

Trattamento delle risorse bibliografiche in alfabeti non latini e *discoverability* delle collezioni digitali: problemi e prospettive

di Domenico Ciccarello

L'informazione multilingue in biblioteca

La globalizzazione sociale in atto e il ruolo dominante assunto da Internet hanno riportato in primo piano la concretezza dell'impegno necessario al servizio bibliotecario pubblico per mantenere fede alla sua tensione sommamente inclusiva, che si traduce nell'aspirazione di costituire la 'porta di accesso locale alla conoscenza globale'. Sgombrando il campo da ogni ideale - oggettivamente inattuabile - di controllo bibliografico (e offerta) universale, l'ampiezza di orizzonti che i principi internazionali, a partire dal *Public Library Manifesto* fin dalle prime edizioni, attribuiscono alla biblioteca pubblica, va riferito non tanto al grado di esaustività di copertura della conoscenza prodotta (il contenuto), quanto alla garanzia dell'effettivo grado di *empowerment* dell'intera comunità servita (il target degli utenti da raggiungere): «I servizi della biblioteca pubblica sono erogati sulla base delle pari opportunità di accesso per tutti, indipendentemente dall'età, dall'etnia, dal genere, dalla religione, dalla nazionalità, dalla lingua, dallo status sociale e da qualsiasi altra caratteristica»¹.

DOMENICO CICCARELLO, Università degli studi di Palermo, Dipartimento di Culture e società, e-mail: domenico.ciccarello@unipa.it.

Ultima consultazione siti web: 20 dicembre 2024.

Questo lavoro è stato supportato dal progetto PNRR 'Italian Strengthening of Esfri RI Resilience' (ITSERR), finanziato dall'Unione Europea – NextGenerationEU (CUP: B53C22001770006), Principal Investigator: Fabrizio D'Avenia; Deputy Principal Investigator: Gianmarco Braghi. I punti di vista e le opinioni espresse sono esclusivamente dell'autore, e non riflettono necessariamente quelli dell'Unione Europea, che non può esserne in alcun modo considerata responsabile. L'autore, in qualità di RTD-A presso l'Università di Palermo, partecipa in particolare alle attività del Work Package 5, *Digital Maktaba* (WP Leader: Fabrizio D'Avenia; Product owners: Federico Ruozzi e Sonia Bergamaschi, Università di Modena e Reggio Emilia). Ai professori D'Avenia, Braghi, Ruozzi, Bergamaschi, e agli altri membri del WP5 (Sania Aftar, Domenico Beneventano, Amina El Ganadi, Luca Gagliardelli, Luca Sala, Giovanni Sullutrone, Riccardo Vigliermo), esprimo il mio più sentito ringraziamento. Desidero inoltre manifestare la mia gratitudine a Mauro Guerrini e Carlo Bianchini per gli incoraggiamenti ricevuti, e a Manuela Grillo e Roberto Raieli per il prezioso supporto bibliografico.

1 IFLA-UNESCO, *Manifesto IFLA-UNESCO delle biblioteche pubbliche 2022*, 2022, <<https://repository.ifla.org/handle/123456789/2022>>.



Gli estensori del *Manifesto* sanno bene che il fattore linguistico immanente alle risorse bibliografiche richiederà, di norma, un input differenziale ai gestori delle biblioteche perché si ottenga l'equità di servizio auspicata: «Si devono fornire servizi e materiali specifici a quegli utenti, ad esempio le minoranze linguistiche [...] che, per qualsiasi motivo, non possono utilizzare i servizi e i materiali ordinari»². Rilevando quindi, in premessa, che la questione dell'informazione multilingue concerne, per ovvii motivi che riguardano i suoi stessi presupposti, non solo le biblioteche pubbliche di base ma anche le biblioteche scolastiche, universitarie, specializzate, ecclesiastiche etc., proverò nei paragrafi che seguono a fare il punto su come stiano evolvendo i fattori abilitanti all'uso efficace delle risorse multi-alfabeto, mantenendo una distinzione d'ambito – per motivi che saranno chiariti nel corso del ragionamento – tra le raccolte fisiche descritte nei cataloghi di biblioteca e le collezioni ibride rese accessibili tramite discovery service o piattaforme/interfacce di biblioteche digitali.

Decolonizzare il catalogo: una questione biblioteconomica

Per iniziare il discorso, mi preme sottolineare come la questione della catalogazione multilingue dovrebbe trovare la sua collocazione più idonea nell'ambito della cosiddetta *biblioteconomia critica*, piuttosto che soltanto nella sfera delle tecniche catalografiche, che naturalmente vi sono implicate a pieno titolo. O perlomeno sarà questa la prospettiva entro cui, cercando di raccogliere alcuni degli spunti recentemente proposti da Rossana Morriello³, vorrei provare a riflettere sull'argomento. Se è vero che – per tornare alle dichiarazioni del *Manifesto* IFLA-UNESCO – «i servizi bibliotecari devono essere adattati alle diverse esigenze [...] degli utenti multilingui»⁴, dovremmo anzitutto verificare in che misura, nel nostro Paese, tale adattamento ai bisogni delle comunità di madrelingua straniera sia stato finora realizzato. In anni ormai lontani, tale approccio squisitamente biblioteconomico è stato condotto ad ampio raggio dall'AIB, inizialmente in seno alla CNBP⁵, e poi con l'attività del Gruppo di studio Biblioteche multiculturali⁶, affrontandone sulla scorta delle linee guida internazionali le principali questioni politico-programmatiche. Non solo veniva approfondito il tema della catalogazione in alfabeti non latini, che

2 *Ivi*.

3 Rossana Morriello, *Biblioteconomia critica e decolonizzazione delle collezioni*, «Biblioteche oggi», 39 (2021), n. 2, p. 25-38, DOI: 10.3302/0392-8586-202102-025-1.

4 *Manifesto* cit.

5 Cfr. *Linee guida per i servizi multiculturali nelle biblioteche pubbliche*, a cura della Commissione nazionale biblioteche pubbliche [AIB]; testi di Lucia Bassanese ... [et al.]; traduzione italiana del documento IFLA *Multicultural communities: guidelines for library services* a cura di Alberta Dellepiane e Adriana Pietrangeli. Roma: Associazione italiana biblioteche, 2003.

6 Nell'ambito dell'attività di coordinamento di questo Gruppo di studio AIB e come membro e poi corrispondente della relativa sezione IFLA Library Services to Multicultural Populations, chi scrive, tra il 2004 e il 2008, ha tradotto dall'inglese diversi documenti di advocacy, tra cui le *Dieci ragioni per offrire servizi bibliotecari multiculturali*, e il *Manifesto* IFLA. Si veda: Clara M. Chu, *Raison d' être per i servizi bibliotecari multiculturali*, «AIB notizie», 17 (2005), n. 3-4, p. III; *Manifesto IFLA per la biblioteca multiculturale: la biblioteca multiculturale. Porta di accesso a una società di culture diverse in dialogo*, «AIB notizie», 21 (2009), n. 3, p. 14-15, anche online nel sito dell'IFLA, <<https://repository.ifla.org/handle/20.500.14598/743>>.

pure è stato oggetto di un apposito seminario⁷, ma si individuava l'analisi di comunità come strumento utile a fare emergere meglio i bisogni dei nuovi arrivati⁸; si guardava alla cooperazione con le scuole, i primi istituti educativi ad essere stati investiti dall'"ondata multi-etnica"⁹, e con gli enti del terzo settore; si studiavano e confrontavano diverse formule per la costruzione dello 'scaffale multiculturale' e la sua promozione verso gli utenti¹⁰; si osservavano le buone pratiche internazionali, con lo sguardo rivolto ai metodi più efficaci per offrire informazione di comunità multilingue, anche tramite le directory istituzionali di siti web (*virtual reference desks*)¹¹. Le riflessioni prodotte dai bibliotecari che hanno partecipato a quelle elaborazioni¹² convergevano su conclusioni pressoché unanimi, ribadendo a più riprese la necessità di superare prioritariamente due problemi (o se si vuole, due anomalie) della tradizione bibliotecaria italiana: il primo, quello di rimanere ancorata, adattandosi in modo forse troppo acritico e salvo poche eccezioni, a una prassi di descrizione e indicizzazione delle risorse in cui i titoli delle opere e ogni altro punto di accesso preferito ai documenti venivano definiti esclusivamente tramite la forma traslitterata delle voci di autorità (l'unica ammessa nell'indice SBN); il secondo, quello della mancanza di soluzioni innovative, di fronte ai limiti della normativa nazionale sull'as-

7 Il seminario "Da Babele a Alessandria? Catalogazione e alfabeti non latini", organizzato dalla Commissione nazionale biblioteche pubbliche dell'AIB, si è tenuto a Venezia, Fondazione Querini Stampalia, il 2 ottobre 2003. Un resoconto agile in Stefania Marzocchi, *Biblioteche di pubblica lettura e materiali in alfabeti non latini: un po' di teoria e un po' di pratica*, «Bibliotime», 6 (2003), n. 3, <<https://www.aib.it/aib/sezioni/emr/bibtime/num-vi-3/marzocch.htm>>.

