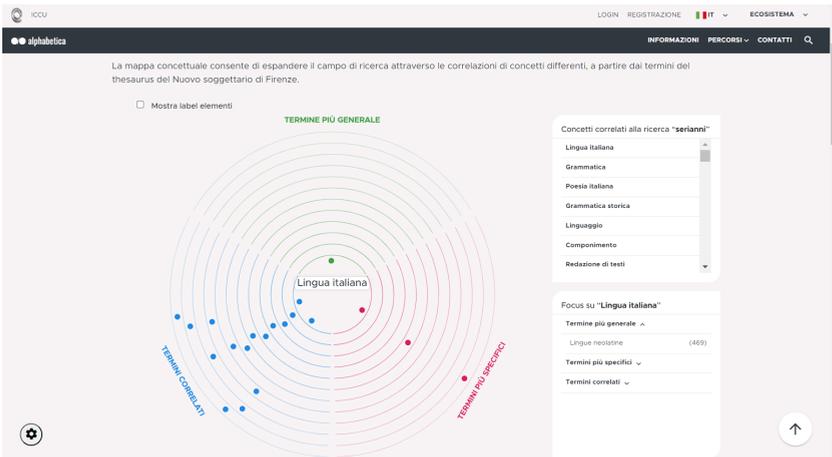


Fig. 6. L'immagine rappresenta la mappa concettuale relativa alla chiave di ricerca “Serianni”

Nello stesso canale, inseriamo la chiave di ricerca “teatro greco”.

Al centro del grafico si colloca il termine del *thesaurus* “teatro greco”, e tra i termini correlati rintracciamo “teatri greci”, termini che hanno due distinte accezioni all’interno del Nuovo Soggettario⁷. Proseguendo la nostra navigazione tra vari termini correlati e termini più specifici, da un’iniziale *focus* su una forma teatrale dell’antica Grecia, approdiamo al termine “versi degli animali”.

⁷ Una lettura attenta del *thesaurus* del Nuovo soggettario annullerebbe questa ambiguità, poiché nella nota d’ambito è chiarito l’uso e la differenza tra i termini. Si consulti il manuale della nuova edizione, disponibile anche online: Biblioteca nazionale centrale di Firenze (2021), *Nuovo soggettario: guida al sistema italiano di indicizzazione per soggetto*, <https://www.bncf.firenze.sbn.it/wp-content/uploads/2020/01/Nuovo-soggettario_Guida.pdf>.

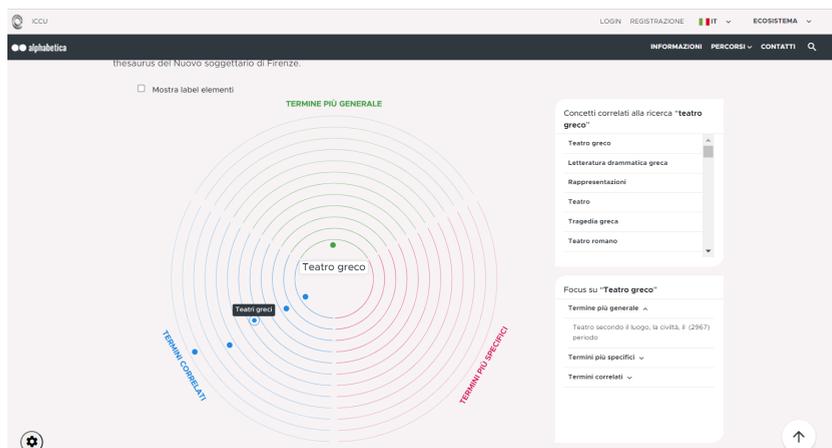


Fig. 7. L'immagine rappresenta la mappa concettuale relativa alla chiave di ricerca "teatro greco"

Gli algoritmi sottesi a questo tipo di risultato sono basati fondamentalmente su relazioni semantiche, tra opere e tra autori.

Una ricerca effettuata sui cataloghi commerciali restituisce un ampio ventaglio di scelta all'utente, che ha la percezione di poter scegliere tra tanti elementi di suo interesse, che in realtà sono stati già stati selezionati dai motori di raccomandazione grazie alla profilazione dei propri dati. Questo tipo di strategia, che certamente ottimizza un ritorno economico degli investimenti, come evidenziato in un recente studio condotto dall'Ufficio Studi Rai⁸, deve adattarsi alla *mission* del servizio pubblico.

