

# Between Intertextuality and Paratextuality: Gérard Genette teacher of method

### Valentina Sestini(a)

a) Sapienza, Università di Roma

Contact: Valentina Sestini, <u>valentina.sestini@uniroma1.it</u>
Received: 06 February 2023; Accepted: 27 February 2023; First Published: 15 May 2023

#### **ABSTRACT**

Tracing the birth and development of the concept of paratext in Gérard Genette (1930-2018), the contribution underlines the effectiveness of Genetti's investigation which has become, over the years, a multidisciplinary methodological approach.

#### KEYWORDS

Gérard Genette; Paratextuality; Intertextuality; Methodological approach.

# Tra intertestualità e paratestualità: Gérard Genette maestro di metodo

#### ABSTRACT

Nel ripercorrere la nascita e lo sviluppo del concetto di paratesto in Gérard Genette (1930-2018), il contributo rimarca l'efficacia dell'indagine genettiana divenuta, nel corso degli anni, un approccio metodologico multidisciplinare.

#### PAROLE CHIAVE

Gérard Genette; Paratestualità; Intertestualità; Approccio metodologico.

<sup>© 2023,</sup> The Author(s). This is an open access article, free of all copyright, that anyone can freely read, download, copy, distribute, print, search, or link to the full texts or use them for any other lawful purpose. This article is made available under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. JLIS.it is a journal of the SAGAS Department, University of Florence, Italy, published by EUM, Edizioni Università di Macerata, Italy, and FUP, Firenze University Press, Italy.



Il pensatore non è creatore. È veggente. Roger Munier

## 1. La nascita del concetto di paratesto

Critico letterario, filosofo, saggista, esponente di spicco dello strutturalismo francese insieme a Roland Barthes e Claude Lévi-Strauss, Gérard Genette (1930-2018)¹ ebbe, tra i molti meriti scientifici, anche quello di definire il paradigma teorico della paratestualità, dimensione metodologica poi applicata in diversi comparti speculativi.

Il concetto fu frutto di una lunga gestazione, iniziata alla fine degli anni '70 in una forma ancora molto embrionale nel volume del critico Introduction à l'architexte (Genette 1979), sviluppata poi nei suoi Palimpsestes. La littérature au second degré (Genette 1982) e terminata, come noto, nel saggio interamente dedicato alla tematica dal titolo Seuils (Genette 1987). D'altra parte, lo stesso Genette avrebbe più volte dichiarato che ogni suo nuovo libro era nato da uno spunto contenuto in uno precedente, come peraltro dimostrano le numerose concatenazioni della sua produzione scientifica. Dallo sviluppo di alcuni dettagli contenuti nell'Architesto derivò, quindi, Palinsesti e, analogamente, da una costola di quest'ultimo nacque Soglie, dove l'autore offrì un'analisi completa e sistematica della paratestualità letteraria, divisa in peritestualità (copertina, avvisi, dediche, indici, ecc., che dipendono dalla responsabilità dell'editore o, più astrattamente, dell'edizione) ed epitestualità (testi che non si trovano annessi al testo nello stesso volume, ma che circolino in qualche modo in libertà, in uno spazio fisico e sociale virtualmente illimitato). Il punto di partenza fu, come detto, l'Introduction à l'architexte, pubblicato in Francia alla fine degli anni '70 e tradotto in Italia nel 1981, dove Genette, esaminando i generi letterari e partendo dalla *Poetica* di Aristotele, introdusse il concetto di transtestualità, definendola come ciò che trascende dal testo e che include, quindi, anche il paratesto (termine, quest'ultimo, che però non viene citato):

Ma sta di fatto che *per il momento* il testo (non) mi interessa (che) per la sua *trascendenza testuale*, cioè tutto ciò che lo mette in relazione, manifesta o segreta, con altri testi. La chiamo *transtestualità* e vi includo l'*intertestualità* in senso stretto [...] vale a dire la presenza letterale (più o meno letterale, integrale o no) di un testo in un altro: la citazione, vale a dire la convocazione esplicita di un testo presentato e distanziato allo stesso tempo con delle virgolette, è l'esempio più evidente di questo tipo di funzioni, che ne comporta altri ancora (Genette 1981, 69-70).

La transtestualità nasceva dall'idea che nessun testo fosse concepibile nella sua unicità, poiché ogni testo – in effetti – contiene altri testi, fa riferimento ad altri testi, nasce da altri testi e produce altri testi.<sup>2</sup> Alla base vi era chiaramente la formazione strutturalista di Genette e l'applicazione di alcune categorie strutturaliste alla teoria letteraria. La stessa "intertestualità" era stata, infatti, già introdotta, nella critica letteraria, dalla psicoanalista francese di origine bulgara Julia Kristeva, nel 1978, in un

<sup>1</sup> Per un approfondimento sulla figura di Genette si leggano: (Montalbetti 1998); (Bollino 2006); (Turgeon 2018); (Ballerio e Pennacchio 2021).

<sup>2</sup> Il rimando è a Michel Foucault e alla sua definizione di testo come un "nodo di un reticolo" e "meccanismo di rimandi". Cfr. (Foucault [1969] 1971).



significato parzialmente diverso da quello attuale. Esso derivava, a sua volta, dalla lettura di un saggio di Michail Bachtin (*Dostoevskij. Poetica e stilistica*), dove l'autore aveva evidenziato che spesso, nella stessa pagina o addirittura nella stessa frase, si verifica la compresenza di due diverse "voci", a loro volta rappresentanti due diversi punti di vista, che dialogano implicitamente tra loro e che arricchiscono il tessuto semantico dell'opera. Bachtin chiamò questo fenomeno, che è proprio della lingua, 'plurivocità': il fenomeno analogo, ma applicato ai testi scritti, divenne poi l'intertestualità (Bachtin 1968). Sul finire degli anni settanta a questa riflessione si intrecciò la cosiddetta estetica della ricezione, che spostò l'attenzione dal testo – e dall'autore – al lettore. Nel suo volume *Lector in fabula*, Umberto Eco infatti sostenne che il testo è "una macchina pigra", che per liberare la sua ricchezza di significati necessita dell'attivazione di un lettore (Eco 1979). Quanto più il lettore è competente – cioè possiede una ricca biblioteca di altri testi e una ricca enciclopedia di conoscenze – tanto più il testo potrà manifestare la sua ricchezza intertestuale, potrà cioè rivelare la quantità di testi che sono, più o meno consapevolmente, dietro alla sua creazione da parte dell'autore.

In un certo senso, quindi, Genette non inventò il concetto di intertestualità e transtestualità, ma lo ampliò, declinandolo in varie forme e dando vita a nuovo approccio metodologico.

Il concetto, come dicevamo, solamente anticipato nelle ultime pagine del volumetto dedicato all'architesto, venne poi ripreso e sviluppati nei suoi *Palinsesti. La letteratura al secondo grado* dove Genette per la prima volta usò il termine 'paratestualità', affermando che:

Oggetto di questo studio è chiamato altrove, "in mancanza di meglio", *paratestualità*. Da allora ho trovato definizioni migliori – o peggiori, si giudicherà -, e ho utilizzato "paratestualità" per designare tutt'altra cosa. L'insieme di questo ardito programma è quindi da riprendere. Riprendiamo, dunque. Oggetto della poetica, dicevo pressappoco, non è il testo considerato nella sua singolarità (di questo si occupa piuttosto la critica), bensì l'*architesto* o, se preferiamo, l'architestualità del testo (più comunemente, ed è un po' la stessa cosa, "la letterarietà della letteratura"), cioè l'insieme delle categorie generali o trascendenti – tipi di discorso, modi d'enunciazione, generi letterari, ecc. – cui appartiene ogni singolo testo (Genette 1997, 3).

Nel volume dedicato ai palinsesti il critico spiegò che questa trascendenza testuale o transtestualità, si può dividere in cinque tipi di relazioni diverse, ossia, intertestualità, transtestualità, paratestualità, metatestualità e architestualità, sottolinenado che:

Il secondo tipo di transtestualità è costituito dalla relazione, generalmente meno esplicita e più distante, che nell'insieme formato dall'opera letteraria il testo propriamente detto mantiene con quanto non può essere definito che il suo *paratesto*: titolo, sottotitolo, intertitoli, prefazioni, postfazioni, avvertenze, premesse, ecc.; note a margine, a piè di pagina, note finali; epigrafi; illustrazioni; *prière d'insérer*, fascetta, sovraccoperta, e molti altri tipi di segnali accessori, autografi o allografi, che procurano al testo una cornice (variabile) e talvolta un commento, ufficiale o ufficioso, di cui il lettore più purista e meno portato all'erudizione esteriore non può sempre disporre come vorrebbe e domanda. Non intendo qui avviare o anticipare lo studio, forse a venire, di questo campo di relazioni, che avremo del resto parecchie occasioni di incontrare e che può essere considerato uno dei luoghi privilegiati della dimensione pragmatica dell'opera, vale a dire della sua azione sul lettore – luogo particolare di quanto, dopo gli studi di Philippe Lejeune sull'autobiografia, si è soliti chiamare *contratto* (o *patto*) generico" (Genette 1997, 5).



Nell'ultima riflessione, il critico faceva riferimento – come noto – al famoso saggio di Lejeune (Lejeune 1975), pubblicato nella stessa collana editoriale dei volumi di Genette (Edizioni Seuil), dove l'autore aveva definito l'autobiografia come un genere in cui autore, narratore e protagonista coincidono, sottolineando che nell'autobiografia l'autore è legato al lettore da un "patto autobiografico", che lo impegna non soltanto a dichiarare "fattualmente la verità", ma anche a cercare di comunicare il senso della propria esistenza.

I concetti riguardanti la paratestualità, anticipati come già detto in forma embrionale nell'*Introduzione all'architesto*, e poi ripresi e ampliati nei *Palinsesti*, trovano una loro identità finale e definita nel volume *Seuils* (*Soglie*), che in certo senso rappresenta la sintesi hegeliana delle opere precedenti. Attraverso il paratesto, infatti, si manifestano tutte le altre categorie transtestuali citate da Genette, nonché una serie di informazioni che permettono il riconoscimento e la corretta ricezione del testo. La funzione svolta dal paratesto è determinante, perché introduce il lettore all'interno del testo orientando la sua lettura:

L'opera letteraria è, interamente o essenzialmente, costituita da un testo, vale a dire (definizione minima) da una serie più o meno lunga di enunciati verbali più o meno provvisti di significato. Questo testo, però, si presenta raramente nella sua nudità, senza il rinforzo e l'accompagnamento di un certo numero di produzioni, esse stesse verbali o non verbali, come un nome d'autore, un titolo, una prefazione, delle illustrazioni, delle quali non sempre è chiaro se debbano essere considerate o meno come appartenenti ad esso, ma che comunque lo contornano e lo prolungano, per *presentarlo*, appunto, nel senso corrente del termine, ma anche nel suo senso più forte: per *renderlo presente*, per assicurare la sua presenza nel mondo, la sua «ricezione» e il suo consumo, in forma, oggi almeno, di libro. Questo accompagnamento, d'ampiezza e modalità variabili, costituisce ciò che ho battezzato altrove, conformemente al senso spesso ambiguo di questo prefisso in francese – pensate, dicevo, ad aggettivi come "parafiscale" o "paramilitare" -, il *paratesto* dell'opera. È attraverso il paratesto dunque, che il testo diventa libro e in quanto tale si propone ai suoi lettori e, in genere, al pubblico (Genette 1989, 3-4).

Più che di un limite o di una frontiera assoluta, per Genette il concetto di paratesto è dunque quello di una *soglia*, un "confine tra un dentro e un fuori, un luogo privilegiato, con il compito, più o meno ben compreso e realizzato, di far meglio accogliere il testo e di sviluppare una lettura più pertinente agli occhi" (Genette 1989, 4). Memorabile l'esempio apportato dal critico: ridotto al suo solo testo e senza alcuna istruzione per l'uso, come leggeremmo l'*Ulysses* di Joyce se non si intitolasse *Ulysses*?

Il paratesto quindi non solo come una zona di 'transizione', ma anche di 'transazione', di mediazione tra il testo e il lettore, che ha il compito di orientare la ricezione del testo e di svilupparne una lettura più appropriata.

Di fronte a un oggetto teorico che si presentava alquanto complesso e sfuggente Genette procedette con grande ordine e cautela, con gli strumenti tipici dei grandi maestri. Innanzitutto concentrò la sua attenzione sul paratesto verbale escludendo materiali iconici (ad esempio, le illustrazioni), nonché lo sterminato campo del "paratesto fattuale", cioè tutti i fatti che, noti al lettore, possono condizionare la sua ricezione del testo (l'età o il sesso dell'autore, la sua eventuale appartenenza a un'accademia, l'ottenimento di premi ecc.). In secondo luogo, i vari enunciati del paratesto, e in particolare del peri-



testo (a cui sono dedicati 11 capitoli sui 13 totali) vengono da lui sottoposti a un'analisi rigida: dei vari elementi paratestuali (titoli, dediche, avvisi ecc.) Genette definisce nell'ordine, prima i connotati spaziali (l'ubicazione materiale all'interno del libro), poi quelli temporali (il momento della produzione, rispetto alla produzione del testo), ed infine quelli pragmatici e funzionali (identità del destinatore e del destinatario, intenzionalità dell'enunciato e sua efficacia) (Barenghi 1989, 722-726). Dunque uno studio che spicca per grande chiarezza d'impianto e rigore di metodo.

## 2. L'eredità genettiana negli studi del libro (in Italia)

A distanza di più di trent'anni dall'apparizione di *Soglie*,<sup>3</sup> l'approfondimento speculativo sulle indagini paratestuali e le ricognizioni sulle varie componenti che presentano al lettore ora il singolo libro, ora un gruppo di documenti scritti, si sono evoluti sensibilmente. Impossibile ricordare tutte le numerose iniziative e i pregevoli contributi che hanno avuto come oggetto una delle tante 'soglie' (dai titoli alle copertine, dalle prefazioni agli indici, ecc.) che assolvono funzioni e ruoli particolari, a volte addirittura affrancandosi dai messaggi semantici che contengono.

Indubbiamente tra le componenti paratestuali che hanno suscitato, negli anni, maggior interesse figura sopra a tutte la dedica, nella sua doppia natura di dedica d'opera e dedica d'esemplare. La potenzialità di questa zona paratestuale, in grado di aggiungere ulteriori e preziosi elementi per ricostruire la genesi di un'opera o per enucleare informazioni inedite sugli autori, era molto chiara al critico francese che, dopo aver definito la differenza tra dedicante e dedicatario (Genette 1989, 115-140), precisa:

malgrado la funzione direttamente economica della dedica sia oggi scomparsa, il suo ruolo di patrocinio o di cauzione morale, intellettuale o estetica si è essenzialmente preservato: non si può, alla soglia o alla fine di un'opera, menzionare una persona o una cosa in quanto destinatario privilegiato senza in qualche modo invocarlo, come l'aedo invocava la musa (...) e dunque implicarlo come una sorta di ispiratore ideale (Genette 1989, 133).

Nel citare la funzione di patrocinio o di cauzione, morale intellettuale o estetica, che la dedica mette in atto, Genette aprì la strada, forse inconsapevolmente, a una serie di indagini condotte da studiosi in ambiti disciplinari diversi.

Alle dediche, nel corso degli anni, sono stati rivolti infatti moltissimi studi interpretativi, che hanno collegato le caratteristiche retoriche e formali di tali pratiche ai vari contesti storici, culturali o politici in cui esse furono prodotte.<sup>4</sup> E, d'altra parte, la dedica era stata oggetto di attenzione già

\_

<sup>&</sup>lt;sup>3</sup> Sul trentennale di *Soglie* si legga il fascicolo monografico 23 (2019) di "Interférences littéraires. Literaire interferenties: Multilingual e-Journal for Literary Studies", dal titolo *Seuils, Paratexts. Trente Ans Après*, a cura di Guido Mattia Gallerani, Maria Chiara Gnocchi, Donata Meneghelli e Paolo Tinti, che ospita i contributi tenutisi nelle due giornate di studio *Attention au paratext! Seuils trent'anni dopo* (Bologna, 15-16 febbraio 2018).

<sup>&</sup>lt;sup>4</sup> La dedica ha costituito oggetto di interesse in molti progetti anche d'oltralpe, come attesta l'archivio informatico della dedica "Margini", coordinato da Maria Antonietta Terzoli dell'Università di Basilea, nato allo scopo (nell'agosto del 2002) di studiare e catalogare i testi di dedica della tradizione italiana. Finanziato dal Fondo Nazionale Svizzero per la Ricerca, l'archivio è consultabile online al sito: <a href="https://www.margini.unibas.ch">https://www.margini.unibas.ch</a>. Il progetto è nato in seno al convegno *I margini del libro* (Basilea, 21-23 novembre 2022), di cui furono in seguito pubblicati anche gli atti. Cfr. (Terzioli 2004).



prima degli studi genettiani. Si pensi, per quanto riguarda l'Italia, al noto volume di Carlo Dionisotti su Aldo Manuzio editore (Dionisotti 1975), in cui l'insigne filologo e critico della letteratura italiana ripercorse – come scrisse Barberi – la "vivace, complessa e tuttavia armonica personalità intellettuale di Aldo" (Barberi 1977, 187) proprio attraverso l'analisi delle zone paratestuali delle sue edizioni. Doveroso ricordare che lo stesso Barberi, alcuni anni prima del libro di Dionisotti, aveva dato alle stampe, per i tipi dello stesso editore, un bellissimo volume sul frontespizio nel libro antico (Barberi 1969), e qualche anno dopo, il catalogo si sarebbe arricchito di un fondamentale studio, questa volta sulle prefazioni dei prototipografi Sweynheym e Pannartz di Giovanni Andrea Bussi, curato da Massimo Miglio (Bussi 1978), a testimonianza di un interesse 'antico' per le zone marginali dei libri.

Sicuramente lo studio delle dediche è diventato uno dei capitoli culturali di maggior interesse nelle indagini paratestuali, forse perché considerate preziose componenti "per la non semplice decodifica delle logiche sottese alle scelte editoriali" (Santoro 2002, 60). E non è sicuramente un caso che spesso l'intervento censorio, soprattutto nel periodo di antico regime, si sia accanito proprio su tali zone extratestuali, con espurgazioni molto significative.