8 Una sintesi in Arianna Buson, *La biblioteca multiculturale*. Roma: Associazione italiana biblioteche, 2013, p. 28-34; per una visione più generale sull'argomento, rimando ad Anna Galluzzi, *Analisi di comunità: uno strumento per la pianificazione dei servizi*, «Bollettino AIB», 41 (2001), n. 2, p. 175-188.

9 Richiamo volutamente il termine adoperato in un libretto che in quella fase aveva avuto molta fortuna: Vinicio Ongini, *La biblioteca multi-etnica: libri, percorsi, proposte per un incontro tra culture diverse*. Milano: Editrice bibliografica, 1992.

10 *Id.*, *Lo scaffale multiculturale*. Milano: Mondadori, 1999.

11 Una breve rassegna con esempi italiani in Domenico Ciccarello, *I servizi multiculturali delle biblioteche. Quali tendenze in Italia e nel mondo*. «Biblioteche oggi», 27 (2009), n. 4, p. 25-30. Sui VRD in particolare cfr. *Id.*, *Virtual reference desks e informazione di comunità multilingue: una rassegna internazionale*, «Culture del testo e del documento», 10 (2009), n. 29, p. 135-145.

12 Alcuni esempi, non esaustivi, del dibattito di quegli anni: Cecilia Cognigni, *Intercultura e biblioteca pubblica: il dibattito internazionale*, «Bollettino AIB», 42 (2002), n. 4, p. 451-468; *Ead.*, *La comunicazione interculturale: alcune riflessioni sul ruolo delle biblioteche pubbliche*, «Biblioteche oggi», 22 (2004), n. 1, p. 37-40; Alessia Ceccarelli, *Nella mia biblioteca nessuno è straniero*, «Biblioteche oggi», 21 (2003), n. 7, p. 25-31; Lorenzo Luatti, *Biblioteche e intercultura*. In: *La città plurale: trasformazioni urbane e servizi interculturali*, a cura di Lorenzo Luatti. Bologna: EMI, 2006, p. 299-339. Di indubbio interesse anche i materiali della sessione *Biblioteche pubbliche: la sfida multiculturale del 49° convegno AIB*, con interventi, tra gli altri, di Laura Bassanese, Manuela Venturelli, Claudia Parmeggiani, Antonio di Vieste, cfr. *Bibliocom 2002. Atti del 49° Congresso nazionale dell'Associazione italiana biblioteche, Roma, 15-17 ottobre 2002*. Roma: Associazione italiana biblioteche, 2006, p. 127-176. Ne riassumevo parzialmente i termini in Domenico Ciccarello, *Biblioteche pubbliche: i servizi multiculturali*. In: *Rapporto sulle biblioteche italiane 2004*, a cura di Vittorio Ponzani. Roma: Associazione italiana biblioteche, 2004, p. 45-51.

sunzione nei ruoli del pubblico impiego, per consentire la presenza e il contributo operativo di staff madrelingua, multilingue e multiculturale, nelle nostre biblioteche¹³. Si tratta, come si può vedere, di un complesso di questioni interconnesse, di etica e di policy prima ancora che tecnologiche, riconducibili senz'altro a quelli che con terminologia più attuale oggi vengono chiamati 'decolonialità' o anche 'principi DEI'¹⁴, e che senz'altro si ponevano già allora molto oltre un orizzonte puramente catalografico. Rimane la constatazione di una mancanza di fluidità nel percorso di sviluppo che, grazie ai progressi dell'automazione, avrebbe dovuto condurre in pochi anni le nostre biblioteche a una piena integrazione degli script non latini nei cataloghi elettronici, fattore che continua a incidere in modo fortemente negativo sulla qualità dell'informazione offerta dai nostri istituti¹⁵. Un nodo che, nonostante le dichiarazioni di intenti che accomunavano le istituzioni partecipanti al seminario di Venezia del 2003, non è stato ancora sciolto. In proposito, svilupperemo qui di seguito una serie di ulteriori osservazioni, anzitutto ripercorrendo nuovamente qualche passo indietro.

Traslitterazione e script originale. Rivalità o convivenza necessaria?

Il documento conclusivo del citato seminario "Babele e Alessandria", scaturito dalle relazioni e dal confronto tra i partecipanti, e infine redatto da Chiara Rabitti per la CNBP dell'AIB, accanto alla questione centrale delle lingue e delle scritture del catalogo bibliografico (gestione delle registrazioni nei caratteri propri della lingua originale e uniformità dei sistemi di traslitterazione), richiamava l'attenzione anche sulla necessità di alcuni interventi di politica bibliotecaria necessariamente connessi all'obiettivo dell'equità di trattamento sul piano linguistico e culturale. Tali interventi, come concordavano i partecipanti, avrebbero dovuto essere indirizzati: 1) alla formazione di figure professionali competenti, in un *blending* mirato di competenze linguistico-culturali e tecnico-catalografiche; 2) all'adozione di un approccio di tipo cooperativo-consortile, basato sullo *human resource sharing* ovvero su programmi a sostegno della diffusione del know-how dei punti di eccellenza e della circolazione delle migliori pratiche; 3) alla creazione di gruppi di lavoro, coordinati a livello ministeriale, per l'individuazione delle soluzioni tecniche più idonee a migliorare le regole di catalogazione con riguardo ai sistemi di traslitterazione, come anche ad adeguare il sistema SBN secondo le esigenze degli utenti relative alla ricerca e al recupero delle risorse bibliografiche con l'impiego delle lingue e scritture non latine¹⁶.

13 Domenico Ciccarello, *La professione del bibliotecario nella società multiculturale*. In: *Le politiche delle biblioteche in Italia: la professione. Atti del 53° Congresso nazionale AIB*. Roma: Associazione italiana biblioteche, 2007, p. 107-114.

14 L'acronimo DEI sta per "diversity, equity, and inclusion"; cfr. Rossana Morriello; Lucia Sardo, *Biblioteconomia e decolonialità: collezioni e cataloghi per una visione altra*. In: *Visioni future: next generation library, relazioni di giovedì 30 marzo*, convegno delle Stelline (Milano, 30-31 marzo 2023). Milano: Editrice bibliografica, 2023, p. 165-174 (2 vol.).

15 Per un confronto con il dibattito internazionale, può essere utile segnalare che in ambito IFLA la questione era già sotto i riflettori trent'anni fa, cfr.: *Multi-script, multilingual, multi-character issues for online environment. Proceedings of a Workshop sponsored by the IFLA Section on cataloguing, Istanbul, Turkey, August 24, 1995*, edited by John D. Byrum and Olivia Madison. München: Saur, 1998.

16 Il documento, a suo tempo pubblicato in AIB-WEB col titolo *Catalogazione e alfabeti non latini*, è tuttora reperibile all'URL: <<https://www.aib.it/aib/commiss/mc/rabchio1.htm>>

La rielaborazione di uno dei contributi a quel convegno, presentata da Barbara Poli¹⁷ in questa rivista a qualche anno di distanza (2009), offriva ulteriori stimoli di non poco conto. In quel periodo, la recente adozione dello standard UNICODE nei principali sistemi gestionali di biblioteca aveva favorito una svolta sul piano tecnico, aprendo la strada a una decodifica più completa e uniforme dei caratteri non latini¹⁸. La bibliotecaria della Querini Stampalia aveva quindi sviluppato un'indagine sulla realtà statunitense, che già una quindicina di anni fa esibiva trasformazioni e realizzazioni di un certo impatto, ben fotografate dall'autrice. Tra queste, per il mondo OCLC, "Connexion" come sw di trattamento in back office dei dati bibliografici, che permetteva la creazione e la modifica di registrazioni del catalogo in dieci script diversi dal latino e dal greco (anche con l'opzione di potere ripetere uno stesso tag per più lingue-script), oltre ad offrire per i catalogatori l'accesso ad alcune authority list onomastiche¹⁹, e – per alcune lingue – ad uno strumento di conversione automatica dei dati bibliografici, dalle forme latinizzate a quelle nella scrittura originale, e viceversa, denominato "Transliteration tools"; inoltre i "Language sets" per aiutare biblioteche anche medio-piccole a costruire delle collezioni bibliografiche multilingue per le esigenze delle comunità locali; e "FirstSearch" come strumento di interrogazione del metaOPAC consortile con interfaccia utente in inglese, francese, spagnolo, arabo, cinese, giapponese e coreano.

Quanto agli orientamenti della Library of Congress, Poli annotava la scelta del *merging* delle registrazioni prodotte in diverse lingue/scritture originali (giapponese, cinese, coreano, arabo, persiano, ebraico e yiddish) all'interno di un database unico, "MDS-Books All"; evidenziava il recente impiego dei tag 4XX del MARC per registrare le voci d'autorità in caratteri non latini (non preferite), accanto al tradizionale trattamento mediante il tag 880 (*alternate graphic representation*), il che aveva portato al riversamento su NACO²⁰ (e quindi alla piena disponibilità per i catalogatori) di circa mezzo milione di intestazioni nominali nella scrittura originale; non trattandosi, tuttavia, di forme pienamente controllate, il catalogo presentava abbondanza di duplicati.