Alphabetic, servizio di informazione e fruizione bibliografica, deve porsi come strumento neutrale e indipendente, per questo motivo non utilizza motori di raccomandazione presentando a tutti gli utenti i medesimi risultati, proponendo suggerimenti basati su meri criteri biblioteconomici, anche mediante l'utilizzo dei *widget* sopra descritti.

Attualmente in Indice sono soggettate circa il 20 % delle risorse e hanno un legame alla classificazione Dewey poco meno dell'15%. Questi dati, che grazie al lavoro recentemente svolto dall'istituto sono in fase di crescita, non consentono ancora una copertura esaustiva per tutto il patrimonio bibliografico; un algoritmo che si basa sui dati semantici, costruito con saldi criteri scientifici necessita in ogni caso di lavorare su dei dati che rispondano alle *query* effettuate, che a volte potrebbero deluderci in termini di quantità, ma raramente in termini di qualità. Per usufruire a pieno di questi servizi, la comunità SBN deve continuare a lavorare incessantemente seguendo le norme

⁸ Si tratta di Ufficio Studi Rai (2022).

di catalogazione, continuamente integrate da nuova documentazione, come le recenti norme relative alla semantica pubblicate sulla pagina *wiki*, ovvero le Linee guida sull'indicizzazione per soggetto e la classificazione nel Servizio bibliotecario nazionale (SBN)⁹.

La mappa concettuale, e gli altri strumenti simili, si basano sui dati presenti in Indice, per cui la catalogazione del patrimonio bibliografico da parte di bibliotecari preparati e scrupolosi ridurrà il margine di errore, in presenza del quale l'algoritmo non potrà discriminare due termini che all'interno del Nuovo Soggettario assumono valori differenti, come nel caso citato di "teatro greco" e "teatri greci".

Un ulteriore incremento delle risorse digitali in Alphabetic potrebbe apportare anche un ampliamento delle risorse da presentare all'interno della mappa concettuale, grazie alla diffusione di progetti di indicizzazione automatica mediante l'impiego del *thesaurus* del Nuovo Soggettario di Firenze. Le metodologie impiegate per automatizzare parzialmente o totalmente il processo di indicizzazione si basano su tecnologie di intelligenza artificiale e *machine learning*, estraendo automaticamente parole/frasi chiave e procedendo a rilevazioni statistiche delle occorrenze di un termine all'interno di documenti testuali. In questo modo, tramite Alphabetic si amplia lo spettro di ricerca, aprendo nuovi orizzonti anche mediante l'integrazione di altri portali, con collegamenti a siti web interni o esterni.

Area utente

Nel network dei servizi bibliografici nazionali è presente l'area utente, accessibile tramite la *login* a partire dai singoli portali del *network*. All'interno del proprio spazio personale, è possibile salvare informazioni relative a singole risorse o l'intera ricerca effettuata¹⁰ ed inserire questi elementi all'interno di collezioni. In fase di creazione della raccolta o della memorizzazione di singole risorse possiamo aggiungere notazioni, ad esempio, dopo aver effettuato la ricerca di un manoscritto, possiamo annotare negli appunti il *manifest* di un manoscritto della Biblioteca Apostolica Vaticana che vogliamo confrontare con quest'ultimo. Successivamente, richiamando la risorsa, possiamo visualizzarla all'interno del *viewer* integrato e aggiungere uno o più elementi provenienti da una teca esterna per il confronto, grazie all'uso del protocollo III F¹¹. Le

⁹ Le linee guida sulla semantica sono accessibili al seguente link: <https://norme.iccu.sbn.it/index.php?title=Norme_comuni/Linee_guida_sull%27indicizzazione/Indicizzazione_per_soggetto>.

¹⁰ La memorizzazione delle risorse era già possibile effettuarla con la precedente versione dell'OPAC SBN, ma era una funzionalità legata al browser e alla sessione.

¹¹ Il IIIF è un protocollo interoperabile per la condivisione delle immagini,

raccolte così create possono essere condivise tra gruppi di lavoro.