In Italia, dopo l'apparizione del volume di Genette, l'attenzione critica e bibliofilica nei confronti del paratesto ha trovato un rinnovato impulso soprattutto nel corso degli anni Duemila, anche grazie alla pubblicazione del volume di Cristina Demaria e Riccardo Fedriga, che sotto forma di glossario illustra "il paratesto in ogni sua parte, con un'attenzione particolare al peritesto editoriale, senza però dimenticarsi quello autoriale" (Demaria e Fedriga 2001, 22). Per quanto riguarda le discipline del libro, invece, l'interesse è stato sicuramente destato dalle iniziative e gli studi di Marco Santoro e Maria Gioia Tavoni,<sup>6</sup> cui nel corso degli anni si sono aggiunti contributi di colleghi diversi,<sup>7</sup> molti dei quali ospitati all'interno della rivista internazionale "Paratesto".<sup>8</sup>

Con il rilancio della prospettiva paratestuale in Italia, sono fiorite molteplici indagini che si sono occupate dei 'dintorni' del testo con analisi sempre più puntuali e circostanziate. Si pensi a quelle dedicate agli indici, gli apparati informativi che consentono di orientarsi all'interno del libro grazie alla loro funzione di bussola, ma che in alcuni casi divennero un pretesto per un tentativo

<sup>&</sup>lt;sup>5</sup> Sullo stesso argomento si legga anche (Infelise e Plebani 2016).

<sup>&</sup>lt;sup>6</sup> Doveroso ricordare le iniziative che nacquero all'interno del progetto nazionale Cofin 2003 "Oltre il testo", coordinato da Marco Santoro, che vide coinvolte, oltre all'Università di Roma "La Sapienza", altre cinque unità di ricerca: Università di Bologna (responsabile Maria Gioia Tavoni), Università della Calabria (responsabile Carmela Reale), Università di Genova (responsabile Anna Giulia Cavagna), Università di Messina (responsabile Giuseppe Lipari) e Università di Verona (responsabile Giancarlo Volpato). Fra i molti risultati scientifici del progetto vanno segnalati: il convegno internazionale svoltosi a Roma e a Bologna dal 15 al 19 novembre 2004, (Santoro e Tavoni 2005); la realizzazione della mostra tenutasi presso la Biblioteca Universitaria di Bologna "Sulle tracce del paratesto", il cui catalogo, a cura di Biancastella Antonino, Marco Santoro e Maria Gioia Tavoni, è stato edito dalla Bononia University Press nel 2004; la fondazione, nel 2004, della rivista annuale «Paratesto», da parte di Marco Santoro e Maria Gioia Tavoni (diretta fino al 2016 da Marco Santoro, con la condirezione di Maria Gioia Tavoni e, dal 2017 a oggi, da Rosa Marisa Borraccini e Valentina Sestini). Infine da segnalare la nascita della collana «Biblioteca di "Paratesto"» per Serra Editore.

<sup>&</sup>lt;sup>7</sup> Si ricordino quelli di Marco Paoli, in particolare (Paoli 2004) in cui l'autore, nel ricostruire la condizione dell'autore italiano settecentesco, analizza la prassi consolidata dell'anticipo delle spese tipografiche da recuperare, dopo la stampa, grazie proprio al sistema delle dediche. Ma l'indagine paratestuale si è prestata a molte sollecitazioni esegetiche diverse. Si leggano ad esempio gli studi di Mauro Guerrini e Rossano De Laurentiis sugli FRBR e la paratestualità (De Laurentiis e Guerrini 2005) o quelli di Paola Castellucci sull'ipertesto e il paratesto (Castellucci 2018).

<sup>&</sup>lt;sup>8</sup> Si veda la nota precedente.



di una vera e propria classificazione del sapere, come nel caso della *Table* di Mouchon nell'*En*cyclopedie (Tavoni 2009). O si pensi a una delle soglie più interessanti della paratestualità ossia il frontespizio, strumento per eccellenza di mediazione intellettuale e relazionale tra produttore, committente e fruitore, cui sono stati dedicati innumerevoli approfondimenti scientifici. Con la sua posizione privilegiata ad apertura di libro, il frontespizio rappresenta, infatti, soprattutto nel libro antico, un potente mezzo espressivo per il tipografo, sia a scopo pubblicitario che per la conoscenza e la diffusione del prodotto, come bene evidenziato da molti studiosi (Baldacchini 2004). Impossibile poi non citare l'ampia letteratura scientifica dedicata allo studio delle copertine, che nel processo di ricezione e consumo dei testi giocano – come sappiamo – un ruolo fondamentale. Elemento collante tra l'intentio editoris e l'intentio auctoris, la copertina condiziona infatti la ricezione del testo da parte del lettore finendo per influenzarne il consumo. Una soglia, come evidenziato già da Genette, che costituisce una cerniera tra il momento di produzione e quello della fruizione e che può sviluppare una propria sintassi paratestuale subito riconoscibile (inutile ricordare il caso emblematico del giallo in Italia, dove il colore, inizialmente funzione dell'identità di collana, è poi rapidamente passato all'uso editoriale e all'uso comune per identificare un intero genere).

Lungi dal voler elencare tutti i contributi dedicati all'analisi paratestuale nell'ambito delle discipline del libro, l'intento di queste brevi riflessioni è solo quello di rimarcare l'importanza dell'approccio metodologico teorizzato da Genette, che "può e deve confrontarsi e coniugarsi con altre metodologie ma nella sua organica, convinta e convincente autonomia esegetica" (Santoro 2005, 11).

Qualche studioso ha obiettato che non tutte le opere di Genette siano invecchiate bene, soprattutto se confrontate con quelle del più noto strutturalista francese Roland Barthes. Tra l'altro, la fortuna di Genette in Italia conobbe, diversamente da quanto è accaduto altrove, una fase "piuttosto prolungata di appannamento, dovuta certo all'esaurimento della voga strutturalista, ma fors'anche a un approccio in qualche caso riduttivo, perché, da un lato poco interessato a cogliere quelle sfaccettature della sua riflessione e della personalità di studioso meno coerenti con l'immagine divulgata (...) e, dall'altro, disattento agli sviluppi ulteriori del suo percorso intellettuale" (Bollino 2006, 22), come sottolineò anche Pierluigi Pellini nel bellissimo necrologio apparso sul "Manifesto" all'indomani della scomparsa del critico (Pellini 2018).

In realtà, la grande potenza del pensiero genettiano non ci sembra tanto legata alla fortuna delle sue singole opere, quanto all'essere diventato un indiscutibile riferimento di metodo. Nel 1989, in una recensione a *Soglie* apparsa su "Belfagor", il critico Mario Barenghi affermava che le future indagini sul paratesto non sarebbero potute non partire dal prezioso contributo di Genette e buona parte della terminologia da lui coniata avrebbe trovato sicuramente cittadinanza nel comune linguaggio critico. La previsione di Barenghi ha trovato pieno compimento, come dimostra l'applicazione del metodo paratestuale a molteplici ambiti disciplinari, dalle scienze storiche a quelle sociologiche, dalla critica cinematografica al marketing editoriale, di cui Genette non si è mai occupato personalmente. Il campo paratestuale si è rivelato un terreno incredibilmente proficuo, in cui le conoscenze hanno potuto esaltare le proprie potenzialità esegetiche per la comprensione di ulteriori forme comunicative. In questa prospettiva l'apparato concettuale e terminologico proposto da Genette ha esercitato "compiutamente la sua capacità euristica e la sua efficacia esplicativa" (Barenghi 1989, 726).



Quello che Genette non avrebbe potuto immaginare è che l'informatica avrebbe sparigliato le carte, aprendo nuovi scenari in cui rivedere ruoli e funzioni del paratesto per adattarli alle caratteristiche del nuovo testo digitale. Il paratesto di Genette nasceva, infatti, come elemento della materializzazione del testo, cioè dell'edizione tradizionale in forma di libro. Il venir meno della 'materialità' a favore della nuova dimensione digitale – che, aprendo nuove possibilità di accesso e di gestione dei contenuti, "ha inevitabilmente modificato la natura stessa del testo così com'era concepito nell'era della stampa, cioè come oggetto statico – non ha certo annullato il ruolo del paratesto" (Bonincontro 2001). Gli elementi del paratesto sono cambiati, o meglio se ne sono aggiunti di nuovi, che hanno funzione di guida nella navigazione all'interno del testo, funzione determinante proprio quando il testo è costruito per una fruizione non più, o non principalmente, lineare. L'aspetto grafico del paratesto assume dunque un'importanza notevole, che in alcuni casi potrebbe diventare perfino maggiore del contenuto testuale, soprattutto perché può condizionarne o limitarne l'accesso. Eppure il moltiplicarsi di iniziative di digitalizzazione di testi a stampa, cui si assiste in questi anni, non sembra ancora aver avviato una riflessione seria sul trattamento del paratesto presente nell'edizione tradizionale. Il centro della discussione dovrebbe quindi spostarsi sul quesito se il "paratesto presente in un libro (per ipotesi a stampa) conservi la sua funzione una volta separato dal supporto materiale originario e trasferito su un nuovo canale di trasmissione, quello appunto digitale" (Bonincontro 2001).

Su questo aspetto forse la comunità scientifica dovrebbe continuare a confrontarsi, nella convinzione, come sosteneva Genette, che una soglia debba essere sempre e comunque "attraversata" (Genette 1989, 404).



## Riferimenti bibliografici

Antonino, Biancastella, Marco Santoro, e Maria Gioia Tavoni, a c. di. 2004. *Sulle tracce del paratesto*. Catalogo della Mostra tenuta a Bologna (2004). Bologna: Bononia University press.

Bachtin, Michail Michailovič. 1968. *Dostoevskij. Poetica e stilistica*. Traduzione italiana di Giuseppe Garritano. Torino: Einaudi.

Baldacchini, Lorenzo. 2004. Aspettando il frontespizio: pagine bianche, occhietti e colophon nel libro antico. Milano: Edizioni Sylvestre Bonnard.

Ballerio, Stefano e Filippo Pennacchio, a c. di. 2021. *Il conoscibile nel cuore del mistero: dialoghi su Gérard Genette*. Milano: Ledizioni.

Barberi, Francesco. 1969. Il frontespizio nel libro italiano del Quattrocento e del Cinquecento. Milano: Il Polifilo.

Barberi, Francesco. 1977. "Recensione di Aldo Manuzio editore: dediche, prefazioni, note ai testi." *La Bibliofilia* 79 (2): 187-190.

Barenghi, Mario. 1989. "Recensione di Gérard Genette, Soglie." Belfagor, 30 novembre, 44, 722-726.

Bollino, Fernando. 2006. L'arte in opera: itinerari di Gérard Genette. Bologna: CLUEB.

Bonincontro, Ilaria. 2011. "Il paratesto nella digitalizzazione di edizioni tradizionali." *Margini. Giornale della dedica e altro*, 5. <a href="https://www.margini.unibas.ch/web/rivista/numero\_5/saggi/articolo5/bonincontro.html">https://www.margini.unibas.ch/web/rivista/numero\_5/saggi/articolo5/bonincontro.html</a>.

Bussi, Giovanni Andrea. 1978. *Prefazioni alle edizioni di Sweynheym e Pannartz prototipografi romani*. A cura di Massimo Miglio. Milano: Il Polifilo.

Cadioli, Alberto. 2005. "Il patto editoriale nelle edizioni moderne e contemporanee." In *I dintorni del testo. Approcci alle periferie del libro. Atti del convegno internazionale (Roma, 15-17 novembre 2004; Bologna, 18-19 novembre 2004).* A cura di Marco Santoro e Maria Gioia Tavoni. Vol. II, 663-672. Roma: Edizioni dell'Ateneo.

Castellucci, Paola. 2018. "Dove finisce il racconto? Ipertesto digitale e paratesto strutturalista." *Paratesto*, 15 189-195.

De Laurentiis, Rossano, e Mauro Guerrini. 2005. "FRBR e paratesto." In *I dintorni del testo. Approcci alle periferie del libro. Atti del convegno internazionale (Roma, 15-17 novembre 2004; Bologna, 18-19 novembre 2004).* A cura di Marco Santoro e Maria Gioia Tavoni. Vol. II, 641-653. Roma: Edizioni dell'Ateneo.

Demaria, Cristina, e Riccardo Fedriga, a c. di. 2001. *Il paratesto*. Milano: Edizioni Sylvestre Bonnard.

Dionisotti, Carlo. 1975. Aldo Manuzio editore: dediche, prefazioni, note ai testi. Introduzione di Carlo Dionisotti. Testo latino con traduzione e note a cura di Giovanni Orlandi. Milano: Edizioni «Il Polifilo».

Eco, Umberto. 1979. Lector in fabula: la cooperazione interpretativa nei testi narrativi. Milano: Bompiani.



Foucault, Michel. (1969) 1971. L'archeologia del sapere. Traduzione di Giovanni Bogliolo. Milano: Rizzoli.

Gallerani, Guido Mattia, Maria Chiara Gnocchi, Donata Meneghelli, e Paolo Tinti, a c. di. 2019. "Seuils, Paratexts. Trente Ans Après." *Interférences littéraires. Literaire interferenties: Multilingual e-Journal for Literary Studies* 23.

Genette, Gérard. 1979. Introduction à l'architexte. Paris: Éditions du Seuil.

Genette, Gérard. 1981. *Introduzione all'architesto*. Traduzione italiana di Armando Marchi. Parma: Pratiche Editrice.

Genette, Gérard. 1982. Palimpsestes: la littérature au second degré. Paris: Éditions du Seuil.

Genette, Gérard. 1987. Seuils. Paris: Éditions du Seuil.

Genette, Gérard. 1989. Soglie. I dintorni del testo. A cura di Camilla Maria Cederna. Torino: Einaudi.

Genette, Gérard. 1997. *Palinsesti. La letteratura al secondo grado*. Traduzione italiana di Raffella Novità. Torino: Einaudi.

Infelise, Mario, e Tiziana Plebani, a c. di. 2016. Aldo Manuzio: la voce dell'editore. Prefazioni e dediche. Venezia: Marsilio.

Lejeune, Philippe. (1975) 1986. *Il patto autobiografico*. Traduzione italiana di Franca Santini. Bologna: il Mulino.

Montalbetti, Christine. 1998. Gérard Genette: une poétique ouverte. Paris: Bertrand-Lacoste.

Paoli, Marco. 2004. L'appannato specchio. L'autore e l'editoria italiana nel Settecento. Lucca: Maria Pacini Fazzi editore.

Pellini, Pierluigi. 2018. "L'anima più concreta dello strutturalismo. (in morte di Gérard Genette)." "Alias", Il manifesto, 4 febbraio, 14.

Santoro, Marco, e Maria Gioia Tavoni, a c. di. 2005. I dintorni del testo. Approcci alle periferie del libro. Roma: Edizioni dell'Ateneo.

Santoro, Marco. 2002. "Appunti su caratteristiche e funzioni del paratesto nel libro antico." In *Libri edizioni biblioteche tra Cinque e Seicento. Con un percorso bibliografico*, 51-92. Manziana (Roma): Vecchiarelli.

Santoro, Marco. 2005. "Nulla di più ma neppure nulla di meno: l'indagine paratestuale." In *I dintorni del testo. Approcci alle periferie del libro*. Roma: Edizioni dell'Ateneo, vol. I, 3-13.

Tavoni, Maria Gioia. 2009. Circumnavigare il testo. Gli indici in età moderna. Napoli: Liguori.

Terzoli, Maria Antonietta a c. di. 2004. *I margini del libro. Indagine teorica e storica sui testi di dedica. Atti del Convegno Internazionale di Studi. Basilea, 21-23 novembre 2002.* Roma-Padova: Editrice Antenore.

Turgeon, David. 2018. À propos du style de Genette: essai. Montréal: Le Quartanier, DL.



# The "cabrei" of the Archivio di Stato di Bologna: a complex reading

## Elena Gonnelli(a)

a) University of Florence, <a href="https://orcid.org/0000-0002-6908-8106">https://orcid.org/0000-0002-6908-8106</a>

Contact: Elena Gonnelli, <u>gonnellielena@gmail.com</u>
Received: 2 February 2023; Accepted: 10 February 2023; First Published: 15 May 2023

#### **ABSTRACT**

The term *cabreo* has been given different meanings over the years: what was originally defined as a legal and then economic instrument, has increasingly acquired an aesthetic connotation that made it more like a piece of art. Starting from its definitions, we will try to show how, at the heart of the cabreo representation, there is a complex system of relationships: the patronage, the drafters, the subscribers, and the apparatuses are all elements that add information to the picture. Taking a few significant examples from the documentary heritage of the Archivio di Stato di Bologna, we will try to show how the examination of the graphic artefacts allows us to reconstruct the field of action of public institutions, private individuals and religious corporations, if analysed in its reference framework.

#### **KEYWORDS**

Historical cartography; Cabreo; Archives; Cartographic documentation.

# I cabrei dell'Archivio di Stato di Bologna: una lettura complessa

#### **ABSTRACT**

Al termine cabreo sono stati attribuiti significati diversi nel corso degli anni: in origine strumento giuridico, poi economico, ha sempre più acquistato una forma estetica che lo ha avvicinato all'opera d'arte. Proprio partendo dalle sue definizioni, dunque, si cercherà di stabilire come al centro della rappresentazione cabreistica vi sia un complesso sistema di relazioni: la committenza, gli estensori, il sottoscrittore, gli apparati sono tutti elementi carichi di un portato informativo. Attraverso alcuni esempi, tratti dal patrimonio documentario dell'Archivio di Stato di Bologna, si cercherà di mostrare come l'esame dei reperti grafici permetta di ricostruire il campo d'azione di istituzioni pubbliche, di privati e di corporazioni religiose, se analizzato alla luce del suo contesto di riferimento.

#### PAROLE CHIAVE

Cartografia storica; Cabrei; Archivi; Documentazione cartografica.

<sup>© 2023,</sup> The Author(s). This is an open access article, free of all copyright, that anyone can freely read, download, copy, distribute, print, search, or link to the full texts or use them for any other lawful purpose. This article is made available under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. JLIS.it is a journal of the SAGAS Department, University of Florence, Italy, published by EUM, Edizioni Università di Macerata, Italy, and FUP, Firenze University Press, Italy.