L'autrice quindi menzionava alcuni interessanti programmi di catalogazione cooperativa in corso nel continente europeo: il "CJK Allegro"²¹, frutto degli sforzi congiunti di alcune biblioteche tedesche e inglesi; lo European Union Catalogue of

17 Barbara Poli, *Le scritture del catalogo: uno sguardo alle esperienze straniere*, «Bollettino AIB», 49 (2009), n. 1, p. 47-53.

18 Riferimenti precoci, in questo senso, si rinvengono in Joan M. Aliprand, *The Unicode standard: its scope, design principles and prospects for international cataloguing*, «Library resources and technical services», 44 (2000), n. 3, p. 160-167, DOI: 10.5860/lrts.44n3.160; Laura Tull, *Library systems and Unicode: a review of the current state of development*, «Information technology and libraries», 21 (2002), n. 4, p. 181-185.

19 Ad esempio, il *Chinese name authority file*, sviluppato dal Joint University Librarians Advisory Committee (JULAC) di Hong Kong.

20 LC-NACO (Name Authority Cooperative Program) è uno dei programmi di catalogazione partecipata della Library of Congress, dedicato all'alimentazione dell'authority file di autori, luoghi e titoli uniformi del catalogo.

21 CJK (Chinese Japanese Korean) è l'acronimo impiegato per il gruppo di lingue/scritture cinese-giapponese-coreano; Allegro è un sw di gestione catalografica sviluppato in Germania negli anni Ottanta.

Japanese Books²², curato da un gruppo di informatici di Tokyo; lo Union Catalogue of Chinese Books britannico (fig. 1)²³;



Figura 1 – Union Catalogue of Chinese Books. Visualizzazione di una descrizione in caratteri cinesi

infine, per la Francia, la stratificazione del trattamento di lingue e scritture diverse in OPALE Plus della Bibliothèque Nationale, e le raccomandazioni espresse in un rapporto tecnico curato da Danièle Duclos-Faure per il Ministero dell'Educazione²⁴, anche in vista dell'istituzione (avvenuta nel 2008) della Bibliothèque Universitaire des Langues et Civilisations. Poli concludeva auspicando nuovi passi in avanti anche per la situazione italiana, con l'occasione dell'implementazione delle nuove *Regole di catalogazione* (REICAT).

Il modello statunitense, basato sull'opzione di servizi specializzati “chiavi in mano”, che includevano sia la consulenza per la selezione e l'acquisizione di pacchetti di materiali bibliografici (utili, a seconda dei casi, per soddisfare le esigenze delle biblioteche pubbliche o di quelle universitarie e scolastiche), sia la gestione e acquisizione delle relative registrazioni bibliografiche, corredate delle pertinenti voci di autorità, ha trovato un corrispettivo molto simile nell'organizzazione bibliotecaria australiana. Con l'impiego del sw Aleph di Ex-Libris e l'implementazione di Unicode, lo staff di catalogatori multilingue della CAVAL Collaborative Solutions, partner della rete bibliotecaria pubblica di Victoria, coordinata dalla State Library of Victoria, ha garantito (per diversi anni, fino alla ristrutturazione dei servizi nazionali nel 2014) la gestione di un ambiente di catalogazione in grado di supportare la redazione di registrazioni bibliografiche in sessanta lingue differenti, nel rispetto delle scritture originali e con la possibilità di scambio bibliografico tra database diver-

22 Il catalogo collettivo, originariamente sviluppato dal NII (National Institute of Informatics), e a cui hanno partecipato biblioteche di diversi paesi europei: Belgio, Germania, Norvegia, Svezia, Svizzera, è recentemente confluito nello Union Catalogue Database of Japanese Texts (detto anche Kokusho Database), ospitato dall'Università di Tokio, a <<https://kokusho.nijl.ac.jp/>>.

23 Il database cumulativo, che riunisce le registrazioni catalogografiche di numerose biblioteche, tra cui la British Library e le universitarie di Oxford e Cambridge, è ancora attivo, <<https://a35.ie/db/rslpchin/a35-app.php>>.

24 Il documento era mirato a supportare il funzionamento della nuova BULAC, che si sarebbe qualificata come “Pôle des langues et civilisations du monde”. Danièle Duclos-Faure, *Réflexions et recommandations pour le catalogage des documents en caractères non latins en France dans le catalogue collectif universitaire, Sudoc*, «International cataloguing and bibliographic control», 32 (2003), n. 2, p. 26-29.

si²⁵. Senza dubbio, il successo del network australiano è stato dovuto alla compresenza di esperti linguistici, informatici, catalogatori nel quadro di un business model che fa del partenariato pubblico-privato una norma, anziché l'eccezione. Il team multidisciplinare ha affrontato da subito, tra gli altri, gli ostacoli dell'ordinamento delle voci d'autorità in un catalogo multilingue, delle problematiche di visualizzazione dei caratteri non latini, della corretta e uniforme resa dei diacritici.

Non si avvertirà mai abbastanza, in ogni caso, sull'estrema delicatezza delle questioni e sull'elevato grado di difficoltà e complicanze che caratterizzano la costruzione del catalogo multilingue e multi-alfabeto: non è una questione che si possa risolvere a buon mercato, con la sola ottimizzazione della decodifica dei caratteri. Ciascuna lingua e scrittura reca con sé specifiche modalità di impiego, particolarità redazionali, implicazioni culturali, i cui effetti arricchiscono l'universo catalografico (ma al tempo stesso, ne complicano non poco il processo descrittivo). Basti pensare, per limitarci ai caratteri di scrittura della lingua araba, al loro orientamento inverso (da destra a sinistra) rispetto a quelli latini, che impegna i sistemisti informatici all'impiego di appositi algoritmi volti a garantire la corretta direzionalità del testo sia in scrittura che in visualizzazione, caratteristica peraltro condivisa con altre lingue²⁶. Oppure si consideri la profonda differenza, invalsa nei Paesi arabi nella costruzione dei nomi di persona, tra autori 'moderni' (dal Novecento in poi) e autori 'classici' (questi ultimi con forme del nome molto estese, che comprendono, oltre al nome proprio e al cognome – derivante da professione, casata o toponimi – anche appellativi spirituali, indicazioni nominali di paternità e di filiazione, dando luogo a molteplici varianti formali a seconda della fonte impiegata), a parte la necessità di qualificarli cronologicamente con una duplice trascrizione di data: calendario islamico e calendario gregoriano²⁷.

Risulta, pertanto, comprensibile come l'accelerazione nelle attività di catalogazione delle collezioni in caratteri non latini sia stata accompagnata, in questi ultimi anni, da una moltiplicazione delle riflessioni in merito alle modalità di replica delle registrazioni delle medesime raccolte nella forma traslitterata, tema che suscita non poche preoccupazioni.

Sono questioni che, per un verso, riguardano le regole nazionali e i protocolli interni ai sistemi adottati, e non è un mistero che in Italia vi sia ancora una forte divergenza tra i due: mentre le nuove REICAT invitano il catalogatore all'equità di trattamento catalografico, ovvero «Informazioni che devono essere riportate o fornite in lingue che utilizzano sistemi di scrittura diversi dall'alfabeto latino si danno, per quanto possibile, sia nella scrittura originale sia in forma traslitterata o trascritta in alfabeto latino»²⁸, nella Guida SBN, pur riconoscendo che per le aree 1, 2 e 4 le informazioni bibliografiche «vengono normalmente ricavate dalla risorsa e sono quindi nella lingua e/o scritture in cui esse appaiono», viene prescritto che «Testi in

25 L'esperienza di CAVAL (Cooperative Action by Victorian Academic Libraries) e della rete VICNET (Victoria Network) è stata illustrata in Cathie Jilovsky; Andrew Cunningham, *Multilingual cataloguing: practical and technical issues*, «Semantic scholar», 2008, <<https://api.semanticscholar.org/CorpusID:203596814>>.

26 In base a tale caratteristica, i catalogatori e gli informatici statunitensi convenzionalmente raggruppano le lingue 'HAPY', cioè l'ebraico, l'arabo, il persiano, l'yiddish.

27 Magda El-Sherbini, *Bibliotheca Alexandrina's model for Arabic name authority control*, «Library resources & technical services», 57 (2012), n. 1, p. 4-17.

28 Paragrafo 0.5.3.

caratteri non latini si traslitterano»²⁹ seguendo l'Appendice F del codice nazionale, e si esclude quindi, di fatto, il ricorso alla scrittura della lingua originale che compare nell'edizione descritta.

Per altro verso, le questioni e le relative riflessioni, spostandoci ora proprio sul versante della traslitterazione, si fanno ulteriormente complicate, dato che, come sappiamo, non vi è affatto soddisfazione, convergenza o accordo sufficiente tra catalogatori, specialisti dei domini linguistici e utenti comuni su quale standard (schema, tabella) sia opportuno adottare. Tale perdurante incertezza, insieme all'accumulo delle prassi difformi pregresse già abbondantemente sedimentatesi nei nostri cataloghi, ha condotto all'aumento delle varianti delle trasposizioni linguistiche, dei duplicati bibliografici ma soprattutto, e ciò che è più grave, moltiplica il rischio di perdita di 'discoverability' cioè inibisce in larga misura le chances di recupero delle risorse bibliografiche a seguito delle ricerche effettuate in catalogo.