Alphabetica è un sistema in continua evoluzione, pronta a recepire le sollecitazioni esterne ed i suggerimenti degli utenti. Recentemente sono stati implementati la possibilità di riordinare le risorse salvate nelle collezioni in base ad un proprio ordine, l'importazione e l'esportazione delle collezioni in pdf. Dall'area utente è inoltre possibile attivare le notifiche per avere gli aggiornamenti riguardanti notizie contenenti determinati soggetti, classi Dewey, autori e luoghi.

Riferimenti bibliografici

- Biblioteca nazionale centrale di Firenze (2021), *Nuovo soggettario: guida al sistema italiano di indicizzazione per soggetto*. Roma: Associazione italiana biblioteche, Firenze: Biblioteca nazionale centrale di Firenze, <https://www.bncf.firenze.sbn.it/wp-content/uploads/2020/01/Nuovo-soggettario_Guida.pdf>.
- Buttò S. (2020), 'Alphabetica, il nuovo portale per la ricerca integrata: un salto di qualità per le biblioteche italiane', *Digitalia*, vol. 15, no. 2, pp. 9-15, DOI <<https://doi.org/10.36181/digitalia-00010>>.
- Buttò S. (2022), 'Aperto, flessibile, integrato: tre parole chiave per l'Ecosistema dei servizi bibliografici nazionali', *Digitalia*, vol. 17, no. 1, pp. 11-17, <<https://digitalia.cultura.gov.it/article/view/2951>>.
- Caminito M. (2018), *La biblioteca nella nuvola: utenti e servizi al tempo degli smartphone*. Milano: Editrice Bibliografica.
- Castro, E. (2022), 'Alphabetica e il reference digitale: suggerimenti e mappe concettuali per gli utenti del portale', *AIB studi*, vol. 62 no. 3, pp. 351-362, DOI <<https://doi.org/10.2426/aibstudi-12055>>.
- Castro, E. (2022), 'Uno strumento poliedrico di accesso ai servizi bibliografici nazionali', *Digitalia*, vol. 17, no. 1, p. 39-47, <<https://digitalia.cultura.gov.it/article/view/2955>>.
- Censis (2015), *I media dopo la pandemia: diciassettesimo rapporto sulla comunicazione*. Milano: Franco Angeli.
- Guarasci R., Guerrini M. (2022), *Cos'è l'indicizzazione*. Milano: Editrice Bibliografica.
- Solimine G. (2005), *La biblioteca: scenari, culture, pratiche di servizio*. Roma:

open source con diversi livelli di API per poter costruire e implementare diversi servizi. Un'immagine pubblicata con le caratteristiche di IIIF lo rendono perfetto per la pubblicazione di immagini in alta qualità sul *web*. Con IIIF un'immagine diventa un oggetto digitale portatile. Una delle funzionalità più utilizzate di IIIF è la possibilità di visualizzare, in uno stesso *viewer*. A titolo esemplificativo, si citano le principali biblioteche digitali che utilizzano il IIIF sono: LOC, Biblioteca Apostolica Vaticana, Bayerische Staatsbibliothek, Gallica, Digital Bodleian, Estense Digital library.

- Laterza, 2005.
- Svenonius E. (2008), *Il fondamento intellettuale dell'organizzazione dell'informazione*, traduzione di Maria Letizia Fabbrini, introduzione di Mauro Guerrini. Firenze: Le lettere.
- Ufficio Studi Rai (2022), *Algoritmi di servizio pubblico: sistemi di raccomandazione ed engagement per le nuove piattaforme multimediali pubbliche*. Roma: Rai Libri.

EDL: LA PIATTAFORMA DIGITALE DELLA BIBLIOTECA ESTENSE UNIVERSITARIA DI MODENA

Grazia Maria De Rubeis*

Il 29 giugno 2020 è stata *rilasciata* l'Estense Digital Library, la piattaforma digitale della Biblioteca Estense Universitaria¹.