#### Premessa

Fin da tempi antichi, l'appropriazione mentale dello spazio ha costituito per l'uomo una necessità: disegni e vere e proprie mappe, fossero esse tracciate su roccia, tavole di argilla, legno o marmo dimostrano l'importanza di tale pratica che richiedeva attenta osservazione dei luoghi e competenze scientifiche. Se Eratostene e Tolomeo tentarono di restituire lo spazio applicando il sapere astronomico e matematico all'arte del disegno per avvicinare il più possibile il modello alla realtà, con l'evoluzione della cartografia storica si assiste all'uso strumentale in termini di controllo e gestione del territorio (Rombai 2002, 19-24; Dematteis 1990). Si tratta di rappresentazioni che, conservate in archivi e biblioteche, erano frutto di attente osservazioni e precise misurazioni topografiche, ma che ancora oggi ci attraggono particolarmente per il loro aspetto estetico. Sarà importante sottolineare come la cartografia storica in generale sia stata ampiamente utilizzata, soprattutto negli ultimi anni, con risultati positivi ai fini della localizzazione, in molte ricerche archeologiche di insediamenti abbandonati e scomparsi, sulle testimonianze delle attività produttive (Panzeri e Farruggia 2009). Da non dimenticare, inoltre, il possibile recupero di microtoponimi (poderi, piccoli appezzamenti di terreno, boschi, acquitrini ecc.), nonché di termini dialettali geografici. Sarà l'avvento della cartografia geometrica, costruita sulla base dei grandi catasti e dell'attività degli enti centrali di Stato nel corso del XIX secolo, ad eliminare tutti gli aspetti pittorici e gli ornamenti artistici, lavorando sull'utilità della sintesi più che sulla leggibilità del contenuto.

Tentare di collocare in questo complesso panorama ciò che comunemente va sotto il nome di 'cabreo' non è cosa semplice, a partire dalla stessa definizione della materia. Nel catalogo della cabreistica toscana, Ginori Lisci oltre a offrire un forte apparato iconografico che mostra situazioni diverse (come, per esempio, la ricostruzione storica delle principali imposizioni sopra i beni immobili) lavora proprio sulla polisemia del termine (Ginori Lisci 1978, 20). Martirologi (dal nome Marte, quale sinonimo di terra), terrilogi, campioni o campioni di beni (Manno Tolu 1981, 100), quaderno di piante, catasto o catastici (in ambito veneto), effetti, o, genericamente, descrizioni geografiche sono tutti modi per denotare una rappresentazione. La dicitura cabreo proviene dall'ambito religioso e, in particolare, dall'Ordine di Malta che già dalla fine del Cinquecento aveva messo in uso questo nome. Il Dizionario del linguaggio di Giulio Rezasco (1881) rimanda inoltre all'etimologia spagnola, descrivendo il cabreo come lo strumento adottato dai cavalieri gerosolimitani per misurare e descrivere i beni delle singole Commende; sarà poi il Dizionario Enciclopedico Italiano (1955) ad aggiungere l'etimo latino caput breve, capi brevium, confermando, tuttavia, il significato di mappa di beni appartenenti a grandi amministrazioni ecclesiastiche e signorili, arrivando poi per estensione alle raccolte organiche di disegni attinenti a singole proprietà. Le raccolte rilegate di mappe, prospetti di edifici e vedute della stessa misura e datazione, rappresentanti beni urbani e rurali di enti o privati possono avere anche finalità diverse da quelle catastali, proprietarie o patrimoniali: si pensi, solo a titolo di esempio, al famoso codice Rustici. In questo caso, le lunghe scritture di cronache, le relazioni dei viaggi che l'orafo fiorentino fa durante il suo ritorno dalla Terra Santa sono preminenti e spostano il fuoco della produzione grafica verso un'asse differente, quello della memorialistica. Allora, quando effettivamente nasce l'esigenza di produrre e, soprattutto, di conservare questa particolare tipologia documentaria? A che scopo e con quale valore? Allo stesso modo, perché si esaurisce la consuetudine dei cabrei? Solo ponendosi queste domande e cercando risposte al riguardo è possibile apprezzare il valore intrinseco della fonte (cosa dice



quella rappresentazione e come lo fa; che linguaggio usa) e individuarne le connessioni esterne (quali altre fonti possiamo e/o dobbiamo usare per arrivare a un portato informativo completo e quanto più avvicinabile alla realtà storica).

#### Una fonte per lo studio del territorio bolognese

Le varie e diverse zone paesistiche dell'Emilia Romagna, già oggetto di numerosi studi soprattutto rivolti all'evoluzione del suo paesaggio agrario, nonché la presenza dei cavalieri di San Giovanni a Bologna (Boris 2009, 487-507) sono coordinate che consentono di osservare il complesso cabreistico conservato presso l'Archivio di Stato di Bologna, quale felice sintesi di conoscenze agrimensorie e di abilità pittoriche e cartografiche. Nella provincia bolognese la rappresentazione planimetrica delle proprietà agricole cominciò ad essere eseguita, con discreta frequenza, nei secoli XVII e XVIII, soprattutto su commissione di importanti famiglie. In generale scarsi gli esemplari cinquecentini, i cabrei sono quasi sempre volumi di grandi dimensioni (dai 35 ai 50 cm di altezza e dai 25 ai 35 cm di larghezza, fino ad arrivare alle forme di atlante anche di 60 cm x 75 cm); presentano coperte in cuoio coevo (le rilegature cartonate sono più tarde) spesso ornate con bulloni o fermagli quadrangolari metallici. Il numero delle carte è totalmente variabile, mentre il frontespizio è librario, più o meno ornato, recante quasi sempre l'autore: normalmente agrimensore. La figura professionale destinata alla realizzazione dei cabrei si andò specializzando fino a stabilire una consuetudine grafica e tecnica: un vero e proprio mondo minore che si affermava svolgendo quotidianamente la propria attività. L'esistenza di agrimensori e di periti, in una data società, è di per sé un fatto molto significativo, non solo perché indica la formazione di un ceto specializzato nella misura e nella descrizione dei terreni, ma anche perché testimonia una fase inoltrata della proprietà e del possesso fondiario (Varignana 1974, 1-37).¹ Nello specifico il territorio bolognese, esposto a frequenti alluvioni, vedeva la necessità di sviluppare e applicare tutte le possibili tecniche in uso per il monitoraggio e la conseguente disciplina delle acque.

"Accanto all'attività più rilevante dell'ingegneria idraulica e civile, va ricordata l'attività degli agrimensori e dei tecnici relativa alla rilevazione e alla descrizione degli immobili rustici e urbani. Anche questo aspetto della vita culturale è da porsi in relazione con lo sviluppo della vita economica. Il moltiplicarsi degli aspetti giuridici relativi al regolamento delle proprietà e dei possessi, il crescere delle esigenze dell'economia di mercato, l'enuclearsi del precapitalismo, una più moderna disciplina delle proprietà e dei possessi ponevano l'esigenza di descrizioni, sempre più accurate dei beni. Questa esigenza si era manifestata in tutti i casi in cui si voleva dare ordine alle amministrazioni pubbliche e private" (Dal Pane 1974, XV-XVI).

I cabrei in questo sono una fonte utilissima: la rappresentazione delle colture arboree (disegnate o accompagnate da didascalie), le affossature per lo scolo dei canali, le differenti qualità dei terreni che prendono coloriture distinte, la casa colonica (Gambi 1976, 479-504), sono solo alcune delle componenti informative che se ne ricavano. Tutti gli elementi venivano eseguiti a mano dai professionisti del disegno che adottavano una tecnica su carta inchiostrata e, da un certo momento in avanti, colorata ad acquerello per rendere sfumature e prospettive. Essi, inoltre, potevano essere

<sup>1</sup> Per una testimonianza coeva sulla confusione all'interno della categoria professionale cfr. anche (Girri 1767).



commissionati da una corporazione ecclesiastica così come da un ente pubblico o privato (Crini 1993, 361-387). In questo senso, la dichiarazione dell'autore ha una valenza intellettuale tanto quanto probativa: nelle piante possono essere indicate le colture messe a rendita, gli affittuari, i proprietari dei beni confinanti; i cabrei possono contenere lo spoglio sistematico dei livelli, con una descrizione analitica ordinata per "rendatari", i titolari della conduzione, e molte altre informazioni di natura descrittiva e tecnica.

"La carta che si riferisce ad aree di uso agricolo dà maggiore enfasi – quando si esprime con moduli realistici, per lo più a colori – alle costruzioni padronali coi loro parchi, e ai tipi di colture di pregio, come quelle arborate, o agli allestimenti onerosi, come i muri di sostegno per i terrazzi, i regolari reticoli che formano i cavi di scolo, gli argini risicoli, le fosse da macerare la canapa, ecc." (Gambi 1976, 479). Questo significa, per il privato, tenere conto della gestione e della conservazione del proprio patrimonio, ma anche controllare l'evoluzione della proprietà immobiliare regionale, soprattutto per gli anni che conobbero le conseguenze derivanti dalle soppressioni degli enti ecclesiastici ricchi di cospicui patrimoni e, al contempo, l'affermarsi del controllo statuale sulla proprietà fondiaria, attraverso l'attivazione del catasto.

Il Chirografo di Nostro Signore Papa Pio VI col quale si ordina e stabilisce il regolamento della pubblica economia di Bologna diretto per l'esecuzione all'e.mo e rev.mo signor cardinale Ignazio Boncompagni Ludovisi legato di detta città del 25 ottobre 1780, costituisce un momento cardine e di svolta per il sistema di rilevamento e per la messa in pianta dei territori rurali della provincia. L'operazione prevedeva un censimento generale dei terreni che dovevano essere stimati non solo per l'estensione o per le colture in atto, ma anche sulla base delle potenziali rendite. In sintesi, la tassazione, o terratico, non sarebbe più stata commisurata sull'effettivo reddito, bensì su quanto i poderi avrebbero fruttato se condotti secondo criteri razionali. Inevitabilmente, questa maggiore oculatezza che i proprietari dovevano assumere nei confronti delle loro pertinenze si tradusse in rilevazioni più attente, ma soprattutto più scientifiche, a scapito, senza dubbio, della resa estetica (Zangheri 1961). Esiste una differenza grafica tra la produzione su commissione privata, atta dunque a documentare i beni esistenti per predisporre migliorie nel sistema di conduzione, per divisioni ereditarie o per semplice revisione dei possessi, e quella statale volta al censimento e al controllo centrale. Se resta vero che almeno fino alla metà del XVIII secolo non si percepisce ancora una simbolizzazione codificata, ovverosia valida per tutti gli estensori, la consuetudine sopra richiamata permetteva di evidenziare alcune tipologie notevoli e comuni, sottolineate dalla cifra stilistica. Dalle commissioni private, perlopiù familiari, ciò che emerge è la volontà dei disegnatori di osservare un certo canone estetico, lo stesso canone che verrà gradualmente sacrificato a seguito della scientificità dei nuovi rilievi catastali, dove, senza alcuna pretesa decorativa prevarrà la stringatezza grafica.

Le fonti disponibili in Archivio di Stato di Bologna relative alla presenza dei cavalieri di San Giovanni sono numerose e appartengono a diverse tipologie: l'Ufficio dei Memoriali, gli archivi privati di famiglie nobili come i Malvezzi e i Grassi, i fondi di grandi magistrature come il Senato e l'Ambasciatore bolognese, ma soprattutto il cospicuo fondo delle Corporazioni religiose soppresse. È infatti del XVII secolo (1694 -1696) il primo cabreo, con coperta originale in cuoio e costola puntonata a secco dorata, che registrava i beni della Commenda (1694) (Corporazioni



religiose soppresse, fasc. 18/2097). Fatto compilare dal commendatore Sacchetti al notaio Vecello e al pubblico perito agrimensore Egidio Maria Bordoni, esso conteneva un indice, una trascrizione degli atti e un frontespizio che anticipava le tavole grafiche. Un aspetto da sottolineare è come, in origine, il cabreo rappresentasse un protocollo notarile, usato per registrare i titoli giuridici del possesso e dei diritti d'uso della terra. Solo dal XVI secolo, da strumento giuridico si trasformò in strumento economico, ma l'apparato cartografico mantenne comunque l'esigenza del valore legale di prova degli atti di terminazione e di misurazione del possesso fondiario, per questo veniva sottoscritto da un notaio (Briante 2009, 350-351). Del resto, la meticolosità delle norme dell'Ordine imponeva la tenuta di un catasto aggiornato e valido per i vari beni, urbani ed extraurbani, che la Commenda possedeva. Non è casuale, pertanto, la presenza, nelle pagine iniziali degli apparati celebrativi del titolo del beneficio e, nel frontespizio, il segno tabellionale con la firma del notaio delegato che esercitava formali funzioni di controllo e garanzia sull'incaricato agrimensore.

L'indice e sommario del processo sopra efformazione del presente cabreo òsia campione di tutti li Beni spettanti alla Baliale Commenda di Santa Maria del Tempio di Bologna detta della Masone (1741) (Corporazioni religiose soppresse, fasc. 19/2098) è un cabreo in coperta cartonata (oggi restaurata), cronologicamente successivo, sottoscritto dal pubblico perito e agrimensore Antonio Conti di Bologna che firmerà tutte le tavole grafiche a seguire, riprese e aggiornate, rispetto all'esemplare precedente. I momenti documentari e giuridici, da apparato, diventano vera e propria parte testuale che ci introduce al mondo bolognese, tanto che il cabreo si chiude con annotazione di "Memorie diverse" (Corporazioni religiose soppresse, fasc. 19/2098, cc. 102v-103r) riguardanti rogiti avvenuti tra l'anno 1730 e il 1747, quindi anteriori e posteriori alla composizione complessiva dell'esemplare e che dovevano servire da memorandum. È poi molto interessante che, talvolta, per meglio comprendere taluni aspetti della Commenda sia necessario integrare con altri documenti non iconografici, ma manoscritti: dalle visite del commendatore agli inventari dei beni; dagli atti notarili di affitti e compravendite ai privilegi, scritture varie e recapiti. Fosse per scopi patrimoniali o per valenze probative questi volumi venivano redatti ab origine insieme ad altri strumenti: lo spoglio dei beni livellari (che di solito contiene la storia di ogni livello e i passaggi intervenuti nella conduzione); le piante dei beni rilevate dalle mappe di un eventuale catasto (Conti 1966, 199); le stime per livello (volumetti redatti e compilati da appositi periti ingegneri); e così via.

Abbiamo già avuto modo di notare come e perché, a partire dalla metà del Settecento, le rappresentazioni cabreistiche si facciano progressivamente più grossolane, con cartigli meno elaborati nei chiaroscuri che diventano solo finte cornici a variazione cromatica: l'ultimo dei tre cabrei conservati nel fondo dedicato alla Commenda dell'Ordine di Malta ripete la struttura dei precedenti libri, ma in maniera nettamente più sintetica, inserendo perfino il frontespizio nella stessa pagina dell'indice (1767) (Corporazioni religiose soppresse, fasc. 20/2099). Elemento di novità è invece la rappresentazione artistica della scala, disegnata sullo strumento del perito ed esplicitata con le pertiche. La scala è e rimane un elemento costitutivo dell'esemplare: i testimoni bolognesi riportano una misura in pertiche per il terreno e in piedi per gli edifici, ma in generale si può ben dire che nei cabrei le scale hanno un piccolo denominatore, ossia riportano superfici del terreno di dimensioni piuttosto limitate.

Se il primo cabreo di provenienza religiosa risale alla fine del XVII secolo, altri, di origine familiare, si attestano presso l'Archivio di Stato, nel fondo Salina Amorini Bolognini (Dehò 1993). Il registro n. 540, datato 1702-1757 e indicato in inventario come cabreo pur appartenendo alla serie "matrici",



riporta rappresentazioni in dettaglio molto semplici, corredate da strumenti descrittivi e regesti patrimoniali (Matrici di relazioni e piante, 1592-1791, unità 540). I confini vengono tratteggiati con l'indicazione delle misure; i terreni individuati sommariamente con la dicitura lavorativo ovvero incolto; al limite sono i rimandi didascalici (pezzo di vigna, ecc.). Trattandosi di uno strumento utile a tenere memoria delle gestioni patrimoniali, sotto, a lato o comunque intorno alle rappresentazioni grafiche vengono trascritti gli atti notarili o le perizie che interessano: proprio per la sua funzione strumentale esso contiene anche un "indice dei capi". Molto interessanti sono, infine, i rimandi interni: il registro, conservato in uno studio, era apposto sulla scansia 11, ma, rispondendo alla finalità pratica di cui sopra, riporta molti collegamenti per reperire velocemente altri strumenti sia interni ("vedasi la scansia n. 8") che esterni ("vedasi il rotolo nella cassetta X al 36").

Ancora diverso il registro che segue (n. 541), dove sono contenute le scritture relative alle eredità, per la gestione e la conservazione del patrimonio della famiglia e che ha la forma descrittiva del memoriale ("Qui cominciano le piante delli terreni della heredità..." c. 1), seppur accompagnata dalla dimensione tecnica altamente specializzata e certificata dai due periti agrimensori che sottoscrivono le pagine (Piante mappe e cabrei, 1575-sec. XIX, unità 541). Un dato che non allontana l'apporto grafico dalle caratteristiche generali e consolidate di raffigurazione, ma le mette in secondo piano, investendo in altri apparati interessanti: un "elenco manoscritto delle piante dei terreni spettanti all'eredità che si contengono in questo libro" (Piante mappe e cabrei, 1575-sec. XIX, unità 541, cc. 25-26); "le piante de' terreni spettanti al Sig. Marchese Antonio Bolognini Amorini transuntate dall'Ufficio Censuario negli anni 1810, 1818 e 1822" (Piante mappe e cabrei, 1575-sec. XIX, unità 541, cc. 32-62); "l'estimo catastale de Beni ed effetti del Marchese Antonio Bolognini Amorini" (Piante mappe e cabrei, 1575-sec. XIX, unità 541, c. 65) e "l'elenco delle piante de terreni spettanti al sig. Marchese del 1822 che contengono in questo libro dalla pagina 33 alla pagina 61". Allegate alle mappe sono anche le missive relative ai rapporti con il catasto, ai censimenti sui beni posseduti e alle varie imposte. L'aspetto contenutistico prevalente ammette tuttavia ancora un realismo grafico corredato da cartigli didascalici con indicazione della proprietà, delle misure dei campi, delle colture caratteristiche e, in alcuni casi, con i nomi dei contadini cui era affidata la conduzione; un'attenzione grafica maggiore sui diplomi o sulle attestazioni (piante e misure di possessioni), la cui accuratezza misurava anche il prestigio familiare.

Molto diverso il cabreo inedito e recentemente ritrovato sul mercato antiquario e restituito dalla Soprintendenza Archivistica del Lazio all'Archivio di Stato di Bologna come appartenente all'antica famiglia gentilizia dei Pepoli. Questo cabreo non censito nell'inventario del Salterini, descrive graficamente solo i livelli della tenuta di Giazzano, appartenuta a Cornelio Pepoli e probabilmente è l'unico esemplare superstite di una intera serie perduta. La ricchezza del frontespizio, l'ariosità della composizione della pagina, la pulizia della linea grafica, l'attenzione alla didascalia e al cartiglio, sono tutti elementi che vanno facilmente denotando la ricchezza e il prestigio della famiglia committente.