Lo dimostrano gli innumerevoli problemi documentati, soprattutto nella rivista «Cataloging & classification quarterly», in merito alle prassi di traslitterazione relative a specifiche lingue/scritture latinizzate nei cataloghi occidentali: tra queste il cinese³⁰, il coreano³¹, il persiano³², il giapponese³³, l'arabo³⁴, il russo³⁵, a testimoniare come la strada da percorrere per la semplificazione dell'accesso alle raccolte multilingue sia ancora molto lunga. Nel contesto italiano, analogamente, sui numerosi (e ulteriori) aspetti e problemi specifici della catalogazione e traslitterazione in specifiche lingue e scritture sono intervenuti in momenti diversi, tra gli altri, Gabriele Mazzitelli (a proposito del cirillico, in questa rivista)³⁶, Federica Olivotto (con riferimento al cinese)³⁷, Chiara Camarda (con riguardo all'ebraico)³⁸.

29 Paragrafo oF.

30 Yue Li, *Consistency versus inconsistency: issues in Chinese cataloguing in OCLC*, «Cataloging & classification quarterly», 38 (2004), n. 2, p. 17-31, DOI: 10.1300/J104v38n02_04.

31 SungKyung Kim, *Romanization in cataloguing of Korean materials*, «Cataloging & classification quarterly», 43 (2006), n. 2, p. 53-76, DOI: 10.1300/J104v43n02_05.

32 Fereshteh Molavi, *Main issues in cataloguing Persian language materials in North America*, «Cataloging & classification quarterly», 43 (2006), n. 2, p. 77-82, DOI: 10.1300/J104v43n02_06.

33 Yoko Kudo, *A study of romanization practice for Japanese language titles in OCLC WorldCat records*, «Cataloging & classification quarterly», 48 (2010), n. 4, p. 279-302, DOI: 10.1080/01639370903338352.

34 Magda El-Sherbini, *Bibliotheca Alexandrina's model* cit. Si veda in particolare l'esempio dedicato alla proliferazione di forme dell'autore 'William Shakespeare' per effetto delle varianti anche minime di traslitterazione (p. 8).

35 Joy DuBose, *Russian, Japanese, and Latin Oh my! Using technology to catalog non-English language titles*, «Cataloging & classification quarterly», 57 (2019), n. 7/8, p. 496-506, DOI: 10.1080/01639374.2019.1671929.

36 Gabriele Mazzitelli, *Ancora sul problema della traslitterazione dei caratteri cirillici*, «Bollettino AIB», 48 (2008), n. 4, p. 343-358.

37 Federica Olivotto, *La catalogazione del libro cinese: considerazioni pratiche e teoriche*, «Bollettino ABEI», 30 (2021), n. 3, p. 53-65.

38 Chiara Camarda, *Le problematiche descrittive della catalogazione dei libri ebraici*, «Bollettino ABEI», 30 (2021), n. 3, p. 47-52.

Sempre in ambito italiano, merita attenzione un ulteriore contributo molto recente al dibattito, in cui oltre a rilievi su aspetti particolari, vengono discusse alcune anomalie generali che caratterizzano la situazione italiana. Oscar Nalesini, dopo avere espresso, in premessa, le motivazioni di una preferenza per la catalogazione in lingua originale, ed evidenziato le componenti sociologiche sottese al trattamento bibliografico dei materiali in alfabeti diversi dal nostro: «[...] è bene tener presente che ogni scrittura, nonché ogni sua standardizzazione, ha anche connotazioni sociali, religiose, politiche, e di conseguenza il modo di codificare una scrittura in un'altra non è mai una scelta culturalmente neutrale»³⁹, ha affrontato in dettaglio numerose questioni tecniche che si presentano di non facile soluzione con riferimento alle scelte di traslitterazione. La preferenza accordata dalle nuove REICAT agli standard ISO per un buon numero di lingue e/o scritture (tra cui greco, cirillico, arabo, ebraico, cinese), in diversi punti dell'appendice F del codice, non appare esente da contraddizioni; ulteriori difformità applicative, inoltre, vengono ordinariamente introdotte dalle consuetudini operative legate agli ambienti SBN e ai relativi protocolli e vincoli catalografici (insufficiente, ad esempio, il numero complessivo dei caratteri speciali fin qui gestiti dal sistema; imperfette le funzionalità della tastiera virtuale); infine, rimane il fatto che il livello di copertura linguistica delle tabelle di traslitterazione di ALA-LC risulta complessivamente più ampio di quello offerto dalle norme italiane. Ulteriori dubbi e perplessità operative, secondo Nalesini, legittimamente scaturiscono dall'impiego di una medesima scrittura per più lingue diverse, da problemi specifici che si presentano nell'uso dei segni diacritici, nonché dall'impossibilità pratica, dovuta in diversi casi alla natura degli alfabeti stessi, di stabilire corrispondenze biunivoche tra script originali e caratteri latini. Tutto ciò, a sua volta, di fatto impedirebbe di ottenere la 'reversibilità' del processo di traslitterazione (che invece sarebbe fortemente auspicabile, sia per motivi di uniformità, che per facilitare processi automatici di latinizzazione degli script nativi o, viceversa, di 'retroconversione' ovvero ripristino dei caratteri originali a partire dalla forma latinizzata).

In base a tali considerazioni, come ha suggerito Blair Kuntz, «For those opposed to transliteration, especially in an age where computerization has introduced Unicode in which native scripts can be displayed, entered, and searched in library catalogues, the practice is wholly unsatisfactory, serves no-one, and should probably be abolished»⁴⁰. Lo stesso Kuntz, tuttavia, prosegue dimostrando che una qualche forma di traslitterazione, il più possibile affidabile e standardizzata, è di aiuto non solo al catalogatore, per facilitarlo nel controllo bibliografico delle collezioni e nella loro indicizzazione, ma anche agli utenti, in genere molto diversi tra loro quanto a provenienza, quanto a esigenze e motivazioni di ricerca, e infine quanto alla padronanza della lingua in cui sono redatti i documenti cercati. Forse, allora, come spesso acca-

³⁹ Oscar Nalesini, *Libri d'altri, cataloghi di noialtri: le scritture non latine in SBN*, «JLIS.it», 13 (2022), n. 2, p. 128-142: p. 131, DOI: 10.36253/jlis.it-462.

⁴⁰ Blair Kuntz, *Is the Organization for Standardization (ISO)'s Arabic transliteration scheme an improvement over Library of Congress'?*, «MELA notes. Journal of Middle Eastern librarianship», 78 (2005), p. 55-68: p. 55. Scopo dell'articolo è comparare l'efficacia della norma ISO 233-2:1993 con quella delle tavole redatte e aggiornate dall'American Library Association e dalla Library of Congress per l'arabo: *ALA-LC Romanization tables: transliteration schemes for non-roman scripts*, Washington: Library of Congress Cataloging Distribution Service, 1997 (e successive ed.); l'esito del confronto mette in rilievo pregi e difetti di entrambi gli schemi ai fini della resa catalografica. La versione online delle tavole nelle diverse lingue e scritture si trova attualmente a partire da <<https://www.loc.gov/catdir/cpso/roman.html>>.

de, la verità sta nel mezzo, e cioè è necessario sapere sfruttare l'artificio della traslitterazione di titoli, autori e altri elementi descrittivi, ma ciò obiettivamente non appare sufficiente: «Why should a native speaker even bother searching for an item using transliteration, when searching using the original script is so much more reliable and efficient?»⁴¹. Dello stesso avviso James E. Agenbroad, secondo il quale la traslitterazione, in realtà, dovrebbe essere usata solo come punto di accesso aggiuntivo (per maggiore comodità dell'agenzia catalografica ed eventualmente degli utenti) rispetto all'unica modalità di accesso che andrebbe garantita sempre e comunque, quella della forma originale della descrizione e delle voci di autorità: «To provide equal and effective access, a library with a multiscript collection for readers of different scripts needs a multiscript catalog based on cataloguing rules that specify access in the original script. It may also need access via the script most used by most of its readers»⁴². E a proposito di equità, vorrei aggiungere nel paragrafo seguente qualche considerazione su uno dei molti paradossi del digitale.

Tra OPAC e strumenti di discovery delle raccolte digitali

Tra i servizi offerti dalle biblioteche in ambiente digitale, ormai da diversi anni hanno progressivamente preso piede, soprattutto nei contesti delle istituzioni accademiche e di ricerca, i discovery tool. Come sappiamo, si tratta di strumenti di ricerca integrata, 'federata' delle informazioni bibliografiche, in grado di facilitare e massimizzare a beneficio dell'utente finale l'information retrieval dei contenuti bibliografici indicizzati dalle biblioteche⁴³. Il serbatoio informativo a cui attingono i discovery service è costituito da una knowledge base in continuo aggiornamento, che viene alimentata a) dai metadati delle risorse disponibili tramite il catalogo bibliografico, b) dai metadati relativi al flusso delle collezioni digitali in abbonamento attivo, e c) dai metadati della letteratura scientifica liberamente disponibili in modalità open access di cui sia stato dinamicamente effettuato un harvesting a cura del soggetto commerciale produttore, eventualmente con il concorso attivo dell'organizzazione interessata.