Prima di illustrare il progetto ritengo utile richiamare il *Piano nazionale di digitalizzazione del patrimonio culturale*² e il *Nuovo Manifesto per le biblioteche digitali*³ dell'AIB, per sottolineare come EDL ne condivida, anzi ne anticipi, principi e obiettivi. Nel PND si legge:

«Il digitale offre inedite possibilità per la disseminazione culturale: le piattaforme di accesso al patrimonio propongono efficaci strumenti per l'accesso ai dati dei beni culturali, la condivisione dei risultati della ricerca, l'impiego per finalità didattiche e formative, la divulgazione scientifica, la partecipazione alla creazione di contenuti digitali con nuovi linguaggi: tutto

* Grazia Maria De Rubeis - Biblioteca Estense Universitaria, Modena. graziamaria.derubeis@cultura.gov.it.

¹ Il contributo riprende fedelmente le riflessioni condivise in occasione del II Seminario italo-spagnolo di Biblioteconomia e Documentazione, ma registra i cambiamenti occorsi e le migliori apportate dopo la conferenza, nel corso del 2023, alla Estense Digital Library: EDL, infatti, «si mette in discussione e si aggiorna» (cfr. *infra* AIB 2020).

² Il *Piano nazionale di digitalizzazione del patrimonio culturale* (PND) è stato redatto dall'Istituto centrale per la digitalizzazione del patrimonio culturale - Digital Library del Ministero della cultura. Accompagnato da 5 linee guida, è stato sottoposto a consultazione pubblica nel giugno 2022 e pubblicato nella versione 1.1 in aprile 2023. «Frutto di un processo di condivisione e confronto con diverse istituzioni culturali, il Piano costituisce la visione strategica con la quale il Ministero intende promuovere e organizzare il processo di trasformazione digitale nel quinquennio 2022-2026, rivolgendosi in prima istanza ai musei, agli archivi, alle biblioteche, agli istituti centrali e ai luoghi della cultura statali che possiedono, tutelano, gestiscono e valorizzano beni culturali». Cfr. MiC, 2023; i documenti si leggono anche su docs.italia.it e su github.com, sul sito di partecipa.gov.it sono pubblicati gli esiti della consultazione pubblica, con il Report finale e il dossier che raccoglie i contributi aperti ricevuti.

³ AIB 2020.

ciò allarga l'impatto generato dalla fruizione culturale, che non è più solamente passiva»;

e ancora

«il digitale offre la possibilità di operare un cambio di prospettiva: chi interagisce con il patrimonio non è più un fruitore passivo, ma un soggetto attivo, capace di arricchire l'universo informativo e di co-creare in modo partecipato, mettendosi in rapporto con gli altri. Non è più sufficiente coinvolgere gli utenti nella fruizione del patrimonio culturale, ma è necessario "attivarli" come agenti consapevoli dei processi di produzione culturale».

Vedremo subito che è esattamente quello che EDL si propone di fare ed esamineremo le iniziative messe in atto per raggiungere un tale risultato, perché EDL vuole essere non più solo un centro della memoria, ma anche un centro di creatività.

«EDL non è solo luogo di pubblicazione delle collezioni fisiche digitalizzate, ma anche una piattaforma per la condivisione e l'interoperabilità dei dati grazie all'utilizzo di formati standard aperti. Il sistema consente l'aggregazione e il riuso dei metadati. Questo vuol dire che su EDL non sono condivise le sole immagini, ma che i metadati stessi possono essere oggetto di aggregazione da parte di soggetti terzi. I nostri dati (come quelli di molte altre biblioteche del mondo) sono una miniera di informazioni utili per creare altro sapere, altre informazioni»⁴.

Senza elencare tutti i principi, modelli e funzioni enunciati dal *Nuovo Manifesto per le biblioteche digitali*, esaminiamo velocemente i seguenti, che, senza bisogno di commento, si applicano con immediatezza all'Estense Digital Library, come l'illustrazione della piattaforma dimostrerà:

1. Le biblioteche digitali forniscono dati e servizi
2. Gli utenti delle biblioteche digitali non sono solo gli esseri umani ma anche agenti software quali programmi, sistemi e applicazioni
3. Le biblioteche digitali promuovono la conoscenza
4. Le biblioteche digitali integrano le comunità
5. Le biblioteche digitali non si piegano a un'unica finalità, sia essa culturale o economica, ma allestiscono l'ambiente in cui qualunque finalità potrà liberamente venire perseguita
6. Le biblioteche digitali sono accessibili e si fanno conoscere
7. Le biblioteche digitali si mettono in discussione e si aggiornano
8. Le biblioteche digitali hanno modelli flessibili

⁴ Dalle pagine di presentazione dell'Estense Digital Library, cfr. <<https://edl.cultura.gov.it/pagine/pagina?id=610>>.