#### Conclusioni

Il rapporto tra documentazione cartografica, spesso percepita essenzialmente nella sua dimensione iconografica, e altre tipologie documentali consente la ricostruzione di contesti integrati indiscutibilmente utili alla comprensione storica della città e del territorio. Non solo perché la



ricchezza del patrimonio cabreistico, conservato presso archivi e biblioteche, pone l'Italia in una posizione privilegiata, sollecitando un dibattito utile alla conservazione così come alla divulgazione delle risorse cartografiche, ma anche perché le fonti – immagini, i segni grafici, possono portare un valore informativo ancora maggiore se ricondotte a un più ampio contesto documentario. Il cabreo fu uno strumento di gestione amministrativa, con un ruolo strategico nell'economia del territorio; fu registro di titoli giuridici del possesso così come monumento di auto documentazione per la famiglia interessata alla conservazione e alla trasmissione del proprio patrimonio. I complessi cabreistici non sempre sono di facile lettura: gli studi a riguardo si concentrano, spesso, su documenti attinenti al territorio regionale o provinciale (Romiti 1997, 548-566; Guarducci e Rombai 1994, 137-156) proprio perché essi necessitano di analisi approfondite, rilevazioni attente che possano mostrare eventuali discrepanze tra originali e copie, tra diverse redazioni così come tra interpolazioni e addenda (Carassi 1987, 743-754). Non esiste, infatti, una precisa codificazione formale e simbolica nella formazione dei cabrei figurati: come abbiamo visto, è possibile parlare di consuetudine grafica e di una sua evoluzione, ma tendenzialmente sono le capacità tecniche degli estensori, l'esigenza del committente e lo scopo celebrativo piuttosto che utilitaristico a determinarne le fattezze finali. Attraverso pochi esempi si è cercato di mostrare come l'esame dei reperti grafici permetta di ricostruire il campo d'azione di istituzioni pubbliche, di privati e di corporazioni religiose, se analizzato alla luce del suo contesto di riferimento, mettendo altresì in evidenza un sistema complesso di relazioni: "i cabrei [...] raccontano una storia complessa, costituita da tante storie che si sintetizzano nelle mappe, la cui lettura quindi si dispiega su molteplici piani: innanzitutto le storie e le intenzioni degli attori e le loro reti di relazioni, dal committente all'agrimensore, ma a volte anche di altri personaggi che si agitano come comprimari o come comparse nella numerosa schiera dei convenuti che partecipano alla misura, poi alla storia della misura stessa, dei suoi saperi tecnico scientifici e della relativa formazione professionale di chi esercita l'agrimensura, la storia del paesaggio, in quanto sistema agrario, e del suo doppio, ossia della sua rappresentazione, di quel modello di paesaggio che il cabreo vuole codificare, la storia del contesto sociale e culturale e delle condizioni economiche che presiedono alla formazione e prima ancora alla decisione di realizzare il cabreo" (Sereno 2002, 151).



## Riferimenti bibliografici

Boris, Francesca. 2009. "L'archivio di Santa Maria del Tempio e altri documenti sui gerosolimitani a Bologna." In *Cavalieri di San Giovanni in Liguria e nell'Italia settentrionale: quadri regionali, uomini e documenti*, a c. di. Josepha Costa Restagno, 487-507. Genova: Albenga.

Briante, Paola. 2009. "Una fonte per lo studio del territorio piemontese. I cabrei dell'Ordine di Malta dell'Archivio di Stato di Torino." In *Cavalieri di San Giovanni in Liguria e nell'Italia settentrionale: quadri regionali, uomini e documenti*, a c. di. Josepha Costa Restagno, 349-71. Genova: Albenga.

Carassi, Marco. 1987. "Problemi di schedatura e conservazione della documentazione cartografica e iconografica." In *Cartografia e istituzioni in età moderna*, 743-54. Genova: Società Ligure di Storia Patria.

Conti, Elio. 1966. I catasti agrari della Repubblica Fiorentina e il Catasto particellare toscano (secoli XIV-XIX). Roma: Istituto storico italiano per il Medio Evo.

Corporazioni religiose soppresse, 992. Archivio di Stato di Bologna.

Crini, Pietro. 1993. "La cartografia tra pubblico e privato." In *Imago et descriptio Tusciae. La Toscana nella geocartografia dal XV al XIX secolo*, a c. di. Leonardo Rombai, 361-87. Venezia: Marsilio.

Dehò, Valeria, a c. di. 1993. Archivi privati e diversi, Fondo Salina Amorini Bolognini, 540, 541, 542. Inventario. Bologna: Archivio di Stato di Bologna.

Dal Pane, Luigi. 1974. "Descrizioni immobiliari e mappe nella storiografia delle strutture." In *Le collezioni d'arte della Cassa di risparmio in Bologna. I disegni. Mappe agricole e urbane nel territorio bolognese dei secoli XVII e XVIII*, vol. II, a c. di. Franca Varignana, XV-XVI. Bologna: Edizioni Alfa.

Dematteis, Giuseppe. 1990. Le metafore della terra: la geografia umana tra mito e scienza. Milano: Feltrinelli.

Dizionario enciclopedico italiano. 1955 -1961. Roma: Istituto della Enciclopedia Italiana.

Gambi, Lucio. 1976. "La campagna: gli uomini, la terra e le sue rappresentazioni visive." In *Storia d'Italia – Atlante VI*, 479-504. Torino: Einaudi.

Ginori Lisci, Leonardo. 1978. Cabrei in Toscana. Raccolte di mappe, prospetti e vedute sec. XVI – sec. XIX. Firenze: Cassa di risparmio di Pistoia e Pescia.

Girri, Francesco Maria. 1767. L'agrimensore istruito. Ferrara: G.A. Coatti.

Guarducci, Anna e Leonardo Rombai. 1994. "I cabrei della Prepositura e del Capitolo di Sant'Andrea d'Empoli (secoli XVII-XIX). Cartografia e territorio." In *Sant'Andrea* a *Empoli*, 137-56. Firenze: Giunti.

Manno Tolu, Rosalia. 1981. "Pergamene «creature» e cabrei degli Spedali di Pistoia", Rassegna degli Archivi di Stato XLI (1-2-3): 92-106.

Matrici di relazioni e piante, 1592-1791, Collezione Tognetti (sec. XVI – 1830), già Salina Amorini Bolognini, Archivi privati e diversi. Archivio di Stato di Bologna.



Panzeri, Matteo e Farruggia, Angela, a c. di. 2009. Fonti, metafonti e GIS per l'indagine della struttura storica del territorio, Torino: CELID.

Piante mappe e cabrei, 1575-sec. XIX, Fondo Salina Amorini Bolognini, 1284-sec XIX, Archivi privati e diversi, Archivio di Stato di Bologna.

Rezasco, Giulio. 1881. Dizionario del linguaggio italiano storico ed amministrativo. Firenze: Le Monnier.

Rombai, Leonardo. 2002. Geografia storica dell'Italia. Ambienti, territori, paesaggi. Firenze: Le Monnier.

Rombai, Leonardo, a c. di. 1993. *Imago et descriptio Tusciae*. *La Toscana nella geocartografia dal XV al XIX secolo*. Venezia: Marsilio.

Romiti, Beatrice. 1997. "La cartografia cabreistica lucchese e il cabreo ottocentesco della famiglia Bottini", *L'Universo*, 77 (4): 548-66.

Sereno, Paola. 2002. "Rappresentazione della proprietà fondiaria: i cabrei e la cartografia cabreistica." In *Rappresentare uno Stato. Carte e cartografi degli stati sabaudi dal XVI al XVIII secolo*, a c. di. Rinaldo Comba e Paola Sereno, vol. I. Torino: Allemandi.

Varignana, Franca, a c. di. 1974. Le collezioni d'arte della Cassa di risparmio in Bologna. I disegni. Mappe agricole e urbane nel territorio bolognese dei secoli XVII e XVIII, vol. III. Bologna: Edizioni Alfa.

Zangheri, Renato. 1961. La proprietà terriera e le origini del Risorgimento nel Bolognese. Vol. I, (1789-1804). Bologna: Zanichelli.



# Neurodegenerative clinical records analyzer: detection of recurrent patterns within clinical records towards the identification of typical signs of neurodegenerative disease history\*

Erika Pasceri<sup>(a)</sup>, Mérième Bouhandi<sup>(b)</sup>, Claudia Lanza<sup>(c)</sup>, Anna Perri<sup>(d)</sup>, Valentina Laganà<sup>(e)</sup>, Raffaele Maletta<sup>(f)</sup>, Raffaele Di Lorenzo<sup>(g)</sup>, Amalia C. Bruni<sup>(h)</sup>

a) Department of Culture, Education and Society, University of Calabria (CS), Italy, <a href="https://orcid.org/0000-0001-9917-2184">https://orcid.org/0000-0001-9917-2184</a>
b) LS2N, UMR CNRS 6004, Nantes Université, Nantes, France, <a href="https://orcid.org/0000-0001-7899-8198">https://orcid.org/0000-0001-7899-8198</a>
c) Department of Culture, Education and Society, University of Calabria (CS), Italy, <a href="https://orcid.org/0000-0002-3018-1987">https://orcid.org/0000-0002-3018-1987</a>
d) Department of Culture, Education and Society, University of Calabria (CS), Italy, <a href="https://orcid.org/0000-0002-2852-0919">https://orcid.org/0000-0000-0002-2852-0919</a>
e) Association for Neurogenetic Research (ARN), Lamezia Terme, CZ, Italy, <a href="https://orcid.org/0000-0001-8379-400X">https://orcid.org/0000-0001-8379-400X</a>
f) Regional Neurogenetic Centre, ASP CZ, Lamezia Terme, CZ, Italy, <a href="https://orcid.org/0000-0001-5848-9097">https://orcid.org/0000-0001-5848-9097</a>
g) Regional Neurogenetic Centre, ASP CZ, Lamezia Terme, CZ, Italy, <a href="https://orcid.org/0000-0003-3471-3343">https://orcid.org/0000-0003-3471-3343</a>

Contact: Erika Pasceri, <u>erika.pasceri@unical.it</u>; Mérième Bouhandi, <u>merieme.bouhandi@ls2n.fr</u>;
Claudia Lanza, <u>claudia.lanza@unical.it</u>; Anna Perri, <u>anna.perri@unical.it</u>; Valentina Laganà, <u>valelagana@gmail.com</u>;
Raffaele Maletta, <u>maletta@arn.it</u>; Raffaele Di Lorenzo, <u>dilorenzo@arn.it</u>; Amalia C. Bruni, <u>amaliaceciliabruni@gmail.com</u>
Received: 17 November 2022; Accepted: 09 January 2023; First Published: 15 May 2023

#### **ABSTRACT**

When treating structured health-system-related knowledge, the establishment of an over-dimension to guide the separation of entities becomes essential. This is consistent with the information retrieval processes aimed at defining a coherent and dynamic way – meaning by that the multilevel integration of medical textual inputs and computational interpretation – to replicate the flow of data inserted in the clinical records. This study presents a strategic technique to categorize the clinical entities related to patients affected by neurodegenerative diseases. After a pre-processing range of tasks over paper-based and handwritten medical records, and through subsequent machine learning and, more specifically, natural language processing operations over the digitized clinical records, the research activity provides a semantic support system to detect the main symptoms and locate them in the appropriate clusters. Finally, the supervision of the experts proved to be essential in the correspondence sequence configuration aimed at providing an automatic reading of the clinical records according to the clinical data that is needed to predict the detection of neurodegenerative disease symptoms.

#### **KEYWORDS**

Alzheimer; Categorization; Electronic health records (EHR); Machine learning; Semantic annotation.

<sup>\*</sup> Authors all contributed in the realization of this paper, although Mérième Bouhandi specifically dealt with sub-section "Classifying Alzheimer-related indicators", "Results" and "Discussion" Sections, Claudia Lanza dealt with the "Introduction" Section, covered the "Materials and methods", "Text recognition" Sections and detailed in sub-section "Neurodegenerative categories matching with expressions" the categories schematization to run the association between medical expressions and pre-established categories; Erika Pasceri mainly focused on Section "Related works" dealing with the preliminary related works on the clinical health records digitization process and annotation as well as the KOSs frameworks, on sub-section "Sample definition and clinical records digitization" addressing the case of study framework digitization and on "Conclusion" Section , Anna Perri dealt with the sub-section "Objectives and context framework" and integrated the process of validating the correct correlation within the tags, Valentina Laganà, Raffaele Maletta, Raffaele Di Lorenzo and Amalia Cecilia Bruni provided the whole set of clinical data giving access to the archive and supported the evaluation process of the isolated linguistic structures with their domain field of knowledge.

<sup>© 2023,</sup> The Author(s). This is an open access article, free of all copyright, that anyone can freely read, download, copy, distribute, print, search, or link to the full texts or use them for any other lawful purpose. This article is made available under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. JLIS.it is a journal of the SAGAS Department, University of Florence, Italy, published by EUM, Edizioni Università di Macerata, Italy, and FUP, Firenze University Press, Italy.



#### Introduction

This paper presents a multidisciplinary research activity dealing with the realization of a semantic analyzer tool for the management of information contained in the digital clinical records of patients affected by Alzheimer's Diseases (AD), a progressive and disabling neurodegenerative disorder that, rarely, can be inherited in an autosomal dominant way (Bruni, Bernardi, and Maletta 2021; Alzheimer's Association 2016). AD is characterized by cognitive deficits, e.g., memory loss and behavioral and psychological symptoms of dementia (BPSD), including a wide range of non-cognitive symptoms involving perception, mood, behavior, personality, and basic functioning (Bruni, Bernardi, and Maletta 2021; Bruni, Bernardi, and Gabelli 2020).

The disposal of structured clinical data referring uniquely to specific under-treatment or post-treatment records results to be a leading task (Mills 2019). Indeed, the decision-making operations undertaken by doctors within specialized health sectors are generally based on the reference to specific parameters structured over the clinical documents (Shellum et al. 2016)in part by providing the capability for a broad range of clinical decision support, including contextual references (e.g., Infobuttons. Therefore, when it comes to understanding the logic behind a medical set of procedures, it becomes essential to evaluate them under the lens of the clinical information structure that can facilitate the appropriate data input as well as the subsequent inference processes over the acquired knowledge base. This study specifically describes the steps followed towards the construction of a semantic analyzer for the electronic health records (HER) referring to the AD. In particular, the paper is subdivided into several sections reflecting the different stages pursued to reach the development of a clinical supporting reader from a semantic perspective capable of retrieving the AD-related categories from the analysis of the linguistic expressions included in the anamnesis.

#### Related works

In the healthcare literature, considerable attention has been directed toward the clinical decision support (CDS) tools in the way they can provide medical operators with a pre-settled clinical workflow meant to orientate the health data insertion and the subsequent execution of specialized tasks (Kharbanda et al. 2018; Spineth, Rappelsberger, and Adlassnig 2018; Tolley et al. 2018; Beeler, Bates, and Hug 2014). To this end, it is necessary to apply information and knowledge management advanced techniques and methodologies which allow users to understand, share and use available information and transform data into knowledge. In this study, a focus will be given to semantic annotation, classification and evaluation of the clinical data with respect to a reference corpus made of the anamnesis referred to patients suffering from AD syndrome. The concept "semantic annotation" is intended as "the process of attaching to a text document or other unstructured content, metadata about concepts (e.g., people, places, organizations, products or topics) relevant to it». Specifically referring to the biomedical domain, it is worth mentioning three biomedical annotators: (i) Clinical Text Analysis and Knowledge Extraction System (cTAKES) (Savova et al. 2010) based on Unstructured Information Management Architecture (UIMA) and OpenNLP frameworks; (ii) MetaMap (Stewart, von Maltzahn, and Abidi 2012) which exploits the Unified Medical Language System (UMLS) Metathesaurus to process the mapping with the med-



ical entities of the electronic health records and the concepts contained in the classification systems; (iii) MedCATTrainer annotator (Searle et al. 2019) which works in conjunction with Named Entity Recognition and Linking (NER+L) operators to extract medical information from texts. Despite the relevant outcomes found in exploiting the semantic annotator tools and the facility to apply their main functions to the source biomedical texts, the Natural Language Processing (NLP) tasks executed on unstructured texts via machine learning techniques intrinsically provide a more fine-grained systematization of the categories to be retrieved and used as tagging segmentation of clinical datasets. In this way, users can meet specific medical needs by collecting several important sequences of clinical records characterized by a medical recursive writing schema, offering early detection work on the patients' medical history. The detection of key textual units has also been addressed by (Hassanzadeh, Nguyen, and Koopman 2016), who exploit external officially shared semantic resources (MetaMap, NCBO annotator, Ontoserver, and QuickUMLS) to map the medical information in the EHR and obtain a more reliable set of data framework. (Klassen, Xia, and Yetisgen 2016) built NLP schemes to identify medical events in clinical notes in order to detect the diagnosis or coordination changes, as well as (Patel et al. 2018) who created a clinical entity recognition (CER) process using machine learning techniques classifying the desired outputs in categorized sequences to be retrieved. Tou et al. (2018) describe a study on the isolation of medical forecastable clusters referring to personal data, vital signs, or diagnosis results to detect the main forms of infections. The biomedical domain is also rich in Knowledge Organization Systems (KOSs), which differ in various aspects: their type (classification systems, thesauri, ontologies, etc.); their function and purpose (information retrieval, information sharing, indexing, etc.) (Mazzocchi 2018). Among these, Alzheimer's Disease Thesaurus is used for indexing and searching the ADEAR (Alzheimer's and related Dementias Education and Referral Center) database, which was created in 1990 by the US Congress as Alzheimer's Disease Education and Referral with the aim to «compile, archive, and disseminate information concerning Alzheimer's disease for health professionals, people with AD and their families, and the public». The OWL Ontology includes 156,869 classes belonging to different categories, such as organism, anatomy, biological process, neurological disease, neurological disorder, cellular anatomy, and so on.