La compresenza, accanto alle registrazioni del catalogo, di contenuti digitali nativi che normalmente sono stati prodotti, metadati e archiviati da soggetti esterni alla biblioteca e con sistemi informativi diversi, ha creato in molti casi, tuttavia, il paradosso di una disuguaglianza nelle opportunità di *discoverability* e di presentazione al pubblico delle risorse bibliografiche in caratteri non latini, e ciò vale non solo per i titoli e gli autori, ma anche per i soggetti espressi con il linguaggio naturale. Ci spiegheremo meglio con un esempio. Una ricerca a caso, volutamente effettuata con la parola chiave generica الإسلامية (= islamico/a) tramite un discovery service di Ateneo, tra i molti risultati restituisce quello mostrato come esempio nella fig. 2.

41 *Ivi*.

42 James A. Agenbroad, *Romanization is not enough*, «Cataloging & classification quarterly», 42 (2006), n. 2, p. 21-34: 30, DOI: 10.1300/J104V42N02_03.

43 Cfr. Roberto Raieli, *Web-scale discovery services: principi, applicazioni e ipotesi di sviluppo*. Roma: Associazione italiana biblioteche, 2020.



Figura 2 – Visualizzazione della descrizione relativa a un contenuto digitale in abbonamento, a seguito di query effettuata tramite discovery tool nello script originale

Viceversa, l’utente dell’OPAC del medesimo Ateneo (dando per assodato che quell’OPAC segua le norme nazionali e ne condivida i limiti) sarà costretto a cercare unicamente nella forma traslitterata ‘al-Islāmīyah’ perché possa recuperare dei documenti cartacei in lingua originale pertinenti all’area tematica di suo interesse, seppure a suo tempo tali documenti sono stati appositamente selezionati e acquisiti dal SBA nella lingua e nello script nativo, in accordo con i docenti delle discipline e/o gli esperti madrelingua coinvolti.

In altre parole, nel caso delle risorse in alfabeti non latini, le registrazioni del nostro catalogo basate sulla sola traslitterazione, per quanto dettagliate, accurate e dotate di collegamenti a strumenti di indicizzazione raffinatissimi, sconteranno sempre un gap linguistico che non appare in alcun modo fungibile. Si considerino, ad esempio, i punti di accesso per soggetto alla descrizione mostrata in precedenza (fig. 2). Se possiamo tutti concordare che l’indicizzazione per parole chiave proposta dal produttore commerciale della risorsa digitale è ridondante e qualitativamente inferiore, per precisione e capacità di richiamo, rispetto all’indicizzazione controllata tramite thesaurus che i catalogatori sarebbero capaci di offrire per la medesima registrazione bibliografica (ad esempio con il Nuovo Soggettario BNCf), allora bisognerà pure ammettere che la questione dello script nativo si pone in modo decisivo anche con riguardo all’effetto abilitante che offrono le chiavi di accesso di tipo semantico in lingua originale alle descrizioni in catalogo. Anche in questo caso, come hanno dimostrato Magda el-Sherbini e Sherab Chen in un ampio contributo sul tema, che comprendeva un’indagine di customer satisfaction sull’impiego delle LCSH per i testi in caratteri non latini, realizzata nel 2010 e rivolta sia a bibliotecari che a studenti / utenti finali, la dialettica traslitterazione vs. lingua nativa presenta rilevanti aspetti critici che, di fronte al 68% di preferenze per la ricerca semantica mediante lo script originale, gli autori riassumono così: «In general, the number of statements regarding the difficulties librarians and end users encountered when they searched non-Roman items by controlled English subject terms indicates a significant level of dissatisfaction with the current controlled English subject heading system»⁴⁴.

⁴⁴ Magda El-Sherbini; Sherab Chen, *An assessment of the need to provide non-Roman subject access to the library online catalog*, «Cataloging & classification quarterly», 49 (2011), n. 6, p. 457-483: 479, DOI: 10.1080/01639374.2011.603108. La tabella con gli esiti delle risposte alla domanda n. 9: *Would you prefer to search subjects in your own (non-Roman scripts) language?* si trova alla p. 473.

Tornando al confronto tra i set di metadati degli OPAC e quelli dei discovery service, la disparità nelle opportunità di accesso alle informazioni sulle risorse bibliografiche disponibili in script diversi dal latino, che a lungo andare rischia di tradursi in un vero e proprio digital divide (paradossalmente, a dispetto degli strumenti – entrambi elettronici – impiegati in fase di ricerca), appare trasversale a numerose biblioteche accademiche e di ricerca in Italia, e sarebbe molto opportuno lavorare in forma cooperativa per risolverla nel modo più efficace. Lungimirante quanto promettente, in proposito, appare l'esito, ancorché non definitivo, del progetto 'DREAM', condotto qualche anno fa da un team di bibliotecari della Sapienza Università di Roma⁴⁵, con il coordinamento scientifico del sinologo Federico Masini, e avente come obiettivo dichiarato quello di «creare un catalogo partecipato con descrizioni in scritture non latine (cinese, arabo, cirillico, giapponese, coreano, ebraico, ecc.)»⁴⁶, per migliorare l'impatto informativo in favore degli utenti madrelingua, o comunque esperti in lingue scritte in alfabeti diversi dal latino. Partendo dagli esiti di un'indagine interna, condotta nell'ambito del progetto stesso, che si è focalizzata sugli OPAC delle più cospicue collezioni delle biblioteche accademiche e di ricerca italiane, i membri del progetto hanno potuto evidenziare che nel nostro paese almeno mezzo milione di risorse documentarie prodotte in alfabeti non latini è stato descritto in forma traslitterata, e verosimilmente almeno un altro milione attende ancora di essere catalogato online. Tale constatazione ha spinto il team progettuale ad affermare, in termini a nostro parere assolutamente condivisibili, che «the extensive transliteration practice in cataloguing has as a direct consequence the underrepresentation of our library collections both at national and international level»⁴⁷. Mossi da tali motivazioni, i membri di DREAM hanno deciso di lavorare alla creazione di un ambiente di lavoro open source, indipendente dai sistemi di gestione bibliografica (ILS) commerciali, in cui ospitare una massa critica di dati catalografici relativi alle collezioni in caratteri non latini provenienti da più biblioteche, con registrazioni redatte secondo le norme, e con l'impiego degli script originali mediante la codifica Unicode UTF-8. Coerentemente a tale scopo, è stato compiuto anche un certo investimento per provare a implementare alcune caratteristiche utili (es. la bidirezionalità delle informazioni nei campi, la gestione degli identificativi VIAF per le voci di autorità, secondo la logica dei linked data) in dialogo con il sistema gestionale impiegato da Sapienza, Sebina Next; e per i dati di autorità è stato progettato anche un apposito modulo di raccolta in formato OAI-PMH. Il repository di informazioni così ottenuto, alimentato mediante strumenti e algoritmi di 'search and match' tra il catalogo di Sapienza e alcuni grandi cataloghi collettivi di biblioteche straniere (es. il catalogo elettronico della BNF, il multiOPAC del sistema bibliotecario bavarese, berlinese e brandeburghese B3Kat, e il SUDOC francese), ha richiesto un notevole sforzo di adattamento e di affinamento in quanto i dati importati – seppure standardizzati – pervenivano pur sempre a DREAM da fonti e con protocolli

45 Il team è costituito da Antonella Fallerini, Agnese Galeffi, Andrea Ribichini, Mario Santanché, Mattia Vallania.

46 L'acronimo DREAM sta per Data Recording Entry Alternative Multi-script. Alcune informazioni generali in merito al progetto, unitamente a una breve relazione tecnica, sono accessibili a <<https://csb.web.uniroma1.it/progetto-dream>>.

47 Antonella Fallerini [et al.], *DREAM. A project about non-Latin script data*, «JLIS.it», 13 (2022), n. 1, p. 347-355: p. 349, DOI: 10.4403/jlis.it-12727. Ringrazio Antonella Fallerini e Agnese Galeffi per gli ulteriori chiarimenti generosamente offerti durante appositi colloqui intercorsi.

MARC diversi. Obiettivo finale, sottolineano le autrici e gli autori, non è di creare un sistema alternativo a SBN ma, al contrario, anticipare la disponibilità di dati bibliografici ‘puliti’ per renderli disponibili, oltre che nell’OPAC di Polo, anche in Indice SBN non appena saranno state risolte le note problematiche tecniche che attualmente inibiscono il multilinguismo del catalogo nazionale.

Un altro percorso di sviluppo molto interessante, articolato a partire da una visione del tutto simile, è in atto presso la rete delle biblioteche romane ecclesiastiche URBE⁴⁸. Tralasciando le tappe dell’evoluzione dei sistemi catalografici in uso presso la rete, tutti costituiti in modo indipendente da SBN, l’aspetto che mi sembra più rilevante per gli scopi di questo contributo consiste nell’innovatività delle pratiche di condivisione dei dati bibliografici (soprattutto per quanto riguarda gli authority data), che ha conseguenze significative anche con riguardo al trattamento delle registrazioni negli alfabeti originali. Lo si può constatare partendo dal più recente approdo degli sviluppi cooperativi di URBE, costituito da Parsifal⁴⁹. Come ha spiegato recentemente il Direttore del Consiglio Direttivo di URBE Fabio Tassone, la novità di questo sistema per l’accesso e lo scambio di dati bibliografici, con i servizi connessi, «è avere prodotto, a partire dalla ‘clusterizzazione’ delle entità agenti, un authority file condiviso da mettere a disposizione dei partecipanti al progetto», in una logica progettuale che «si regge sul presupposto dell’uso degli identificativi provenienti da fonti autorevoli (VIAF, Wikidata, ISNI, ecc.)», seguendo «una visione della cooperazione tra le biblioteche nel contesto delle nuove tecnologie (LOD, BIBFRAME, ecc.)», il tutto «conciliando nel miglior modo possibile le esigenze degli utenti e dei bibliotecari con i presupposti teorici biblioteconomici e bibliografici e questi con le caratteristiche tecniche degli strumenti del progetto»⁵⁰.