9. Le biblioteche digitali sono definite da dati, contenuti e servizi
10. Le biblioteche digitali contemplan e si fanno carico dell'intero ciclo di vita del digitale
11. Le biblioteche digitali adottano standard per la gestione dei dati, sei servizi e dei sistemi implementati
12. Le biblioteche digitali promuovono il riuso dei contenuti in molteplici contesti, mediante protocolli uniformi e con modalità di accesso configurabili.

I tempi, le persone, i numeri del progetto

A marzo 2019 risale la presentazione del progetto⁵, a giugno 2020 data la prima pubblicazione di EDL, mentre a novembre 2020 vengono conclusi i lavori di digitalizzazione e chiusi i cantieri in Biblioteca, almeno quelli della prima fase (quello dell'Estense Digital Library è un ciclo continuo di pianificazione, esecuzione e controllo, e alla prima sono già seguite altre due fasi). Per la revisione dei metadati e il caricamento finale di tutti i materiali bisogna aspettare marzo 2021.

A prestare la loro opera e la loro competenza specifica sono oltre 50 professionisti, 3 team di informatici e numerose figure di supporto, provenienti da più di 10 tra aziende e istituzioni.

9705 sono gli oggetti digitali pubblicati: 633 carte geografiche, 4171 manoscritti musicali, 3442 fascicoli, di cui 3277 dell'Archivio Muratoriano, cui si aggiungono 277 manoscritti e 1382 libri antichi, risultato delle digitalizzazioni effettuate ex-novo e del recupero delle digitalizzazioni pregresse, realizzate in campagne passate, ma mai pubblicate in rete. I numeri, naturalmente, crescono velocemente e a settembre 2023, al momento di licenziare questo contributo, sono 48734 gli *items* pubblicati, destinati a diventare oltre 145.000 nel prossimo dicembre, quando sarà conclusa la pubblicazione dell'intera Autografoteca Campori⁶.

⁵ Il progetto è stato presentato il 15 marzo 2019 al Convegno delle Stelline, cfr. Zanni, 2019. Finanziato dalla Fondazione Cassa di Risparmio di Modena (FCRMO), è stato aggiudicato all'associazione temporanea di imprese Extense, costituita da Franco Cosimo Panini, Horizons unlimited, Hyperborea e Mida.

⁶ L'Autografoteca Campori è ricca di oltre 135.000 autografi, raccolti nella seconda metà del XIX secolo dal nobile collezionista modenese marchese Giuseppe Campori (1821-1887); lasciata da questi in eredità al Comune di Modena, con obbligo testamentario di deposito permanente presso la Biblioteca Estense, la collezione abbraccia cinque secoli di storia e conserva corrispondenze di nobili e principi di tutta Europa, lettere di poeti, artisti, musicisti, grandi scienziati e medici, ma anche diplomi di sovrani, papi e imperatori. La maggior parte della documentazione è di natura epistolare, ma si trovano anche altri materiali, quali incisioni, disegni, opuscoli, poesie,

Le fasi del progetto

L'intervento ha previsto le seguenti macro-attività:

- Coordinamento del progetto
- Censimento, stabilizzazione e restauro dei materiali
- Recupero del pregresso e digitalizzazione ex-novo di manoscritti musicali e mappe (fogli sciolti, piegati, grandi formati)
- Metadattazione e ricerca
- Pubblicazione on-line, crowdsourcing

Sono stati alimentati quotidianamente due laboratori di digitalizzazione con 4 macchine attive e un laboratorio di restauro, appositamente allestito all'interno dell'Istituto. L'entità del progetto ha reso naturalmente imprescindibile un tavolo di coordinamento, composto da Martina Bagnoli, direttrice delle Gallerie Estensi e responsabile scientifico del progetto e dagli esponenti di Franco Cosimo Panini Editore, Hyperborea e MLOL di Horizons unlimited.