## Objectives and context framework

The neurodegenerative categorization system from which this study has taken its ground has been forged from the one created by the Italian Institute of neurodegenerative diseases (Laganà et al. 2022)from 2006 to 2018, were studied. Symptoms have been extracted from Neuropsychiatric Inventory (NPI located in the South of Italy, with the purpose of enhancing it by executing machine learning operations. The research tasks will be conducted through a semantic analysis of the expressions contained in a sample of clinical records related to the AD-patients. Indeed, the expected achievement of this activity is the development of the automatic classification of the typical expressions contained in the clinical records with the enhanced version of the categories' systematization. The pre-existing categories concerning cognitive and motor signs/symptoms, as well as the BPSD, have been developed and validated by the neurologists



and psychologists working at the Neurogenetic Center of Calabria Region<sup>1</sup>, Italy, where the archive containing the clinical records that have been used in this study (Laganà et al. 2022) from 2006 to 2018, were studied. Symptoms have been extracted from Neuropsychiatric Inventory (NPI is located. The archive consists of 12,860 paper-based handwritten medical records and each of them consists of a folder with an extremely variable number of sheets, as the pages are incremented after each follow-up visit carried out on patients (including diagnostic tests, other laboratory tests, structured or instrumental tests). Texts contained in medical records are handwritten, so for the integrated use of information clinical data need to be extracted in a structured formal way. Data collected are essentially made from narrative texts (Coronato et al. 2014) describing the patients' everyday life, cognitive disorders, and all signs and symptoms that in most cases lead to the disease outbreak.

The approach adopted could be considered interdisciplinary as it requires interaction with knowledge organization experts, natural language technicians, and medical experts. The final product, which will integrate the results and the resources developed during the project, is represented by a repository accessible both from the members of the project staff and by the final users, mainly represented by domain experts. In the next section, the methodological approach will be described.

#### Materials and methods

The work starts from a sample of clinical records stored in the CRN database about patients suffering from AD<sup>2</sup>. Specifically, the total number of records is 12,860<sup>3</sup>. In order to make the textual information shareable for the automatic medical entities detection (signs, symptoms), all paper-handwritten clinical records referring just to dead patients suffering from AD have been considered for the digitization and processing through a software for handwritten text recognition. This software has been semi-automatically trained to allow the recognition of several handwriting styles of the doctors who wrote clinical records<sup>4</sup>.

## Sample definition and clinical records digitization

The first phase concerned the medical records sample acquisition and, consequently, their extraction from the CRN archive for the digitization activity. The arrangement of the paper-based handwritten clinical records in the CRN's archive follows a shelving disposition (Casanova 1928), within which the documents are organized according to chronological order (Lodolini 2011). For the purpose of this study, a sample of medical records of dead patients has been selected, but only the anamnesis section has been taken into account.

<sup>&</sup>lt;sup>1</sup> Regional Neurogenetic Centre (CRN), Lamezia Terme, Catanzaro (CZ), Italy.

<sup>&</sup>lt;sup>2</sup> The study has been approved by the Ethics Committee of University of Calabria, Italy (Protocol Number 2022-UCAL-PRG-0008403 08/02/2022).

<sup>&</sup>lt;sup>3</sup> The distribution is of 40,5% male and 59,5% female, the majority of them coming from the province of Catanzaro (Calabria Provincial Capital), and from other provinces of Calabria. Only 13.8% come from other Italian regions.

<sup>&</sup>lt;sup>4</sup> To adequately train the tool, two different training sets have been set up corresponding to two different doctors' handwriting styles.



## Text recognition

In this second stage, part of the digitized records has been transcribed in order to obtain a reusable file format to be treated in the categories' identification task. To carry out this process Transkribus software has been employed. This text recognition tool specifically works on handwritten documents, and it offers a way to transcribe line by line the sections of these latter by providing a set for training the association of the characters' recognition every time a new document is imported written by the same authors (see Figure 1).

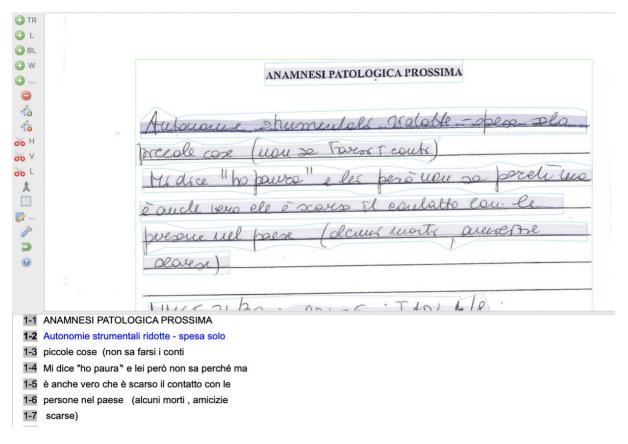


Figure 1. Extract from the Transkribus working environment on a clinical digitized record.

As depicted in the previous image, each line of the digitized clinic record corresponds to a region of the document. In this way, Transkribus allows users to insert the matching transcription of the characters and, consequently, learns how to identify the future writing styles. For this very case of study, Transkribus has been deployed to perform the model training over 100 clinical records of dead patients<sup>5</sup> with confirmed AD syndrome<sup>6</sup>, consisting of 243 pages subdivided as follows:

<sup>5</sup> The sample number has been set to 100 since this activity took its basis from an initial phase of clinical records' digitization, the number will be increased.

<sup>&</sup>lt;sup>6</sup> The digitization will be enriched by other clinical records and the research activity will describe the progression of the symptoms from Mild Cognitive Impairment (MCI) to AD for the identification of the categories referring to patients suffering from these syndromes (Petersen and Negash 2008).



Table 1. Details of the training model for Transkribus.

Train set	Validation set	n. words	n. lines	n. epochs
233	10	32730	5624	50

On Transkribus this procedure is named Handwritten Text Recognition (HTR+). It implies the training of a set that successively tests itself over a test set. It runs over 50 document regions: changes in the numbers impact the length of the process: the more epochs users choose, the longer the activity will take. Figure 2 shows the accuracy percentages of the model trained to process the documents of neurodegenerative records automatically. The y-axis represents the "Accuracy in CER", where CER means the Character Error Rate detected during the transcription process by the model, this curve begins at a level of 100, and it decreases alongside the improvement of the model performances (indeed, the blue line is the progress of the training and red that of the evaluations over the test set). As indicated by the software main webpage:

The value for the Test Set is the most significant as it shows how the HTR+ performs on pages that it has not been trained on. Results with a CER of 10% or below can be seen as very efficient for automated transcription. Results with a CER of 20-30% are sufficient to work with powerful Keyword Spotting technology.<sup>7</sup>

In this case, the CER on the train set corresponds to 13.07%, and this can be due to the fact that parts of the scanned clinical records were marked by several blank sections or some letters, such as the 'p' or the 'g' and 'q', the software was not able to correctly identify for the overlapping line of the letter with the others on the next rows. With this sample the CER on the validation set can be considered sufficiently at a good level considering that the two lines (the blue and the red) match at the end of the curve, meaning that the error is minimized as long as the training progresses.

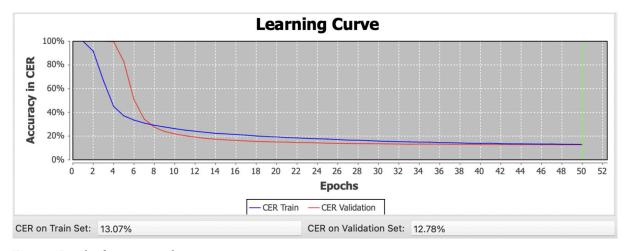


Figure 2. Levels of accuracy in the transcriptions.

<sup>&</sup>lt;sup>7</sup> For more details: <a href="https://readcoop.eu/transkribus/howto/how-to-train-a-handwritten-text-recognition-model-in-transkribus/">https://readcoop.eu/transkribus/howto/how-to-train-a-handwritten-text-recognition-model-in-transkribus/</a>



## Neurodegenerative categories matching with expressions

One goal of the study has been targeted to increase the clinical information about patients treated at CRN by extracting them through NLP techniques, since, to date, data about patients are manually imported into CRN database. The database has represented a solid starting point for implementing a network connection system between the clinical symptoms and signs sentence descriptions within the records and the corresponding categories. Along with the supervision of the CRN physicians and the analysis of the previous works on this subject, this study focused on a categories framework systematization onto two levels: (i) three top categories that have been, in turn, declined in (ii) sub-categories. The following list is meant to show the subdivision employed to reach an automatic identification of AD signs and symptoms descriptions.

	main categories cognitive	BPSD <sup>8</sup>	motor
subcategories	memory orientation speech disorders agnosia apraxia planning handwriting	behavior affective disorders psychosis emotionality sleep disorders	extrapyramidal signs

Once defined this flat top-down signs and symptoms structure, the methodology pursued in this study has been based on the identification of the expressions used by doctors in their descriptions of clinical events within the clinical records to be linked to the categories and sub-categories, as shown in the following figures (Figure 3, Figure 4, Figure 5).

<sup>&</sup>lt;sup>8</sup> For more details about Behavioral and Psychological Symptoms of Dementia see Cerejeira et al. (2012).



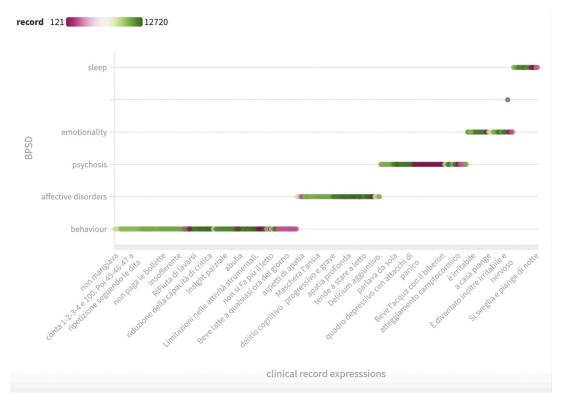


Figure 3. Distribution of subclasses in BPSD.

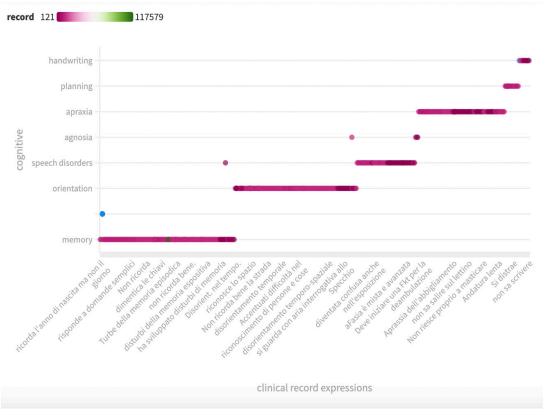


Figure 4. Distribution subclasses Cognitive.



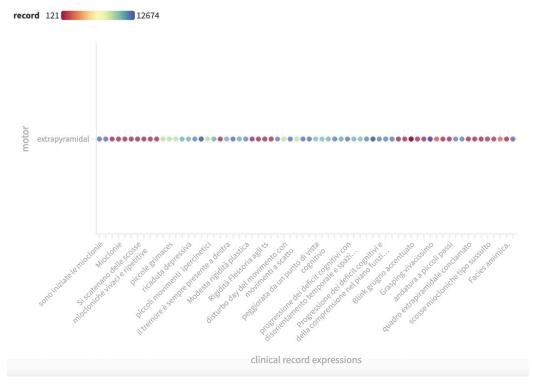


Figure 5. Distribution subclasses Motor.

The association of these expressions has been supported by a preliminary investigation of typical sentences used by the physicians in their descriptions within the clinical records' anamnesis compilation. This task implied the supervision of medical experts who supported the creation of a list of phrases per each category in order to develop a reliable training set for the future automatic identification of the matching expression plus categories. The total number of expressions retrieved in the train set sample is partitioned as the Table 2 shows:

Table 2. Clinical expressions partitioning.

		n. expressions		n. expressions		n. expressions
Top categories	cognitive	452	BPSD	379	motor	71
Sub-categories	memory	145	behavior	165	extrapyramidal signs	71
	orientation	128	affective disorders	75		
	speech disorders	58	psychosis	76		
	agnosia	4	emotionality	40		
	apraxia	88	sleep disorders	23		
	planning	16				
	handwriting	12				



Figure 6 depicts a scatter plot for the expressions related to each sub-categories. The following sections will detail the whole process developed to set a methodology aimed at automatically discovering the phrase segmentations related to the neurodegenerative signs and symptoms by implementing a machine learning schema.

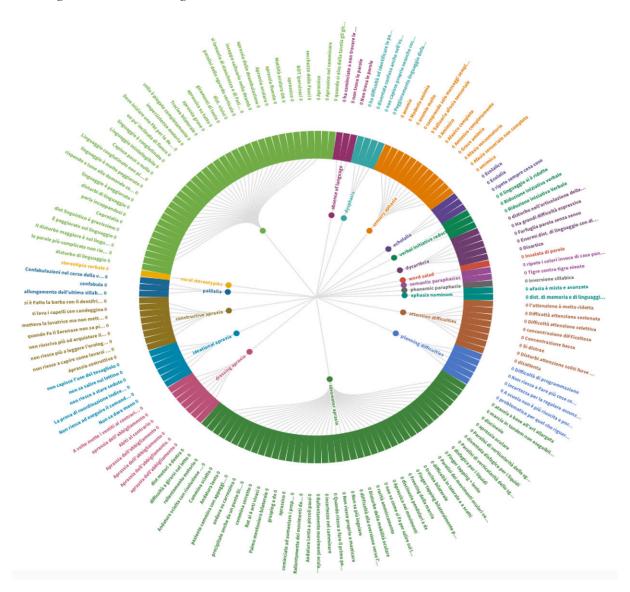


Figure 6. Scatterplot depicting the tags included in the source EHR corpus.

## Classifying Alzheimer-related indicators

In electronic medical records, health indicators, medications, laboratory values, symptoms, and personal history are typically embedded in free text form as clinical, hospitalization, and intervention reports, progress notes, and discharge summaries. Many NLP tasks can be conducted on these corpora, we will focus on extracting cognitive, BPSD, and motor Alzheimer-related indicators. Different NLP methods can automate the identification and classification of linguistic



entities that describe these essential concepts for a given domain, but they are quite challenging to be applied given the unstructured nature of linguistic data in medical records in the healthcare domain (Li et al. 2021). Less rigid methods, such as rule-based ones (Mykowiecka, Marciniak, and Kupść 2009), use token rules and regular expressions with some characteristics of the entities of interest to extract said entities. Finally, corpus-based methods use indicators from text corpora such as statistical information coupled with machine learning approaches for identifying and extracting these entities. Named-entity recognition tasks, knowledge extraction, and biomedical entities extraction, to cite a few, are all tasks that heavily rely on these processes (Lafferty, McCallum, and Pereira 2001; Settles 2004; Wu et al. 2015; Huang, Xu, and Yu 2015; Chalapathy, Borzeshi, and Piccardi 2016; Si et al. 2019).

Rule-based methods can be time-consuming to build and are prone to contextual conflicts, especially with more complex data, requiring a significant amount of human effort to build a complete set of tags, patterns, and domain-specific rules. For this, it results difficult to create a comprehensive and thorough list of rules due to the ever-evolving variability of the terms contained in the documentation under study. With these methods, however, the results are often satisfactory from an accuracy point of view, in terms of correlation between the exact expressions to be retrieved from the clinical records and the association to a pre-defined set of categories. Secondly, unknown and novel terms or rules are introduced unceasingly in active domains such as the healthcare, clinical or biomedical fields. In order to avoid the drawbacks of manual rules, machine learning approaches were proposed quite early on for NER and extraction tasks, with the usage, among other methods, of SVMs (Wu et al. 2015) and CRFs (Lafferty, McCallum, and Pereira 2001; Settles 2004; Si et al. 2019) for the classification and categorization step.

The neural approach to construct word representation (as well as sentences or document representations) can be seen as a crucial breakthrough in machine learning for NLP. Several methods exist for obtaining the word representations of all words in a predefined vocabulary of fixed size from textual corpora (Mikolov et al. 2013; Bojanowski et al. 2017; Devlin et al. 2019). Learning these representations is done in conjunction with training a neural network on a task, such as a document classification one. Thus, a matrix of weights from the network is called an embedding matrix. It can also be an unsupervised process, using statistical methods to represent the words in the corpus, as done in the earliest distributional methods.

Lower computation complexity is one of the main advantages of using the dense, low-dimensional vectors obtained from these methods compared to those obtained with classical distributional methods, eliminating the "curse of dimensionality" problem that early distributional methods based on high-dimensional co-occurrence matrices had. Furthermore, most neural methods output dense vector representations. The main advantage of these dense representations is their power of generalization. By choosing a small size for the word embeddings, the model is forced to choose the most relevant descriptors to populate the embedding matrix, discarding a good amount of the noise naturally existing in the corpus (Mikolov et al. 2013). These word representations are then used as input for actual task-oriented methods. In recent years, deep neural networks helped secure significant progress in NER and medical concept extraction by eliminating the necessity of feature engineering.



As shown in Figure 7, to process the  $X_t$  element, the model combines the representation of the input sequence up to the  $X_{t-1}$  element with the information of this new  $X_t$  element, thus creating a new state representing the input sequence up to the  $X_t$  element. For this reason, by maintaining a state vector that represents each element after it has been processed, it is impossible to parallelize the calculations, which is one of the major drawbacks of these recurrent models.

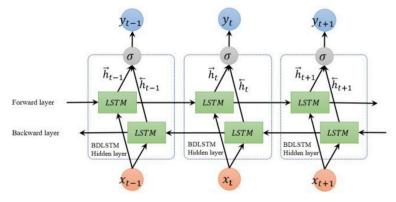


Figure 7. Layer description of a BiLSTM model.

Recurrent Neural Networks (RNNs) can keep track of sentence structure and various dependencies and allow information to be persistent over the network. However, vanilla RNNs often struggle to learn long-term temporal dependencies since their gradients can explode or completely vanish over multiple time steps. The vanilla RNN cell can then be replaced by a Long Short-Term Memory (LSTM) cell (Schuster and Paliwal 1997) or Bidirectional Long-Short Term Memory (BiLSTM) cell (Graves, Fernàndez, and Schmidhuber 2005) to solve this issue via a set of different gates. The addition of a CRF layer was often shown to surpass simple LSTM models for both NER and MCE (Chalapathy, Borzeshi, and Piccardi 2016; Panchendrarajan and Amaresan 2018). Conditional random fields are a class type of statistical modeling methods for prediction tasks where contextual information, i.e., the state of the neighboring tokens, affects the current prediction.

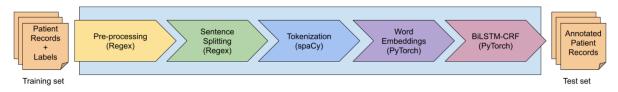


Figure 8. The NLP pipeline of the proposed work.