48 Nata nel 1991, l’Unione romana delle biblioteche ecclesiastiche (URBE) è formata da 20 istituzioni accademiche accreditate dalla Santa Sede: sei pontificie università, un pontificio ateneo, quattro pontificie facoltà, sette pontifici istituti e un centro. Le collezioni di queste biblioteche (per oltre quattro milioni di documenti complessivi) testimoniano un elevato livello di specializzazione non solo per approfondimenti nelle discipline teologiche nell’ambito del cristianesimo, ma anche per studi negli ambiti dell’ecumenismo e delle religioni non cristiane. Obiettivo principale della rete bibliotecaria è la cooperazione per la gestione dei sistemi informativi e degli strumenti catalografici, che serve un bacino di utenza di oltre diecimila studenti universitari.

49 Il nome dato al progetto allude al carattere incessante della ricerca e della sperimentazione, dal nome di uno dei Cavalieri della Tavola rotonda il quale, com’è noto, si era avvicinato alla scoperta del Sacro Graal, pur senza riuscire a toccarlo. L’homepage del metamotores è accessibile a <<https://parsifal.urbe.it/parsifal/homeSearch>>. L’elenco delle biblioteche partecipanti si trova a <<https://parsifal.urbe.it/parsifal/universities>>. L’organizzazione del portale è spiegata sinteticamente qui: <<https://parsifal.urbe.it/parsifal/clusters>>. Tra i partner del progetto, un ruolo fondamentale ha avuto la società @Cult del Gruppo Casalini Libri.

50 Cito da Fabio Tassone, *Premessa*. In: *Parsifal: un modello di collaborazione bibliotecaria per condividere la conoscenza registrata*, a cura di Silvano Danieli, introduzione di Mauro Guerrini. Firenze: Firenze university press, 2024. Il testo citato è a p. 11-13. La pubblicazione ha raccolto gli atti degli interventi alla giornata inaugurale del catalogo Parsifal, svoltasi presso la Pontificia Università Urbana l’11 maggio 2023.

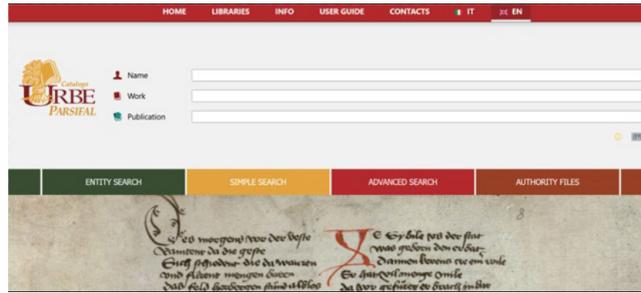


Figura 3 - Maschera di ricerca di PARSIFAL

In concreto, il metamatore Parsifal (fig. 3) offre l'opportunità di interrogare più OPAC di biblioteca, ognuno dei quali è gestito in proprio da una o più biblioteche aderenti, mettendo a valore il know-how dei catalogatori e le caratteristiche peculiari di ciascuna collezione (anche con riferimento alle lingue, scritture, culture che vi sono rappresentate), e di restituire in forma aggregata i risultati bibliografici esito delle ricerche degli utenti, grazie al pieno rispetto dei più recenti e aggiornati standard internazionali in ambito catalografico: RDA, LRM e BIBFRAME. Ne consegue che, nel display dei risultati della ricerca federata, si avrà il pieno rispetto degli alfabeti originali delle pubblicazioni, se così sono state catalogate nell'OPAC di provenienza (proprio l'esito che l'indice SBN al momento non permette di ottenere). Nel caso di Parsifal, il vantaggio di partenza è che diverse biblioteche della rete per anni avevano già condotto prassi operative positivamente improntate alla catalogazione di migliaia di risorse bibliografiche nello script originale⁵¹, come si può vedere dall'esempio random seguente, la copia dell'edizione in arabo del messaggio di Papa Giovanni Paolo II per la Giornata mondiale della pace 2001, posseduta da una delle biblioteche URBE⁵² e descritta come nella fig. 4.



Figura 4 – Esempio di descrizione relativa a un documento in arabo posseduto dalla Biblioteca del PISAI

⁵¹ E ciò anche prima dell'automazione dei cataloghi, con lo sviluppo di cataloghi a schede mobili multialfabeto! Ne danno una curiosa e molto interessante testimonianza Fabio Tassone; Simone D'Ambrosi, *Cataloghi multialfabeto ed esigenze degli utenti. L'esperienza della biblioteca del Pontificio istituto orientale*. In: Nessuno poteva aprire il libro... *Miscellanea di studi e testimonianze per i settant'anni di fr. Silvano Danieli, OSM*, a cura di Mauro Guerrini, con la collaborazione della Pontificia Facoltà teologica 'Marianum' e della rete bibliotecaria URBE. Firenze: Firenze university press, 2019, p. 323-336. Si vedano, in particolare, le schedine cartacee in formato internazionale riprodotte a p. 328.

Inoltre, in Parsifal, l'impiego di identificatori persistenti per le entità del catalogo (autori, opere, pubblicazioni): a) permette il raggruppamento (clustering) di tutte le forme varianti ai fini del recupero corretto delle descrizioni bibliografiche collegate; b) fa funzionare correttamente l'applicazione dei linked open data (LOD) alle entità del catalogo (autori, opere, pubblicazioni); c) può facilitare l'import/export 'pulito' dei dati bibliografici da/verso altri formati (es. MARC21, MARC XML, etc.). Con riguardo all'aggregazione, nelle directory del catalogo, delle forme controllate anche per lingue e alfabeti diversi dal latino, esemplificazioni grafiche molto chiare vengono offerte nella *Guida utente* di PARSIFAL, a cui rimando per ulteriori approfondimenti sull'architettura dello strumento⁵³.

Infine, si segnala che una strategia nuova per cercare di individuare strumenti di ausilio ai catalogatori, anche di quelli con una conoscenza non ottimale delle lingue e scritture diverse dal latino presenti nel portfolio di collezioni delle nostre biblioteche, è stata progettata ed è in corso di sperimentazione nel contesto degli studi religiosi, in particolare nell'ambito del progetto PNRR ITSERR, mirato a sostenere l'infrastruttura europea di ricerca RESILIENCE⁵⁴. ITSERR è articolato in 12 sezioni ('work packages', WP), ognuna con un piano di lavoro e obiettivi specifici, accomunati dal tentativo di costruire strumenti di lavoro nuovi a supporto degli studiosi. Anche nel WP5, denominato "Digital Maktaba"⁵⁵ e dedicato alla catalogazione multialfabeto, il tentativo è quello di sfruttare le nuove frontiere delle digital humanities a supporto degli studi religiosi e dell'attività delle biblioteche, mediante l'impiego di tecnologie di intelligenza artificiale (AI), big data (BD) e calcolo ad alte prestazioni (HCP). Nello specifico, le applicazioni sono rivolte in primo luogo a una cospicua raccolta digitale di documenti scientifici in lingua araba (monografie, periodici, altri materiali documentari in forma digitale, non tutti in forma 'editabile' tale da favorire il riconoscimento immediato dei caratteri), posseduta dalla Biblioteca "Giorgio La Pira" della Fondazione FSCIRE, con l'obiettivo di individuare il workflow più efficace, basato su software di Optical Character Recognition (OCR), strumenti di Natural Language Processing (NLP) e con algoritmi definiti dagli esperti informatici che fanno capo al pro-

52 Si tratta della Biblioteca del Pontificio Istituto di Studi Arabi e d'Islamistica (PISAI), che impiega il gestionale Koha.

53 Si veda ad esempio l'illustrazione, a p. 7 della *Guida*, delle varianti dell'entità 'Alberto Moravia', comprendente anche le forme originali in arabo, in cinese e altre lingue e scritture, collegate a Parsifal tramite VIAF. Si può scaricare la guida all'uso del metamotores qui: <https://parsifal.urbe.it/parsifal/files/Guida_Utente_it_PARSIFAL_rev.pdf>.