Se la fase di censimento è stata fondamentale per pianificare, guidare e rendere efficiente l'intero processo di lavorazione (restauro, digitalizzazione, metadattazione), fino ad arrivare all'effettiva pubblicazione delle immagini e dei metadati sulla Digital Library, di particolare importanza e novità⁷ è stato aver previsto un'attività di conservazione propedeutica alla digitalizzazione; a una restauratrice, infatti, è stato sottoposto l'intero corpus dei materiali, per valutarne lo stato di conservazione e gli interventi necessari. La stessa ha lavorato in costante coordinamento con il personale e i team di digitalizzazione, con l'obiettivo di guidare nella manipolazione e limitare lo stress a cui possono essere sottoposti i documenti oggetto delle acquisizioni. La restauratrice, inoltre, ha stabilizzato i documenti più fragili prima di affidarli alle postazioni di scansione, nonché monitorato costantemente i valori termo-igrometrici dei laboratori di scansione, per non sottoporre i documenti al cambiamento improvviso delle condizioni negli ambienti in cui sono stati spostati nelle varie fasi del processo, rispetto ai depositi climatizzati.

trattati accademici e molto altro. Per uno studio complessivo dell'Autografoteca Campori, la sua costituzione e l'accrescimento della raccolta si rimanda al volume curato da Al Kalak & Fumagalli, 2022; vedi, in particolare, Sandoni, 2022.

⁷ Ulteriore elemento di novità degno di nota è l'aver incluso nel progetto l'iniziativa "Digitiamo", laboratorio di inserimento lavorativo per adolescenti e giovani con autismo, in collaborazione con l'associazione Aut Aut. I ragazzi sono stati coinvolti nell'informatizzazione dei cataloghi ottocenteschi dei manoscritti Campori; il lavoro, propedeutico all'estrazione dei metadati, è stato già utilizzato per la pubblicazione dei primi manoscritti sulla piattaforma digitale.

I problemi, le scelte, le soluzioni

Sono stati affrontati problemi sia di natura tecnologica (per la scansione dei grandi formati è stata adattata una macchina modificata ad hoc, in precedenza utilizzata da Haltadefinizione per l'acquisizione verticale di dipinti quali *La Nascita di Venere* di Botticelli), che di natura logistica e organizzativa, dettati dalla grande quantità di materiali da movimentare e dalle difficoltà legate ai locali di una sede storica come quella della Biblioteca Estense Universitaria, dove impiantare i macchinari. Ma sono state senza dubbio le fonti dei metadati a impegnare maggiormente.

Come è naturale che sia, il patrimonio della BEU, descritto in cataloghi storici manoscritti e/o dattiloscritti, è stato fatto oggetto in passato di lavori che hanno portato alla costituzione di banche dati locali, ma anche di più recenti campagne di catalogazione informatizzate nell'ambito di specifici progetti di carattere nazionale in rete (Internet Culturale, Manus, BiblioMO-SBN); a questo si è aggiunto il lavoro di studio e descrizione svolto da 3 assegnisti di ricerca⁸ nell'ambito del progetto stesso per nuclei specifici di materiali.

È stato più volte sottolineato, da ultimo nella Relazione di consegna deliverable *milestone M5* - 31 marzo 2021, che le fonti individuate per ricavare i metadati descrittivi da associare alle immagini e da pubblicare «sono risultate notevolmente eterogenee e difformi, sia per quanto riguarda la semantica delle informazioni contenute, che in parte si completano, in parte si sovrappongono e in alcuni casi si contraddicono, sia per quanto riguarda gli aspetti tecnico-descrittivi, con banche dati estremamente variegate per standard descrittivi adottati e per livello di dettaglio e completezza delle descrizioni, per formalismi e standard utilizzati nella registrazione dei dati, per tecnologie e supporti di conservazione e via dicendo».

A partire dalle informazioni ricavate dalle fonti disponibili, sono state quindi effettuate mappature per riportare quelle informazioni in tracciati descrittivi specifici per le diverse tipologie di materiali, nell'ambiente di back office, per poi procedere, nel front end pubblico, a un ulteriore passaggio di semplificazione, riconducendo i diversi tracciati a un unico schema descrittivo.