Our RNN uses two vertically stacked and fully-connected BiLSTM with a CRF layer on top, each LSTM cell uses 256 hidden units, and its dropout is set to 0.3. We only keep sequences that are 50 tokens longer or shorter and tagged expressions up to 8 tokens. We use the training-test sets with 30% withheld for the test sets. We train our model on 50 epochs, with Adam with Nesterov momentum (NAdam) as an optimization algorithm. Figure 9 displays the details of the BiLSTM-CRF module for sequence labeling:



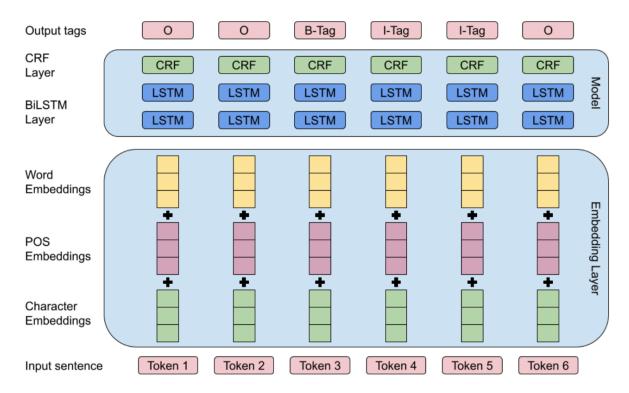


Figure 9. Diagram of the models architecture. The top part covers the language modeling and the tokens extraction and classification, the bottom parts show the different embeddings used to represent the tokens.

In this work, we tackle the problem with an end-to-end architecture. Given a dataset (split into training-test sets) and a set of entities with labels, the steps undertaken have been the following:

- Text preprocessing: uses an extensive set of regular expressions to clean and process the
  text. This step is crucial for any NLP task, and it transforms text into a more digestible
  form so that the methods and algorithms can perform better. This step is even more crucial
  in tasks where records are used since the records are often unstructured, free-form, and
  not normalized.
- 2. **Sentence splitting:** splits the medical record into sentences by relying on a set of regular expression-based rules that define sentence breaks.
- 3. Word tokenizing: splits the sentences into meaningful segments, i.e, tokens, using spaCy.
- 4. **Token embeddings:** each token is represented using three different embedding types. Word embeddings are typically learned using words from the corpus vocabulary during the training phase. We conjointly learn character embeddings and POS embeddings: these two types of embeddings don't encode the same information that word embeddings contain. Character-level embeddings can be considered encoded lexical information, and POS embeddings encode syntactic context.
- 5. **Entity extraction:** the model learns the embeddings of the given tokens and directly uses them to predict the label for each token. We use the tags and sub-tags, the entities provided by the doctors, and the "I-O-B" labels for the tags. "I-O-B" Tagging is a standard tagging format for tagging tokens in tasks like name entity recognition. The "B-" prefix indicates



that the tag is the beginning of a chunk, and an "I-" prefix indicates that the tag is inside a chunk. An "O" tag indicates that a token belongs to no predefined entity and indicates that that token is not to be extracted.

#### Results

In this section we present the result of the evaluation of the proposed methods results using the test corpora. We report the precision, recall, and F1-score, the classical evaluation metrics for entity recognition and sequence labeling tasks. 70% of the dataset is used for training and 30% for testing.



Figure 10. Excerpt from outputs of our algorithm for one of the clinical records. The tagging is learned via the neural network, using "displaCy".

The data is randomly split before training the model and for each run, the seed is randomly initialized too. Each configuration is run three times for our experiments, and the reported results are the average of these runs. Both methods aim to extract Alzheimer's indicators in uploaded clinical records corresponding to the following categories of medical reports: cognitive, BPSD and motor. The following table reports the results for the main categories:

Table 3. Analysis (in % Precision, Recall and F1-score) of the model's outputs for the extraction task. This is an averaging of 3 runs using only the category labels.

Top categories	Precision	Recall	F1-Score
cognitive	0.17	0.37	0.24
BPSD	0.23	0.53	0.31
motor	0.19	0.72	0.30

https://explosion.ai/demos/displacy



#### Discussion

The results differ for most classes, some of these are much more represented in the task corpus than others. The recall is much higher than precision, indicating that too many false positives can be found in the output lists. There can be five main types of problems for each predicted entity and a set of ground-truth entities  $E_r$ :

- 1. Type-1:  $e_p$  is not present in  $E_T$ , false positive.
- 2. Type-2: An  $E_{\tau}$  entity is not predicted, false negative.
- 3. Type-3:  $e_p$  and an  $E_T$  entity have the same span but different labels.
- 4. Type-4:  $e_p$  and an  $E_T$  entity have overlapping spans and different labels.
- 5. Type-5:  $e_p$  and an  $E_\tau$  entity have overlapping spans and the same labels.

Table 4 presents examples of entities extracted by our model:

Table 4. Examples of entities extracted by the model and matching errors.

Entity	Expected Tags (from $E_{\tau}$ )	Actual Tags (for e <sub>p</sub> )	Type of Error
modesta rigidità plastica	B-Motor, I-Motor, I-Motor	B-Motor, I-Motor, I-Motor	No error
riduzione iniziativa verbale	B-Cognitive, I-Cognitive, I-Cognitive	B-Cognitive, I-Cognitive, I-Cognitive	No error
voluta uscire più da sola	O, O, O, O, O	B-Motor, I-Motor, I-Motor, I-Motor, I-Motor	Type-1
episodi di disorientamento spaziale	B-Cognitive, I-Cognitive, I-Cognitive, I-Cognitive	O, O, O, O	Type-2
non sa dare mano	B-Cognitive, I-Cognitive, I-Cognitive, I-Cognitive	B-Motor, I-Motor, I-Motor, I-Motor	Type-3
si guarda circospetto dietro	B-BPSD, I-BPSD, I-BPSD, I-BPSD	O, B-Motor, I-Motor, O	Type-4
sviluppato un atteggiamento disforico	B-BPSD, I-BPSD, I-BPSD, I-BPSD	O, O, B-BPSD, I-BPSD	Type-5

Using morpho-syntactic, lexical, semantic, and distributional information allows us to use much richer information as input data for our model. In future work, an ablation study will be conducted with the different representations used to test the contribution of each and every vector representation in the final scores.

The results can directly suffer from the small amount of input data. Transfer learning (Ruder et al. 2019) can be used here to alleviate this issue, it is the set of methods that allow the transfer of knowledge acquired from solving a given problem to another problem or from a domain to another. Our models require significant resources to be tuned appropriately. However, by using pre-trained models as a starting point, transfer learning allows to train or fine-tune our model without needing much training data.

The perspective is to take advantage of unsupervised training methods (such as BERT language model, Devlin et al. 2019) in future work. It must be noted that our entity extraction work is mainly preliminary, and the methods presented here can be considered baselines for further work.



#### **Conclusion**

This study configured a methodology to retrieve categorized medical expressions to define the correct classification of AD's signs and symptoms. The purpose of this investigation addressed the identification of typical sentences in the digitized clinical records to be automatically mapped with a two-level system categorization. The research work developed a multidisciplinary approach: from paper-based handwritten clinical records to a digitized corpus from which to detect in an automatic way salient medical information to be mapped with normalized neurodegenerative-related categories and sub-categories. The corpus analyzed has been built from anamnesis texts totally written in natural language that for its nature is rich of irregular expressions. This has impacted the configuration of a twofold categorization model meant to contain the mapping between the medical recursive expressions related to AD signs and symptoms and the sub-categories selected with the supervision of the physicians working in this sector. In future work activities, along with the enrichment of the documents, the analyses will be targeted to the classification of the clinical records according to the declared syndromes doctors have assigned to each patient and to the correlation of these diseases with the corresponding automatic symptoms detection.



#### References

Alzheimer's Association. 2016. «2016 Alzheimer's disease facts and figures». *Alzheimer's & Dementia* 12 (4): 459–509.

Beeler, Patrick Emanuel, David Westfall Bates, and Balthasar Luzius Hug. 2014. «Clinical decision support systems». *Swiss Medical Weekly* 144 (w14073): 1–7. <a href="https://doi.org/doi.org/10.4414/smw.2014.14073">https://doi.org/doi.org/10.4414/smw.2014.14073</a>.

Bojanowski, Piotr, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. «Enriching Word Vectors with Subword Information». *Transactions of the Association for Computational Linguistics* 5: 135–46.

Bruni, Amalia Cecilia, Livia Bernardi, and Carlo Gabelli. 2020. «From beta amyloid to altered proteostasis in Alzheimer's disease». *Ageing research reviews* 64: 101126.

Bruni, Amalia Cecilia, Livia Bernardi, and Raffaele Maletta. 2021. «Evolution of genetic testing supports precision medicine for caring Alzheimer's disease patients». *Current Opinion in Pharmacology* 60: 275–80.

Casanova, Eugenio. 1928. Archivistica. Siena: Stab. arti grafiche Lazzeri.

Chalapathy, Raghavendra, Ehsan Zare Borzeshi, and Massimo Piccardi. 2016. «Bidirectional LSTM-CRF for Clinical Concept Extraction». <a href="https://doi.org/10.48550/arXiv.1610.05858">https://doi.org/10.48550/arXiv.1610.05858</a>.

Coronato, Antonio, Giuseppe Di Pietro, Amalia Cecilia Bruni, Erika Pasceri, Maria Teresa Chiaravalloti, and Giovanni Paragliola. 2014. «ALPHA: an eAsy inteLligent service Platform for Healthy Ageing». In *Ambient Assisted Living*, edited by Bruno Andò, Pietro Siciliano, Vincenzo Marletta, and Andrea Monteriù. Springer.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. «BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding». <a href="http://arxiv.org/abs/1810.04805">http://arxiv.org/abs/1810.04805</a>.

Graves, Alex, Santiago Fernàndez, and Jürgen Schmidhuber. 2005. «Bidirectional LSTM networks for improved phoneme classification and recognition». In *ICANN'05: Proceedings of the 15th international conference on Artificial neural networks: formal models and their applications – Volume Part II*, edited by Duch Włodzisław, Janusz Kacprzyk, Zadrozny Sławomi, and Oja Erkku. Berlin, Heidelberg: Springer-Verlag.

Hassanzadeh, Hamed, Anthony Nguyen, and Bevan Koopman. 2016. «Evaluation of Medical Concept Annotation Systems on Clinical Records». In *Proceedings of the Australasian Language Technology Association Workshop* 2016, 15–24. https://aclanthology.org/U16-1002.

Huang, Zhiheng, Wei Xu, and Kai Yu. 2015. «Bidirectional LSTM-CRF Models for Sequence Tagging». http://arxiv.org/abs/1508.01991.

Kharbanda, Elyse O., Steve E. Asche, Alan R. Sinaiko, Heidi L. Ekstrom, James D. Nordin, Nancy E. Sherwood, Patricia L. Fontaine, Steven P. Dehmer, Deepika Appana, and Patrick O'Connor. 2018. «Clinical Decision Support for Recognition and Management of Hypertension: A Randomized Trial». *Pediatrics* 141 (2): e20172954. <a href="https://doi.org/10.1542/peds.2017-2954">https://doi.org/10.1542/peds.2017-2954</a>.



Klassen, Prescott, Fei Xia, e Meliha Yetisgen. 2016. «Annotating and Detecting Medical Events in Clinical Notes». In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, 3417–21. European Language Resources Association. <a href="https://aclanthology.org/L16-1545.pdf">https://aclanthology.org/L16-1545.pdf</a>.

Lafferty, John, Andrew McCallum, and Fernando C. N. Pereira. 2001. «Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data Labeling Sequence Data». In *Proceedings of the Eighteenth International Conference on Machine Learning*, 282–89. Morgan Kaufmann Publishers Inc.

Laganà, Valentina, Francesco Bruno, Natalia Altomari, Giulia Bruni, Nicoletta Smirne, Sabrina Curcio, Maria Mirabelli, Rosanna Colao, Gianfranco Puccio, Francesca Frangipane, Chiara Cupidi, Giusy Torchia, Gabriella Muraca, Antonio Malvaso, Desirèe Addesi, Alberto Montesanto, Raffaele Di Lorenzo, Amalia Cecilia Bruni, and Raffaele Maletta. 2022. «Neuropsychiatric or Behavioral and Psychological Symptoms of Dementia (BPSD): Focus on Prevalence and Natural History in Alzheimer's Disease and Frontotemporal Dementia». *Frontiers in Neurology* 13 (June): 832199. https://doi.org/10.3389/fneur.2022.832199.

Li, Irene, Jessica Pan, Jeremy Goldwasser, Neha Verma, Wai Pan Wong, Muhammed Yavuz Nuzumlalı, Benjamin Rosand, Yixin Li, Matthew Zhang, David Chang, R. Andrew Taylor, Harlan M. Krumholz, and Dragomir Radev. 2021. «Neural Natural Language Processing for Unstructured Data in Electronic Health Records: a Review». http://arxiv.org/abs/2107.02975.

Lodolini, Elio. 2011. Archivistica. Principi e problemi. Milano: Franco Angeli.

Mazzocchi, Fulvio. 2018. «Knowledge organization system (KOS)». *Knowledge Organization* 45 (1): 54–78.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. «Efficient Estimation of Word Representations in Vector Space». <a href="http://arxiv.org/abs/1301.3781">http://arxiv.org/abs/1301.3781</a>.

Mills, Sherri. 2019. «Electronic Health Records and Use of Clinical Decision Support». *Critical Care Nursing Clinics of North America* 31 (2): 125–31. https://doi.org/10.1016/j.cnc.2019.02.006.

Mykowiecka, Agnieszka, Małgorzata Marciniak, and Anna Kupść. 2009. «Rule-Based Information Extraction from Patients' Clinical Data». *Journal of Biomedical Informatics* 42 (5): 923–36. https://doi.org/10.1016/j.jbi.2009.07.007.

Panchendrarajan, Rrubaa, and Aravindh Amaresan. 2018. «Bidirectional LSTM-CRF for Named Entity Recognition». In *Proceedings of the 32nd Pacific Asia Conference on Language, Information and Computation*, 531-540. Hong Kong: Association for Computational Linguistics.

Patel, Pinalkumar, Disha Davey, Vishal Panchal, and Parth Pathak. 2018. «Annotation of a Large Clinical Entity Corpus». In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2033–42. Brussels, Belgium: Association for Computational Linguistics.

Petersen, Ronald C., and Selamawit Negash. 2008. «Mild Cognitive Impairment: An Overview». *CNS Spectrums* 13 (1): 45–53. https://doi.org/10.1017/s1092852900016151.

Ruder, Sebastian, Matthew E. Peters, Swabha Swayamdipta, and Thomas Wolf. 2019. «Transfer



Learning in Natural Language Processing». In *Proceedings of the 2019 Conference of the North*, 15–18. Minneapolis, Minnesota: Association for Computational Linguistics. <a href="https://doi.org/10.18653/v1/N19-5004">https://doi.org/10.18653/v1/N19-5004</a>.

Savova, Guergana K, James J Masanz, Philip V Ogren, Jiaping Zheng, Sunghwan Sohn, Karin C Kipper-Schuler, and Christopher G Chute. 2010. «Mayo Clinical Text Analysis and Knowledge Extraction System (CTAKES): Architecture, Component Evaluation and Applications». *Journal of the American Medical Informatics Association* 17 (5): 507–13. <a href="https://doi.org/10.1136/jamia.2009.001560">https://doi.org/10.1136/jamia.2009.001560</a>.

Schuster, M., and K.K. Paliwal. 1997. «Bidirectional recurrent neural networks». *IEEE Transactions on Signal Processing* 45 (11): 2673–81. https://doi.org/10.1109/78.650093.

Searle, Thomas, Zeljko Kraljevic, Rebecca Bendayan, Daniel Bean, and Richard Dobson. 2019. «MedCATTrainer: A Biomedical Free Text Annotation Interface with Active Learning and Research Use Case Specific Customisation». <a href="https://doi.org/10.48550/arXiv.1907.07322">https://doi.org/10.48550/arXiv.1907.07322</a>.

Settles, Burr. 2004. «Biomedical Named Entity Recognition using Conditional Random Fields and Rich Feature Sets». In *Proceedings of the International Joint Workshop on Natural Language Processing in Biomedicine and its Applications (NLPBA/BioNLP)*, 107–10. Geneva, Switzerland: Coling.

Shellum, Jane L., Robert R. Freimuth, Steve G. Peters, Rick A. Nishimura, Rajeev Chaudhry, Steve J. Demuth, Amy L. Knopp, Timothy A. Miksch, and Dawn S. Milliner. 2016. «Knowledge as a Service at the Point of Care». *AMIA ... Annual Symposium Proceedings. AMIA Symposium* 2016: 1139–48.

Si, Yuqi, Jingqi Wang, Hua Xu, and Kirk Roberts. 2019. «Enhancing Clinical Concept Extraction with Contextual Embeddings». *Journal of the American Medical Informatics Association: JAMIA* 26 (11): 1297–1304. https://doi.org/10.1093/jamia/ocz096.

Spineth, Martin, Andrea Rappelsberger, and Klaus-Peter Adlassnig. 2018. «Implementing CDS Hooks Communication in an Arden-Syntax-Based Clinical Decision Support Platform». *Studies in Health Technology and Informatics* 255: 165–69.

Stewart, Samuel Alan, Maia Elizabeth von Maltzahn, and Syed Sibte Raza Abidi. 2012. «Comparing Metamap to MGrep as a Tool for Mapping Free Text to Formal Medical Lexicons». In *Knowledge Extraction and Consolidation from Social Media (KECSM 2012)*, 63–77.

Tolley, Clare L., Sarah P. Slight, Andrew K. Husband, Neil Watson, and David W. Bates. 2018. «Improving Medication-Related Clinical Decision Support». *American Journal of Health-System Pharmacy* 75 (4): 239–46. <a href="https://doi.org/10.2146/ajhp160830">https://doi.org/10.2146/ajhp160830</a>.

Wu, Yonghui, Jun Xu, Min Jiang, Yaoyun Zhang, and Hua Xu. 2015. «A Study of Neural Word Embeddings for Named Entity Recognition in Clinical Text». *AMIA ... Annual Symposium Proceedings. AMIA Symposium* 2015: 1326–33.