54 RESILIENCE <<https://www.resilience-ri.eu/>> RESILIENCE è l'infrastruttura europea per lo studio delle religioni, e ha l'obiettivo di consolidare, mediante l'azione di strutture e ricercatori di eccellenza, la produzione in forme innovative di nuove competenze, conoscenze, metodologie di lavoro capaci di apportare un significativo impatto nel campo scientifico degli studi religiosi. In quanto tale, RESILIENCE è menzionato nella Roadmap 2021 delle infrastrutture di ricerca di dimensione pan-europea (European Strategy Forum on Research Infrastructures, ESFRI) e ha ricevuto finanziamenti dalla Commissione europea per la sua fase preparatoria (2022-2026). Al progetto PNRR ITSERR, che è stato ideato dalla Fondazione FSCIRE (Bologna/Palermo) ed è di durata triennale (novembre 2022-ottobre 2025), partecipano come partner il CNR-ISTI di Pisa e le Università di Palermo, Modena e Reggio Emilia, Torino, e Napoli l'Orientale.

55 Maktaba è un calco dall'arabo, l'espressione sta per "Biblioteca digitale".

getto, per potere estrarre metadati dai documenti digitali e permettere, come obiettivo finale, ai catalogatori di disporre di strumenti di catalogazione semi-automatica dei materiali in script diversi dal latino, il tutto si spera con output il più possibile rispondenti ai modelli e agli standard bibliografici. I primi step di lavoro del WP5 sono stati presentati in diverse sedi alle comunità degli studiosi di scienze dell'informazione e di linguistica computazionale⁵⁶. Un'ulteriore presentazione di Digital Maktaba, rivolta ai bibliotecari, si è tenuta molto recentemente, nell'ambito del convegno nazionale organizzato dal Gruppo di studio dell'AIB sull'inclusione sociale⁵⁷.

Periodici, quotidiani, magazines e scritture originali

Riprendendo, per concludere, il filo del ragionamento sull'equità di servizio e ricollegandoci ai nuovi scenari venutisi a creare con la svolta della 'biblioteca ibrida', andrà aggiunto, a onore del vero, che oggi alcune delle problematiche gestionali ritenute fino a pochi anni fa di ardua soluzione ci appaiono molto più a portata di mano. Un caso lampante è costituito dalle opportunità di acquisizione/distribuzione di quotidiani e magazines in lingue e alfabeti non latini per gli utenti delle nostre biblioteche. Molti 'non-millennials' come me ricorderanno, in proposito, una proposta di intermediazione commerciale, risalente a una ventina di anni fa (e che ritenevamo lungimirante e innovativa, in quel preciso momento), da parte di uno tra i più attenti editori italiani di biblioteconomia e scienze dell'informazione⁵⁸. Bene, nel giro di pochissimi anni, l'avvento delle licenze per l'edicola digitale a disposizione delle biblioteche italiane ha stravolto del tutto (e in positivo) lo scenario precedente, tanto che, giusto per fare un esempio, il mio Ateneo, con un singolo abbonamento e senza porsi alcun problema linguistico ai fini della catalogazione, ormai da diversi anni è in grado di offrire direttamente, tramite la piattaforma MLOL, a tutti i suoi utenti accreditati l'intero pacchetto digitale *Press Reader* consistente in oltre settemila titoli di seriali in modalità streaming, tra i quali (selezionando con un solo clic la lingua) quasi 300 in lingua cinese e poco meno di un centinaio in lingua araba (fig. 5)⁵⁹.

56 Tra gli altri: Sonia Bergamaschi [et al.], *Novel perspectives for the management of multilingual and multialphabetic heritages through automatic knowledge extraction: the DigitalMaktaba approach*, «Sensors», 22 (2022), article number 3995.

57 Al convegno "Niente su di noi, senza di noi. Biblioteche per l'inclusione" (Napoli, Biblioteca universitaria, 21-22 novembre 2024), all'interno della sessione dedicata ai contributi selezionati dal Comitato scientifico a seguito di 'call for papers', sono intervenuti Federico Ruozi e Domenico Ciccarello, con la comunicazione: *Digital Maktaba. Rafforzare l'inclusività nelle biblioteche attraverso l'estrazione automatica di conoscenza dalle collezioni bibliografiche digitali multi-alfabeto*.

58 Ne ha trattato Barbara Fiorentini, *I giornali degli stranieri. Alcune indicazioni per le biblioteche multiculturali*, «Bibliotime», 8 (2005), n. 2, <<https://www.aib.it/aib/sezioni/emr/bibtime/num-viii-2/fiorenti.htm>>.

59 Per una descrizione del servizio offerto dall'Università degli studi di Palermo, si veda il portale del SBA a <<https://www.unipa.it/biblioteche/collezioni/quotidiani-magazine/>>.

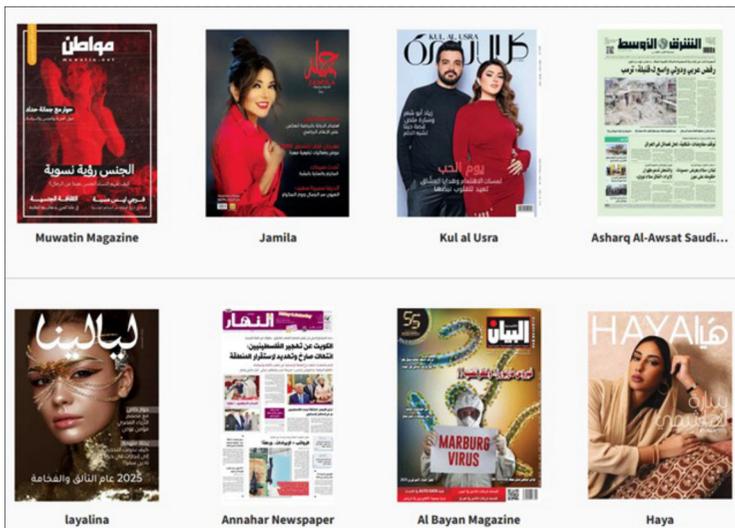


Figura 5 – Edicola digitale Unipa. Una selezione di magazines e quotidiani in lingua araba

Si potrebbe discutere a lungo, anche in questo caso, sulla condizione di disallineamento che si viene a creare tra accesso alla collezione digitale e accesso ai libri fisici tramite le descrizioni nel catalogo elettronico di Ateneo; fatto sta che il servizio c'è, funziona, vi si accede facilmente (nonostante i titoli vengano presentati nella piattaforma in forma traslitterata), e pertanto con una buona e costante attività di reference verso il target maggiormente interessato, l'edicola digitale può essere sempre meglio conosciuta e apprezzata.

Va aggiunto che, fortunatamente, per questa tipologia di materiale (e prevalentemente per i periodici di tipo scientifico), complessivamente il *digital divide* appare un po' più circoscritto, al contrario le opportunità di ricerca dei periodici in alfabeti non latini nei nostri cataloghi si sono allargate considerevolmente. Tale affermazione, ancora prematura con riguardo all'ambiente di lavoro SBN (per il quale valgono le stesse limitazioni tecniche nella trasposizione dei dati catalografici già osservate per le monografie a stampa), risulta ormai pienamente valida con riferimento al repertorio preferito dalle biblioteche accademiche e di ricerca, e cioè il Catalogo italiano dei periodici ACNP, attualmente gestito dall'Area biblioteche e Servizi allo studio dell'Università di Bologna insieme alla Biblioteca centrale "Guglielmo Marconi" del CNR di Roma, con la partecipazione di un nutrito Comitato di biblioteche partecipanti, che con la loro quota di sottoscrizione contribuiscono a rendere finanziariamente sostenibile lo sviluppo dei servizi.

Nel giugno 2024, i gestori di ACNP hanno dato notizia dell'attivazione, nel sistema di gestione del database, di un nuovo campo denominato "titolo originale", che può essere valorizzato anche con i caratteri non latini. Passata piuttosto in sordina nelle liste di discussione professionali, l'informazione appare invece estremamente rilevante nella prospettiva della catalogazione partecipata e del miglioramento del recupero dell'informazione, dal momento che apre a opportunità di immissione dati e di *discoverability* dei periodici in lingue e scritture diverse di cui, nel panorama catalografico italiano, non si riscontrano precedenti minime paragonabili.

Il funzionamento, sia per i catalogatori che per gli utenti finali, è molto semplice. In fase di descrizione, analogamente a quanto avviene con tutte le altre registrazioni dei seriali, la forma originale del titolo, qualora non sia già presente nel catalogo, può essere derivata dal registro ISSN collegato al gestionale ACNP, insieme con altri elementi della descrizione eventualmente presenti in lingua originale, come l'ente autore. L'attuale versione del *Manuale ACNP*⁶⁰ ha già recepito le novità, segnalando le opportune prescrizioni nella sezione delle regole catalografiche⁶¹. Man mano che le biblioteche partecipanti ad ACNP proseguiranno la loro attività di popolamento delle notizie dei periodici in lingua originale nel database, queste saranno utilizzabili come chiave di accesso nella maschera di interrogazione dell'OPAC. Alternativamente, si può cercare il titolo con l'opzione, ugualmente disponibile, della forma traslitterata, che sarà sempre presente. Questo significa che già adesso l'interrogazione per titolo in ACNP consente di recuperare, ad esempio, una rivista irachena di lunga tradizione dedicata ai beni culturali, che in Italia risulta posseduta solo dalla Biblioteca dell'Accademia Nazionale dei Lincei e Corsiniana, sia ricercandola nella forma originale in arabo المورد (= La risorsa) che nella forma traslitterata 'Al-Mawrid', come nell'esempio mostrato nella fig. 6:

Field	Value
Titolo parallelo	المورد, Al-Mawrid
Altri titoli	Quarterly journal of culture and heritage
Altra Forma dell'Ente	دار التحرير الثقافية - وزارة الإعلام, وزارة الإعلام, وزارة الثقافة
Luogo - Editore	Dār al-tū'ūn al-taḥqīfīyāt al-'āmma / Baghdad
Da anno - Ad anno	1971-
Natura	Periodico
Periodicità	Trimestrale
Lingua	Arabo
Paese di pubblicazione	Iraq
ISSN	1606-4836
ISSN L	1606-4836
Codice CDU	008(567)
Codice rivista	P 00160516
Supporto	Printed text
Proc. cumulativo Acnp	1971-

Figura 6 - OPAC di ACNP. Una risorsa in lingua araba visualizzata a seguito di ricerca tramite il titolo originale

60 Università di Bologna. Area Patrimonio Culturale; Biblioteca Centrale "G. Marconi" [del] Consiglio Nazionale delle Ricerche, *Manuale per la gestione del catalogo italiano dei periodici*, scaricabile dall'area del portale dell'Università di Bologna dedicata alla documentazione su ACNP: <<https://acnp.sba.unibo.it/corsi-e-documentazione/documentazione>>.