Ma per arrivare a una banca dati il più possibile omogenea, sono state necessarie successive attività di revisione sia dei tracciati utilizzati che dei dati.

Resta da sottolineare che l'incompletezza delle informazioni relative a molte delle opere disponibili sulla piattaforma non è stata considerata un ostacolo alla pubblicazione, ma un valore per gli studiosi che si vorranno cimentare con

⁸ Già nel 2017 le Gallerie Estensi, con il supporto della FCRMO hanno avviato 3 assegni di ricerca, in convenzione con l'Università di Modena e Reggio Emilia, sui tre fondi poi oggetto del progetto di digitalizzazione: Fondo musicale estense, Fondo cartografico e Archivio Muratoriano.

nuovi studi e scoperte e che sono invitati a contribuire con le loro segnalazioni. Quanto pubblicato non è punto di arrivo, ma punto di partenza: sono stati pubblicati non solo i risultati di schedature e studi approfonditi, ma anche le informazioni minime che possano diventare occasione per aggregare conoscenze intorno a fondi e collezioni poco studiate.

Architettura della piattaforma

La piattaforma di EDL è costituita da 3 moduli applicativi specializzati:

1. un software per la gestione articolata di informazioni catalografiche multi ambito (Arianna4work di Hyperborea);
2. un DAM (Digital Asset Manager) per la gestione delle riproduzioni digitali (Jarvis di Memooria), infrastruttura di long term preservation;
3. un front end per la pubblicazione e la restituzione di informazioni e immagini (MLOL).

Dopo il rilascio nel giugno 2020 sono state apportate modifiche e migliorie, sia nel back end che nel front end.

Per la compilazione dei metadati, Hyperborea è stata incaricata dell'implementazione di una procedura automatica di importazione dei metadati descrittivi disponibili, con mappatura dei tracciati; in particolare è stata sviluppata l'importazione dei metadati descrittivi dei manoscritti, attraverso l'export delle schede descrittive da ManusOnline.

Da una versione appositamente adattata alle esigenze iniziali specifiche del progetto, EDL è migrata sulla piattaforma *standard* MLOL, a seguito di valutazioni economiche, supportate dall'uso.

È stato, in ogni caso, possibile effettuare una revisione della scheda di pubblicazione dei metadati, con aggiunta di campi e modalità di presentazione e un potenziamento dei campi indicizzati nella modalità di ricerca base e in quella avanzata. In particolare è stato abbandonato lo schema unico semplificato utilizzato per la pubblicazione in piattaforma, e oggi MLOL dialoga direttamente con Arianna e i diversi tracciati.

IIIF

EDL è tra le prime, se non la prima biblioteca in Italia ad adottare l'International Image Interoperability Framework. Proprio al centro del progetto è, infatti, l'utilizzo di IIIF⁹, un set di standard open source per la presentazione, la ricerca e l'annotazione di immagini ad alta definizione, nato alla fine del 2011 e condiviso e adottato dalle più importanti biblioteche internazionali (tra le

⁹ Per una disamina delle caratteristiche salienti di IIIF cfr. Salarelli, 2017.

quali, in Europa, la British Library, la BNF, le Bodleian Libraries di Oxford e la Biblioteca Apostolica Vaticana).

Attraverso viewer specifici come Mirador è possibile non solo manipolare l'immagine con strumenti basilari di photoediting (contrasto, colori...), ma selezionare e comparare risorse provenienti da repository e digital libraries diverse, ritagliare immagini e condividerle con un click.

Funzionalità e potenzialità

EDL garantisce l'accesso, ad alta definizione, allo straordinario patrimonio estense, di codici, libri antichi, carte d'archivio, e ne permette lo scarico in formato PDF, ma non solo. In EDL è possibile creare liste, private o pubbliche, di documenti, creare storie, fare annotazioni.

La piattaforma estense vuole non solo valorizzare lo straordinario patrimonio di fonti raccolte, custodite e messe a disposizione del pubblico, ma anche implementare nuove forme di mediazione, comunicazione e didattica, creando un ambiente fertile per la produzione di contenuti da parte dell'utente, chiamato non solo alla fruizione del patrimonio, ma "attivato" a contribuire alla creazione culturale.