# "Resource not found": cultural institutions, interinstitutional cooperation and collaborative projects for web heritage preservation

## Chiara Storti<sup>(a)</sup>

a) Biblioteca Nazionale Centrale di Firenze

Contact: Chiara Storti, <a href="mailto:chiara.storti@cultura.gov.it">chiara.storti@cultura.gov.it</a>
Received: 31 January 2023; Accepted: 02 March 2023; First Published: 15 May 2023

#### **ABSTRACT**

Awareness of the need to preserve the Web heritage began to spread in the early 1990s, and the first Web archiving initiatives were launched at the National Library of Australia in 1993-1996. However, even today, especially in Italy, it is difficult to find contributions that address the topic of Web and Social media archiving from the perspective of process governance and its relationship with the laws and ethical issues. The purpose of this article is to investigate which actors are involved in these processes, what are the main criticalities or innovations compared to "traditional" cultural heritage management, providing concrete examples where possible.

#### **KEYWORDS**

Web archiving; Social Media archiving; Digital legal deposit; Internet Archive.

# "Risorsa non trovata": istituzioni culturali, cooperazione interistituzionale e progetti collaborativi per la conservazione del patrimonio web

#### ABSTRACT

Vedersi restituire l'errore "risorsa non trovata" è un'esperienza comune a tutti coloro che cercano informazioni sul Web. La conservazione del patrimonio web è, infatti, una delle sfide della società contemporanea. Il presente contributo si propone non tanto di analizzare le ormai consolidate tecnologie per il Web e Social media archiving ma piuttosto il piano della governance complessiva dei processi di archiviazione e quello del quadro legislativo di settore non solo a livello nazionale, ma anche nel rapporto con il sovranazionale. Laddove possibile riportando casi reali, si cerca inoltre di definire chi siano i diversi soggetti coinvolti e responsabili della costruzione di un'eredità digitale nazionale che possa essere fruita e compresa compiutamente dalle prossime generazioni.

#### PAROLE CHIAVE

Web archiving; Social media archiving; Conservazione digitale; Deposito legale.

<sup>© 2023,</sup> The Author(s). This is an open access article, free of all copyright, that anyone can freely read, download, copy, distribute, print, search, or link to the full texts or use them for any other lawful purpose. This article is made available under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. JLIS.it is a journal of the SAGAS Department, University of Florence, Italy, published by EUM, Edizioni Università di Macerata, Italy, and FUP, Firenze University Press, Italy.



Negli Stati Uniti e in Australia una sensibilità verso la necessità di conservare le risorse web ha cominciato a diffondersi già a partire dall'inizio degli anni '90 del '900. Anche in Italia, seppure con evidente ritardo, sono ormai quasi venti anni che si pubblicano contributi che trattano di Web e Social media archiving<sup>1</sup> su riviste e collane specializzate nell'ambito delle Scienze archivistiche e biblioteconomiche, delle Digital Humanities e dell'Informatica applicata. Negli ultimissimi anni anche alcune testate giornalistiche generaliste<sup>2</sup> o divulgative specializzate<sup>3</sup> hanno portato l'argomento all'attenzione di un pubblico più vasto, di frequente partendo da fatti di cronaca quali la chiusura o il cambiamento di management di un sito o servizio online. Tuttavia, almeno in ambito nazionale, è difficile reperire contributi che integrino i diversi punti di vista da cui il Web e Social media archiving può essere compreso, o almeno problematizzato. Nella letteratura specializzata viene esplorato principalmente negli aspetti che riguardano le tecnologie, cercando talvolta di individuare quelle che siano maggiormente in grado di restituirci un prodotto che non risulti "estraneo" alla nostra tradizione storico-culturale e archivistica-biblioteconomica. A questo approccio è legata una prospettiva "istituzionale" o "pubblica". D'altra parte, su testate più generaliste, oltre che sulle tecnologie, ci si concentra sulle implicazioni politiche e sociologiche della perdita o della conservazione di significative moli di dati e informazioni, citando le responsabilità delle grandi aziende del settore tech.

In tutti i casi viene quasi completamente omesso il piano della governance complessiva dei processi di archiviazione e quello del quadro legislativo di settore non solo a livello nazionale, ma anche nel rapporto con il sovranazionale.

L'Italia, rispetto alla maggioranza dei Paesi occidentali, arriva tra gli ultimi a porsi il problema di una strategia nazionale di conservazione del Web ma, proprio per questo, si può avvalere dell'esperienza di quasi tre decenni di iniziative di Web archiving. In particolare, è ormai chiaro come cercare di applicare alle risorse del Web le stesse categorie ontologiche e le stesse procedure di conservazione e accesso finora utilizzate per le risorse tradizionali, o per quelle digitali non diffuse su Web, rischi di farci perdere una grande quantità di informazioni e di restituire alle generazioni future un'immagine distorta o non completa del nostro presente. È pertanto necessario trovare soluzioni che prendano in considerazione non solo le tecnologie ma soprattutto la governance dei processi, e che questi processi siano indirizzati a preservare non solo gli oggetti in sé ma anche il contesto storico, politico e sociologico in cui le risorse sono state create, nel rispetto di vincoli normativi ed etici. Appare altresì inevitabile definire più chiaramente ruoli e forme di collaborazione tra le istituzioni del patrimonio, i proprietari e gestori delle piattaforme, fino ai produttori di contenuti: pubblico o privato, singolo o azienda. In questa sede si prova ad indagare quali siano i diversi soggetti coinvolti, laddove possibile riportando casi reali, per costruire un'eredità digitale nazionale che possa essere fruita e compresa compiutamente dalle prossime generazioni.

<sup>&</sup>lt;sup>1</sup> Per una parziale bibliografia in italiano si vedano: <a href="https://www.bncf.firenze.sbn.it/biblioteca/web-archiving/#Contributi%20online%20in%20italiano">https://www.bncf.firenze.sbn.it/biblioteca/web-archiving/#Contributi%20online%20in%20italiano</a> e <a href="https://site.unibo.it/web-and-social-media-archiving-and-preservation/it/strumenti/pubblicazioni">https://site.unibo.it/web-and-social-media-archiving-and-preservation/it/strumenti/pubblicazioni</a>.

<sup>&</sup>lt;sup>2</sup> Si segnalano in particolare gli articoli de IlPost.it: <a href="https://www.ilpost.it/tag/internet-archive/">https://www.ilpost.it/tag/internet-archive/</a>.

<sup>&</sup>lt;sup>3</sup> Si segnalano in particolare gli articoli di Wired.it. Significativo anche che il tema sia stato trattato di recente in una puntata del podcast "CRASH-La chiave per il digitale" dal titolo "Errore 404, così il web sta perdendo i pezzi": <a href="https://open.spotify.com/episode/7iYIin0mYnzkCKporF9pjN?si=Xod5DaFURlihnvb782JzTA">https://open.spotify.com/episode/7iYIin0mYnzkCKporF9pjN?si=Xod5DaFURlihnvb782JzTA</a>.



## "L'invenzione del Web archiviato"4

Si è soliti far risalire le origini del Web archiving all'ottobre del 1996, quando gli ingegneri di Internet Archive lanciarono il primo web *crawler* allo scopo di catturare tutto il Web allora esistente.<sup>5</sup> Il poi divenuto celebre "errore 404" (Parlangeli 2017) era infatti già molto diffuso anche se, a causa dell'ancora scarsa familiarità della maggioranza delle persone con Internet, non era ritenuto un vero problema. Brewster Kahle e Bruce Gilliat ne intuirono ugualmente la portata e misero a punto un sistema per copiare e conservare le pagine web prima che sparissero. Dal 2001, le pagine web oggetto di periodica archiviazione non solo da parte di Internet Archive, ma anche di numerosi altri partner pubblici o privati che utilizzano i loro strumenti, sono liberamente ricercabili e consultabili tramite la Wayback Machine. Quasi contemporaneamente a quel primo web crawling massivo ad opera di Internet Archive, altre istituzioni della memoria, in particolare biblioteche e archivi nazionali, iniziarono harvesting selettivi delle informazioni in rete,8 sulla base della loro appartenenza ad un determinato dominio o gruppo di domini o per argomento. Servizi che nel tempo le comunità nazionali o le comunità di interesse hanno meglio definito per conservare e rendere accessibile il patrimonio web. Il web crawling o web harvesting è ancora la tecnologia più usata per la copia delle pagine web, così come nel frattempo si sono consolidati il formato WARC<sup>10</sup> per l'archiviazione e la Wayback Machine per il replay delle copie archiviate. Tuttavia, in un articolo pubblicato su *Internet Histories*, Kieran Hegarty (2022) ha riportato i risultati di una ricerca effettuata alla National Library of Australia (d'ora in poi NLA) sul periodo compreso tra il 1993 e il 1996, ovvero su quella che l'autore definisce la preistoria del Web archiving. In quei primissimi anni '90 del '900, infatti, oltre a singole iniziative di piccoli gruppi amatoriali che provarono a salvare alcune risorse web<sup>11</sup> per specifiche finalità, di solito in file HTML o con screenshot, alla NLA si andava costituendo il gruppo di lavoro che avrebbe portato alla creazione del programma nazionale australiano di Web archiving "PANDORA". 12 Dall'indagine sulle testimonianze documentali e dalle interviste ai protagonisti dell'epoca emerge come il fatto che le prime infrastruttu-

<sup>&</sup>lt;sup>4</sup> Il titolo del paragrafo è liberamente tradotto da Hegarty (2022).

<sup>&</sup>lt;sup>5</sup> La video intervista ad un giovane Brewster Kahle, allora impegnato nel primo web *crawling*, è stata pubblicata nel 2021 da Internet Archive in occasione del 25° anniversario dell'evento: <a href="https://archive.org/details/wayback-machine-1996">https://archive.org/details/wayback-machine-1996</a>.

<sup>&</sup>lt;sup>6</sup> All'interno della Wayback Machine sono indicizzate anche tutte le url archiviate e definite "pubblicamente accessibili" dalle istituzioni della memoria che utilizzano la suite "Archive-it" <a href="https://archive-it.org/">https://archive-it.org/</a>, oltre che le url archiviate grazie a iniziative particolari di Web archiving come quelle realizzate dall'Archive Team e di cui si dirà più avanti.

<sup>&</sup>lt;sup>7</sup> https://archive.org/web.

<sup>&</sup>lt;sup>8</sup> Un elenco sufficientemente esaustivo di queste iniziative è la *List of Web archiving initiatives* su en.wikipedia: <a href="https://en.wikipedia.org/wiki/List">https://en.wikipedia.org/wiki/List of Web archiving initiatives</a>.

<sup>&</sup>lt;sup>9</sup> L'unica copia integrale del dominio .it risale al 2006 per iniziativa della Biblioteca Nazionale Centrale di Firenze che l'ha resa di recente disponibile nella propria collezione di Archive-it: <a href="https://archive-it.org/collections/15697">https://archive-it.org/collections/15697</a>. Per approfondimenti (Bergamin 2006).

<sup>10</sup> https://www.loc.gov/preservation/digital/formats/fdd/fdd000236.shtml.

<sup>&</sup>lt;sup>11</sup> In questo stesso periodo si colloca anche l'*Occasio project*, uno dei primi progetti pionieri di Social Media archiving, lanciato nel 1994 dall'International Institute of Social History di Amsterdam, allo scopo di conservare le conversazioni a tema sociale e politico postate, tra il 1988 e il 2002, nei gruppi di discussione online. L'archivio non è più disponibile online ma se ne possono recuperare alcune pagine su Internet Archive: <a href="https://web.archive.org/web/20020602111206/http://newsarchive.occasio.net/">https://web.archive.org/web/20020602111206/http://newsarchive.occasio.net/</a>.

<sup>12</sup> https://pandora.nla.gov.au/.



re<sup>13</sup> per il Web archiving siano state costruite all'interno di biblioteche nazionali abbia fortemente influenzato e ancora influenzi le modalità con cui si sceglie cosa e come archiviare, in ultimo determinando, quasi sempre inconsapevolmente, l'immagine che creiamo del passato recente e come questo potrà essere conosciuto in futuro.

L'assunto da cui parte Hegarty è che nessuna infrastruttura può essere creata dal nulla ma che ogni nuova infrastruttura si innesta su infrastrutture preesistenti, a volte ampliandole e migliorandole, altre semplicemente adattandole a nuovi scopi. E il tipo di infrastruttura su cui un nuovo servizio viene innestato ne determina quasi sempre le caratteristiche fondamentali.

When considering the history of web archiving, is important to reflect on why most web archiving institutions are major public libraries (Gomes et al. 2011). Indeed, the dominant web archiving institution, the Internet Archive, considers itself a library and its founder Brewster Kahle refers to himself as a librarian. Here, the notion of the public record is critical. Libraries lean on their responsibility to collect, preserve, and provide access to the published output (usually of a nation) when incorporating a selection of publicly available web within their collections (Brügger 2016). By asserting the material they collect from the web is publicly available, libraries position web archiving in terms of continuity with past collecting practices (Rauber et al. 2008). This positioning allows web archiving systems and practices to build on top of, extend, or anticipate frameworks that major public libraries have long relied upon to develop their collections, particularly legal deposit exemptions that call on publishers to deposit a copy of any publications to the library. Positioning web material as "publications" and website producers as "publishers" is therefore a key assumption that animates web archiving infrastructure and sees its actors and materials largely situated within major public libraries (Hegarty 2022).

In altre parole, il fatto che sia stato proprio all'interno di una biblioteca nazionale che si sia iniziato a ragionare di Web archiving come infrastruttura ha ancora oggi conseguenze su cosa vogliamo o possiamo conservare. I siti web infatti sono stati da subito considerati come un mezzo attraverso cui diffondere "pubblicazioni" e i produttori delle informazioni alla stregua di "editori", potendo in questo modo interpretare il Web archiving come un'attività senza soluzione di continuità con il deposito legale. Ciò da una parte ha permesso alle biblioteche di salvare milioni di risorse web, dall'altra ha fortemente influenzato i criteri e le modalità di selezione delle risorse da conservare, escludendo o non analizzando adeguatamente tutto ciò che non era riconducibile all'interno di questo canone riconosciuto. Quando ci muoviamo nell'ecosistema del Web, infatti, i confini tra i concetti di "pubblicato", "pubblicamente accessibile", "destinato all'uso pubblico" o "di interesse pubblico" si fanno sempre più sfumati, ma continuano ad esistere, e collassarli nell'aggettivo "pubblico" può certamente rivelarsi comodo ma anche rischioso. Allo stesso modo decidere che tutti coloro che gestiscono una piattaforma o producono contenuti per il Web siano "editori", con tutti i diritti ma soprattutto i doveri che questo status comporta, dovrebbe essere se non altro oggetto di una qualche forma di contrattazione, se non meglio formalmente inquadrato a livello normativo; e non può essere una decisione unilaterale delle istituzioni culturali, pena l'inefficacia di questo as-

<sup>&</sup>lt;sup>13</sup> Hegarty (2022) chiarisce che in questo contesto con infrastrutture si debbano intendere proprio le modalità con cui sono "assemblate" ed utilizzate le singole tecnologie, anche sviluppate da terzi, in una specifica realtà come quella di una biblioteca nazionale, per perseguire le proprie finalità istituzionali.



sunto e l'impossibilità di perseguire le finalità pubbliche per cui quelle stesse istituzioni operano. Inoltre, la scelta di definire i siti web come *medium* per le "pubblicazioni", al fine di ricondurli ad una tipologia bibliografica nota ha, almeno per tutti gli anni '90, limitato molto ciò che è stato conservato. L'impossibilità addirittura, per molto tempo, di descrivere attraverso i tradizionali standard catalografici il formato "sito web" ha comportato il sacrificio di una grandissima quantità di informazioni, quasi come se ciò che non potesse essere descritto, in qualche modo non fosse conservabile.

Solo il cambio di paradigma, che ha portato dalla ricerca di "pubblicazioni" alla ricerca di "informazioni" diffuse sul Web, ha permesso di uscire dalla preistoria del Web archiving e di iniziare a pensare a tali risorse come qualcosa di completamente nuovo e per questo meritevole di uno specifico trattamento.

La Library of Congress, a partire dal 2002, definisce all'interno delle AACR - Anglo-American Cataloguing Rules i siti web come "integrating" resource, 14 continuando ad assimilarli ad oggetti bibliografici noti come le pubblicazioni a dispense o fogli mobili, ma anche potendo finalmente rilevare una delle caratteristiche principali di una risorsa web ovvero il suo, a volte molto frequente a volte meno, aggiornamento. <sup>15</sup> La possibilità di aggiornamento è però solo una delle caratteristiche di queste risorse. Si potrebbe affermare, infatti, che i siti web siano risorse sostanzialmente bibliografiche ma archivistiche nella forma. Riprendendo la terminologia archivistica classica, infatti, un sito web è spesso diretta emanazione del suo soggetto produttore (ente, azienda, persona ecc...), e le sue diverse componenti possono essere individuate singolarmente ma hanno un "vincolo necessario". Allo stesso modo possono essere oggetto di aggiornamenti o modifiche anche minori, che necessitano di un "controllo di versione", ma la cui entità potrebbe non giustificare catalograficamente la definizione di "nuova edizione". D'altra parte i contenuti informativi mediati dai siti possono essere spesso considerati di "interesse culturale" e "destinati all'uso pubblico", riprendendo la definizione che la normativa italiana sul deposito legale<sup>16</sup> fornisce delle risorse destinate alle biblioteche, e agli altri istituti individuati, per la creazione degli archivi (!) nazionali e regionali. Anche l'OCLC – Web Archiving Metadata Working Group (WAM) nel documento Descriptive

\_

<sup>&</sup>lt;sup>14</sup> Nel RDA Glossary una "integrating" resource viene definita: a resource that is added to or changed by means of updates that do not remain discrete and are integrated into the whole (e.g., a loose-leaf manual that is updated by means of replacement pages, a website that is updated continuously). Si veda anche: <a href="https://web.library.yale.edu/cataloging/e-resources/updating-websites">https://web.library.yale.edu/cataloging/e-resources/updating-websites</a>. In Italia, ancora le recenti norme per il "Trattamento in SBN di risorse pubblicate online" non considerano i siti web come oggetti catalografici autonomi ma solo come media per la diffusione di altre pubblicazioni: <a href="https://norme.iccu.sbn.it/index.php?title=Norme comuni/Ulteriori indicazioni e approfondimenti/Trattamento in SBN di risorse pubblicate online</a>. La motivazione è banalmente da ricondursi alla constatazione che nelle istituzioni del patrimonio non si è ancora sviluppata una sufficiente sensibilità su questi temi e che un'attività istituzionale di Web archiving è portata avanti solo dalla Biblioteca Nazionale Centrale di Firenze. Su questo tema si veda anche (Allegrezza 2023).