61 Si vedano le indicazioni del paragrafo 8 (*Titolo originale*): «In questo campo si riporta il titolo nei caratteri originali quando non latini. Da compilare se il titolo in caratteri originali non è già stato riportato nel campo Titolo. Eventuali complementi in caratteri originali non si riportano» (p. 17); e del paragrafo 9 (*Ente autore*): «Il campo contiene la forma dell'ente autore normalizzato secondo le indicazioni delle REICAT [...] È possibile inserire la forma dell'ente autore nei caratteri originali quando non latini» (ivi). Essendo il campo ripetibile, si prescrive ulteriormente: «Se l'ente autore è dato in altri alfabeti o sistemi di scrittura riportare anche la forma traslitterata» (ivi, nota 11).

L'immagine dell'OPAC di ACNP, in questo caso, mostra anche diverse forme varianti del nome dell'editore, che ne arricchiscono le potenzialità di recupero dei titoli collegati tramite la query nella scrittura nativa.

Conclusioni

La sfida per la costruzione di una società multilingue e multiculturale basata su principi di parità di condizioni di accesso alla conoscenza registrata passa per la nostra capacità, come bibliotecari, di realizzare strumenti di mediazione linguistica efficaci e moderni⁶². Un catalogo elettronico in cui non siano stati messi a punto, sul piano delle soluzioni tecnologiche, i prerequisiti funzionali a garantire l'equità di *discoverability* delle risorse documentarie indipendentemente dalla lingua e dalla scrittura originale (un OPAC cioè che non permetta a ciascun utente di recuperare i documenti fisici della biblioteca nella loro forma linguistica nativa, ricercandoli in catalogo con i 'query terms' della medesima lingua e/o scrittura, come si fa ormai ordinariamente con le piattaforme editoriali delle collezioni digitali native), non risponde ancora pienamente alle aspettative degli utenti e rischia di inibire notevolmente, seppure senza alcuna intenzione da parte delle istituzioni che lo sviluppano, l'accesso a una parte selettiva della conoscenza registrata frutto del pensiero umano. Mentre si perfezionano le modalità idonee a garantire forme di accesso 'artificiali' ma pur sempre utili per vari aspetti delle attività di mediazione catalografica e informativa (le operazioni di traslitterazione), occorre lavorare di più e meglio per migliorare l'architettura degli strumenti di authority control, a tutti i livelli, ma soprattutto nei sistemi bibliografici nazionali che dovrebbero costituire l'architrave della struttura del servizio bibliotecario⁶³. Il nuovo modello per l'architettura dei dati bibliografici BIBFRAME e lo standard RDA⁶⁴, l'evoluzione delle mappature dei formati, degli standard di codifica dei caratteri e dei software di gestione informativa delle biblioteche, insieme allo sviluppo di più moderne, ricche e funzionali (per quanto apparentemente più complesse) architetture logiche dei database relazionali per il web (i linked open data, il web semantico)⁶⁵ hanno prodotto una cassetta degli attrezzi profondamente rinnovata, formando un set di mattoncini e di sostanze adesive 'reversibili' con cui è possibile (se ne vedono ampie testimonianze all'orizzonte nella comunità LIS) costruire nuovi edifici bibliografici, forse più sostenibili, certamente più interoperabili rispetto al passato. E chissà che, un giorno neppure troppo lontano, l'oracolo non riesca pure a dare una mano d'aiuto all'architetto⁶⁶.

62 Sull'argomento, in generale, si leggano le osservazioni di Pat Riva, *The multilingual challenge in bibliographic description and access*, «JLIS.it», 13 (2022), n. 1, p. 86-98, DOI: 10.4403/jlis.it-12737.

63 Un interessante (e anche abbastanza recente) caso di studio, è quello israeliano. In proposito, si veda: Ahava Cohen, *Luck is what happens when preparation meets opportunity: building Israel's multilingual, multiscript authority database*, «Cataloging & classification quarterly», 58 (2020), n. 7, p. 632-650, DOI: 10.1080/01639374.2020.

64 Mauro Guerrini; Laura Manzoni, *RDA: resource description and access*, nuova ed. Roma: Associazione italiana biblioteche, 2022.

65 Antonella Iacono, *Linked data*. Roma: Associazione italiana biblioteche, 2014; Laura Manzoni, *Web semantico*. Roma: Associazione italiana biblioteche, 2023.

66 Alludiamo ovviamente alla dialettica tra controllo umano delle forme di comunicazione e capacità generative dell'intelligenza artificiale, sottesa nel titolo della fortunatissima pubblicazione di Gino Roncaglia, *L'architetto e l'oracolo. Forme del sapere digitale da Wikipedia a ChatGPT*. Bari; Roma: Laterza, 2023.

Articolo proposto il 7 febbraio 2025 e accettato il 26 febbraio 2025.

ABSTRACT AIB studi, vol. 64 n. 3 (settembre/dicembre 2024), p. 353-372. DOI 10.2426/aibstudi-14130
ISSN: 2280-9112, E-ISSN: 2239-6152 - Copyright © 2024 Domenico Ciccarello

DOMENICO CICCARELLO, Università degli studi di Palermo, Dipartimento di Culture e società, e-mail: domenico.ciccarello@unipa.it.

Trattamento delle risorse bibliografiche in alfabeti non latini e discoverability delle collezioni digitali: problemi e prospettive

La catalogazione dei materiali in caratteri non latini, argomento discusso in Italia sin dai primi anni del 2000, sta tornando a far parte del dibattito professionale. In termini biblioteconomici, sembra opportuno esaminare la questione nella prospettiva della 'decolonizzazione' degli strumenti di catalogazione, cioè nel pieno rispetto della diversità linguistica e culturale nelle attività di rappresentazione, accesso e visualizzazione dei dati bibliografici. Su un piano più tecnico, l'implementazione dello standard Unicode nei software di gestione delle biblioteche si è rivelata una soluzione valida per la corretta codifica degli script nativi. Tuttavia, alcune questioni critiche sono ancora associate ai limiti che l'Indice del catalogo SBN ha stabilito nel modo in cui le informazioni bibliografiche vengono inserite nel sistema bibliografico nazionale e quindi offerte agli utenti finali. Fino a quando la traslitterazione rimarrà l'unico modo per descrivere nei cataloghi elettronici le risorse in caratteri non latini, gli utenti potranno sfruttare i vantaggi derivanti dalle opzioni di interrogazione nella lingua originale dei testi solo quando effettuano la ricerca nei discovery tool. Questi ultimi, infatti, in aggiunta ai materiali fisici della biblioteca interrogano anche le risorse digitali native (per le quali i metadati negli script originali sono già stati direttamente ricavati dalle piattaforme digitali degli editori). Ciò crea una certa disuguaglianza nella possibilità di recuperare i documenti cercati, e l'esigenza di accedere in modo agevole e completo alle risorse bibliografiche multilingue rimane insoddisfatta.

Processing of bibliographic resources in non-Latin alphabets and discoverability of digital collections: problems and perspectives

The cataloguing of resources in non-Roman scripts, a topic discussed in Italy since the early 2000s, has regained prominence in professional discourse. From a library theory perspective, this issue warrants examination through the lens of "decolonizing" cataloguing tools to ensure full respect for linguistic and cultural diversity in the representation, access, and display of bibliographic data.

On the technical side, the implementation of Unicode standards in library systems has offered a viable solution for correctly encoding native scripts. However, critical issues persist due to limitations imposed by the SBN (Servizio Bibliotecario Nazionale) catalogue Index regarding how bibliographic information is entered into the national library system and subsequently presented to end users.

As long as transliteration remains the only method for describing resources in non-Roman scripts within OPACs, users can only leverage original-language query options through discovery tools. These tools, unlike traditional catalogues, handle born-digital resources with metadata in original scripts harvested from publishers' digital platforms. This discrepancy creates inequalities in resource discoverability, leaving the need for comprehensive and straightforward access to multilingual library resources unfulfilled.