Con particolare attenzione al tema del riuso, della partecipazione e della narrazione, sono stati organizzati quattro laboratori per mettere alla prova il rinnovato e migliorato strumento di creazione delle storie, un nuovo mezzo per utilizzare le fonti, disseminare e condividere la conoscenza.

Le Gallerie Estensi, MLOL e l'Ufficio Scolastico Regionale per l'Emilia Romagna hanno organizzato quattro incontri¹⁰ rivolti agli insegnanti e agli studenti delle scuole di ogni ordine e grado per promuovere la conoscenza e l'utilizzo della piattaforma digitale EDL e in particolare degli strumenti di condivisione.

Con lo strumento delle Storie MLOL¹¹, ogni utente della piattaforma digitale delle Gallerie Estensi può utilizzare tutte le risorse digitali per creare dei percorsi visuali all'interno di un'immagine o tra dettagli di immagini differenti, inserendo descrizioni e annotazioni. Le Storie consentono infatti di realizzare approfondimenti, percorsi didattici, possono essere utilizzate efficacemente come supporto alle lezioni, ma possono rivelarsi divertenti strumenti anche

¹⁰ Sono stati realizzati quattro incontri, ognuno su tematiche diverse, tutti caratterizzati da una marcata lettura interdisciplinare delle opere e dei materiali. I corsi sono stati condotti da Isabella Manchia, digital librarian di MLOL, Lucia Peruzzi, operatrice museale delle Gallerie Estensi e presidente di Prospectiva e i docenti Gabriele Benassi (coordinatore delle Equipe Formative presso USR Emilia-Romagna) ed Elisabetta Siboni (docente del Servizio Marconi TSI presso USR Emilia-Romagna).

¹¹ <<https://iiif.medialibrary.it/stories/>>.

per esercitazioni da parte degli studenti, sollecitando da parte loro la rielaborazione delle fonti e la creatività.

Se le liste, costruite da docenti e studenti, sono state definite «archivi di sapere, scaffali di risorse», le storie possono essere utilizzate dai docenti come risorse tematiche complementari al libro di testo, strumenti per approfondire argomenti disciplinari e interdisciplinari, fino a diventare produzioni di scrittura creativa.

Negli ultimi anni abbiamo messo a disposizione del pubblico una enorme quantità di dati, vogliamo che questi siano utilizzati, arricchiti¹², corretti, se necessario, e disseminati a creare nuova conoscenza e nuovi servizi.

Riferimenti bibliografici

- AIB (Associazione Italiana Biblioteche) 2020, *Nuovo Manifesto per le biblioteche digitali*, <<https://www.aib.it/struttura/commissioni-e-gruppi/gruppo-di-lavoro-biblioteche-digitali/2020/82764-nuovo-manifesto-per-le-biblioteche-digitali/>>.
- Al Kalak, M. & Fumagalli, E. (eds.) (2022), *Collezionare autografi. La raccolta di Giuseppe Campori*, Firenze: Olschki.
- Salarelli, A. (2017), 'International Image Interoperability Framework (IIIF): una panoramica', *JLIS.it*, vol. 8, no. 1, pp. 50-66, DOI <<https://doi.org/10.4403/jlis.it-385>>.
- Sandoni, L. (2022), 'Per una storia del collezionismo d'autografi nell'Ottocento. Il caso dell'Autografoteca Campori di Modena', in: Al Kalak, M. & Fumagalli, E. (eds.) (2022), *Collezionare autografi. La raccolta di Giuseppe Campori*, pp. 75-102. Firenze: Olschki, <<https://iris.unimore.it/handle/11380/1285762>>.
- Zanni, A. (2019), 'Il progetto di digitalizzazione della Biblioteca Estense Universitaria di Modena', *Biblioteche oggi*, vol. 37, no. 4, pp. 63-67, DOI <<http://dx.doi.org/10.3302/0392-8586-201904-063-1>>.

¹² È il caso della collaborazione appena avviata con il prof. Luca Frassinetti dell'Università della Campania "Luigi Vanvitelli" per la revisione dei metadati relativi alla messa on-line dei documenti dell'Autografoteca Campori, per la parte relativa alle circa 2.000 unità riconducibili alla corrispondenza passiva del poeta Vincenzo Monti.