<sup>&</sup>lt;sup>15</sup> Anche la tradizionale modalità di accesso alle risorse archiviate tramite Wayback Machine rispecchia la visione di un archivio web come di una collezione di pubblicazioni in evoluzione: "this imagining of a web archive as a collection of 'evolving' publications endures in the many web archives that use Wayback for the replay of content, where access is given at the level of a single 'title' (webpage) along with a list of 'issues' (snapshots)" (Hegarty 2022).

<sup>&</sup>lt;sup>16</sup> L. 15 aprile 2004, n. 106 *Norme relative al deposito legale dei documenti di interesse culturale destinati all'uso pubblico*: <a href="http://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:2004-04-15;106!vig=2023-01-11">http://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:2004-04-15;106!vig=2023-01-11</a> e D.P.R. 3 maggio 2006, n. 252, Regolamento recante norme in materia di deposito legale dei documenti di interesse culturale destinati all'uso pubblico: <a href="http://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.del.presidente.della.repubblica:2006-05-03;252!vig=2023-01-11">http://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.del.presidente.della.repubblica:2006-05-03;252!vig=2023-01-11</a>.



Metadata for Web Archiving<sup>17</sup>, contenente alcune indicazioni per la formulazione di metadati descrittivi per le risorse catturate dal Web, propone una soluzione "ponte" tra i due più tradizionali approcci della descrizione archivistica e della catalogazione bibliografica/bibliotecaria. Lo standard di struttura, il Dublin Core, è neutrale rispetto al dominio di applicazione e il set minimo di elementi è stato scelto mettendo a confronto i principali standard in uso per la descrizione degli oggetti digitali di interesse per gli istituti della cultura e della ricerca: MARC21, EAD, MODS, Dublin Core e Schema.org. Inoltre la soluzione proposta può essere utilizzata a tutti i livelli descrittivi, in maniera scalabile, da un dominio di alto livello al singolo documento, in qualsivoglia formato, pubblicato su una pagina web.

Si potrebbe continuare ancora a lungo con esempi che avvalorino il fatto, per altro mai messo in dubbio fuori dalle istituzioni del patrimonio, che il Web sia un media informativo completamente diverso da quelli finora trattati e che la sua gestione resti un'attività di confine tra le competenze degli archivi e delle biblioteche, ma pare maggiormente proficuo provare da ciò a trarre delle proposte operative. Un tale *excursus* storico serve infatti a ribadire quanto siano necessari da una parte una fattiva collaborazione tra istituzioni della memoria, in particolare archivi e biblioteche, e dall'altra una normativa nazionale che tratti in maniera organica il tema della conservazione del patrimonio digitale definendo le responsabilità, anche dei produttori di informazione, gli ambiti di esclusività e quelli di collaborazione. Oltre che definendo in concreto i rapporti tra le necessità di conservazione per finalità storiche, culturali e di ricerca e i sempre più pressanti interrogativi etici riguardanti ad es. il diritto all'oblio.

Infine, come si vedrà meglio di seguito, pare altrettanto importante formare gli utenti ad un uso consapevole degli archivi web:

Reflecting on web archiving as a story of continuity as well as discontinuity, this article is a call for a greater account of the strengths and limitations of web archiving infrastructure's "installed base"—the major public library—and the myriad artefacts collected throughout these libraries' histories. This can help users of these sources to approach web archives not as a neutral window into the web's past, but a continuation of an inherently partial attempt to mobilise a representation of publics in archival form (Hegarty 2022).

#### In caso di conflitto

In seguito alla drammatica riconquista dell'Afghanistan da parte dei talebani nell'agosto del 2021, il Partito Repubblicano statunitense ha cancellato un post, pubblicato sul proprio sito il 15 settembre 2020, in cui venivano elogiati i risultati dell'accordo promosso dall'ex presidente Trump con i talebani stessi. Grazie alla *Wayback Machine* di Internet Archive è possibile vedere diverse istantanee del sito antecedenti all'agosto del 2021 e leggere l'articolo successivamente eliminato. Il ruolo di un archivio web in caso di conflitto non si riduce, però, a quello di mostrare l'incoerenza del politico o del personaggio pubblico di turno che, come spesso accade, modifica o cancella post scomodi. Come è purtroppo noto, infatti, il restaurato governo talebano ha da subito iniziato

https://www.oclc.org/research/publications/2018/oclcresearch-descriptive-metadata.html.

https://web.archive.org/web/\*/https://gop.com/president-trump-is-bringing-peace-to-the-middle-east-rsr/.



una campagna persecutoria nei confronti di molti cittadini afghani accusandoli di "collaborazionismo" con il governo statunitense, ad esempio perché impiegati all'interno delle istituzioni americane o in associazioni umanitarie sul territorio. I talebani hanno passato al vaglio centinaia di profili social alla ricerca di post, foto ecc...che avvalorassero tali scellerate accuse. Le piattaforme social, per contro, hanno tempestivamente oscurato, su richiesta degli interessati, i profili e ogni altra traccia online (Tuttosport.com. 2021) potenzialmente dannosa per la loro incolumità. Ma ciò rischiava di non essere sufficiente. In particolar modo Twitter¹9 ha richiesto la collaborazione di Internet Archive per rendere inaccessibili anche le copie archiviate dei profili degli afghani perseguitati, almeno fino a quando la situazione interna del Paese non ne consentirà il ripristino. L'inaccessibilità di queste informazioni, per ora a tempo indeterminato, è un esempio di come giustamente le ragioni della incolumità fisica delle persone possano e debbano essere poste in primo piano rispetto al diritto di cronaca o alla ricerca storica.

D'altronde che il Web e, di conseguenza, gli archivi web siano oramai da considerarsi, da molteplici punti di vista, infrastrutture strategiche per un Paese alla stregua di strade, ponti o ospedali, lo dimostra anche il drammatico conflitto in corso a seguito dell'invasione russa dell'Ucraina nel febbraio del 2022. Nell'ultimo anno, all'interno di una complessiva strategia di innalzamento globale della tensione, sono aumentati esponenzialmente gli attacchi hacker russi non solo contro le infrastrutture di rete ucraine ma in tutto il mondo, così come gli utilizzi impropri di piattaforme e servizi web da parte delle forze dell'Intelligence del Cremlino.<sup>20</sup> Fortunatamente, sul versante opposto, fin dalle primissime fasi del conflitto, si è costituito un movimento volontario internazionale per la salvaguardia del patrimonio culturale online dell'Ucraina, che ha salvato fino ad ora oltre 50 TB di dati. L'iniziativa denominata "SUCHO - Saving Ukrainian Cultural Heritage Online"21 non è interessante solo per la mole di informazioni immagazzinate in un lasso di tempo molto breve, ma anche perché l'insieme di procedure messe a punto potrebbero costituire uno standard di fatto nella salvaguardia del patrimonio culturale digitale in caso di conflitto o di altro tipo di emergenza.<sup>22</sup> Inoltre, anche se forse è l'aspetto più scontato, le attività di Web archiving servono a salvare il racconto, o meglio, i racconti che di quel conflitto vengono fatti. Per questo tutte le maggiori istituzioni culturali nel mondo hanno creato collezioni di risorse web di interesse nazionale riguardanti la Guerra in Ucraina, come è uso per molti grandi eventi non per forza di carattere catastrofico, come le pandemie, i Giochi Olimpici, le elezioni governative o le scoperte scientifiche. In Italia il deposito legale delle risorse digitali diffuse tramite rete informatica, di cui il Web archiving nelle modalità finora descritte è solo uno delle componenti, non è obbligatorio. Ciò comporta che le istituzioni della memoria debbano ottenere un preventivo consenso alla copia e all'archiviazione delle pagine web. Anche a causa delle scarse risorse a disposizione delle biblioteche, le procedure di richiesta di tali autorizzazioni si rivelano spesso lunghe e infruttuose: non solo i permessi che si riescono ad ottenere sono quantitativamente poco significativi, ma questi

<sup>&</sup>lt;sup>19</sup> Non solo la società Twitter collabora da anni con le istituzioni del patrimonio per consentire di conservare gli archivi di tweet di rilevanza pubblica ma, a differenza di altri Social, tutti i suoi profili sono "pubblici" almeno per gli iscritti: ciò rende Twitter la piattaforma social più facilmente archiviabile e, di fatto, la più archiviata.

<sup>&</sup>lt;sup>20</sup> A questo proposito interessante la serie di articoli a tema proposti sul blog Guerre di Rete: <a href="https://www.guerredirete.it/">https://www.guerredirete.it/</a>.

<sup>21</sup> <a href="https://www.sucho.org/">https://www.sucho.org/</a>.

<sup>&</sup>lt;sup>22</sup> Con un tweet del 24 agosto 2022 sul proprio profilo, il gruppo SUCHO ha annunciato di stare lavorando alla redazione di un "*Handbook of Emergency Web Archiving*": https://twitter.com/sucho\_org/status/1562484110802595841.



possono arrivare anche con poca tempestività rispetto all'esigenza di archiviare contenuti per loro natura soggetti a continue modifiche.<sup>23</sup> Il risultato sono collezioni di risorse web archiviate non in grado di rappresentare compiutamente la pluralità di aspetti e punti di vista di cui si compone un evento così drammatico come una guerra.

Altre situazioni conflittuali, seppure non riconducibili a scenari di guerra reale, possono infine fungere da esempio di come possa essere fatto anche un uso strumentale dannoso delle url archiviate:

Archiving services serve a variety of purposes beyond addressing link rot. Platforms like archive.is are reportedly used to preserve controversial blogs and tweets that the author may later opt to delete (Mondal et al. 2016). Moreover, they also reduce Web traffic toward "source URLs" when the original content is still accessible, thus depriving them of potential ad revenue streams (users do not visit the original site, but just the archived copy). In fact, anecdotal evidence has emerged that alt-right communities target outlets they disagree with by nudging their users to share archive URLs instead (Koebler 2014), or discrediting them by pointing at earlier versions of articles (Ralph 2017; Zannettou et al. 2018).

Zannettou et al. (2018) hanno dimostrato che l'utilizzo di url archiviate al posto di quelle originali ancora attive può essere scelto appositamente per diminuire il traffico in entrata verso siti o contenuti che non si vogliono sostenere, o per rimandare volontariamente il lettore a contenuti che l'autore ha scelto di modificare o cancellare, con intenti quindi deliberatamente dannosi. Seppure non sia pensabile poter controllare del tutto l'utilizzo *ex-post* degli archivi, una corretta gestione del patrimonio web dovrebbe considerare anche questi aspetti. La disponibilità di metadati completi in grado di fornire informazioni sul contesto dell'archiviazione, su tutte le fasi del processo di conservazione nonché sull'accesso alle informazioni e sui responsabili delle singole procedure è sicuramente un disincentivo ad un loro uso improprio.

## "Archive first, ask questions later"

L'Archive Team si autodefinisce "a loose collective of rogue archivists, programmers, writers and loudmouths dedicated to saving our digital heritage". <sup>24</sup> Dal 2009, spesso in collaborazione con Internet Archive, il collettivo salva siti e servizi web a rischio di chiusura. <sup>25</sup> Tra i suoi progetti che hanno avuto maggiore risonanza mediatica c'è sicuramente quello dell'archiviazione di alcune migliaia di domande e relative risposte fatte dagli utenti di *Yahoo Answers*, lo storico servizio di web forum di Yahoo! chiuso dall'azienda nella primavera del 2021.

<sup>25</sup> L'elenco dei progetti dell'Archive Team suddivisi per categoria è visionabile al link: <a href="https://wiki.archiveteam.org/index.php/Category:Projects\_status">https://wiki.archiveteam.org/index.php/Category:Projects\_status</a>.

<sup>&</sup>lt;sup>23</sup> La Biblioteca Nazionale Centrale di Firenze, nell'ambito del proprio servizio di Web archiving, ha avviato a marzo 2022 una campagna di raccolta delle risorse web italiane relative al conflitto in Ucraina, tuttavia la generale adesione è stata scarsa: <a href="https://www.bncf.firenze.sbn.it/risorse-web-italiane-sul-conflitto-in-ucraina/">https://www.bncf.firenze.sbn.it/risorse-web-italiane-sul-conflitto-in-ucraina/</a>.

<sup>&</sup>lt;sup>24</sup> https://wiki.archiveteam.org/.

<sup>&</sup>lt;sup>26</sup> L'obiettivo dichiarato dell'Archive Team era quello di archiviare le 84 milioni di domande i cui link erano mostrati nella Sitemap del sito ma per ora sono accessibili tramite la Wayback Machine di IA soltanto alcune migliaia di domande: <a href="https://www.org/web/\*/https://answers.yahoo.com/question/\*">https://answers.yahoo.com/question/\*</a> (Wodinsky and Mehrotra 2021).



Il motto dell'Archive Team "Archive first, ask questions later", come rilevato in (Ogden 2022), consente a questo gruppo di volontari di effettuare attività di archiviazione con una efficienza e tempestività il più delle volte impossibili da eguagliare per le istituzioni culturali pubbliche, che operano necessariamente all'interno di limiti imposti dalla normativa vigente e da uno specifico mandato:

For better or worse, this action-oriented "brute force" approach has enabled AT to proceed where institutions like national web archives are subject to their own mandates and legislative environments that constrain the nature of what can be collected, stored and made accessible (Ogden 2022, 120).

Anche Internet Archive non chiede un permesso preventivo all'archiviazione ma, rispetto all'Archive Team, ha un approccio che si potrebbe definire meno aggressivo e un apposito servizio di *opt-out* attraverso il quale chiunque può richiedere la rimozione dei contenuti archiviati in archive. org.<sup>27</sup> L'archiviazione senza l'acquisizione di un preventivo consenso, infatti, non può essere semplicemente liquidata come una sorta di "male necessario", potendosi in talune circostanze considerare un vero e proprio abuso dal punto di vista del diritto d'autore o del trattamento dei dati, come dimostra il caso Tumblr:

The Tumblr case highlights a mutable AT collective that (though committed to the tenets of practice) is both open to modification and willing to negotiate the often fuzzy boundaries that define what constitutes the "public Web" in web archiving (Ogden 2022, 127).

La piattaforma di microblogging Tumblr, lanciata nel 2007, è sempre stata molto popolare, arrivando a contare 455 milioni di blog, grazie soprattutto alle politiche relativamente poco restrittive in termini di contenuti pubblicabili, che comprendevano anche quelli "per adulti" o "sensibili".<sup>28</sup> Tuttavia, nel 2018, per adeguarsi alle nel frattempo mutate leggi statunitensi, Tumblr ha effettuato una stretta in questo ambito, annunciando la cancellazione dei contenuti non più ritenuti idonei. Il rischio di una massiva perdita di dati era dunque concreto non solo per la loro cancellazione da parte dell'azienda, ma anche per il volontario esodo e conseguente chiusura dei profili da parte di utenti che si erano sentiti traditi dal cambio di rotta della piattaforma. In risposta, nel dicembre dello stesso anno, l'Archive Team avvia il Tumblr NSFW project, chiamando a raccolta i volontari per individuare le url da archiviare, ma anche mettendo a disposizione strumenti di facile utilizzo per consentire una più ampia partecipazione all'archiviazione. Il progetto è stato fortemente ostacolato da Tumblr, che ha lamentato un vero e proprio "attacco" ai suoi sistemi. Inoltre l'archiviazione di contenuti che sono oggetto di deliberata cancellazione per cause legali e etiche può, a sua volta, essere una scelta discutibile almeno dal punto di vista morale. L'Archive Team, a questo proposito, si è sempre dichiarata "neutrale" rispetto alla tipologia di informazioni archiviate, sostenendo che tutto ciò che è su Internet almeno teoricamente potrebbe e dovrebbe essere salvato, per fornire un'immagine completa e diversificata dell'esperienza online in un determinato tempo

<sup>28</sup> In Italia sotto la definizione di "contenuti sensibili" rientrano comunemente quelli violenti e pornografici, ma anche quelli istiganti al consumo di alcol e droghe o promuoventi comportamenti generalmente dannosi per il singolo o la collettività. In ambito anglofono gli stessi contenuti sono identificati con l'acronimo NSFW – Not Safe For Work.

<sup>&</sup>lt;sup>27</sup> https://help.archive.org/help/how-do-i-request-to-remove-something-from-archive-org/.



e contesto (Ogden 2022). A questo punto appaiono però ineludibili alcune domande. Innanzi tutto su cosa si debba intendere per "neutralità". In un contesto come quello della formazione di un archivio, della costituzione di un patrimonio, ogni scelta è comunque una scelta politica. Inoltre può rivelarsi pericoloso perseguire una presunta neutralità durante un processo di conservazione a lungo termine, anche alla luce del fatto che non tutti gli aspetti di questo processo sono realmente controllabili o dipendenti dalla volontà di chi conserva. Forse, più che della neutralità, c'è bisogno di ribadire l'importanza della trasparenza delle scelte e delle procedure adottate dai diversi soggetti, pubblici o privati, che operano nel dominio della costruzione del patrimonio digitale. Garantire cioè che, in ogni momento, chi accede all'archivio possa avere contemporaneo accesso alle informazioni sulla sua costituzione: il contesto in cui le informazioni sono state prodotte, quali sono stati i criteri di selezione e chi li ha determinati e applicati.

AT is actively shaping access to dead and dying platforms, as well as creating a community of practice centred on the preservation of access and "rogue archiving" strategies for saving the Web. Despite their use of common tools and standards, these practices will be seen in stark contrast to other risk-averse approaches to web archiving taken by conventional memory institutions, and community archives projects that centre an ethics of care for content creators and future users. AT's interventions simultaneously illustrate the possibilities of participatory web archiving at scale and the potential risks of such approaches in the face of platform resistance, rights and privacy concerns. Given the scale of AT's collecting activities and their impact on the coverage of the IAWM, understanding their practices offers insights into how the Web is transformed through web archiving, as well as their critical ethical implications for how these platforms are studied in future (Ogden 2022, 129).

Il caso citato della Biblioteca Nazionale Centrale di Firenze fortemente limitata nella possibilità di archiviare le risorse web nazionali relative al conflitto in Ucraina, a paragone con il progetto *SUCHO*, risulta quindi emblematico e non isolato. Proprio la scarsità di strumenti legali a disposizione, oltre ad una certamente poco diffusa consapevolezza sul tema, è una delle cause principali per cui su una questione tanto importante come la salvaguardia del patrimonio culturale digitale, durante una grave emergenza come quella di un conflitto armato, sia intervenuto per primo un gruppo costituitosi spontaneamente<sup>29</sup> e non gli organismi tradizionalmente a ciò deputati.<sup>30</sup>

È pertanto probabilmente arrivato il momento di ragionare su provvedimenti normativi che prevedano anche per le istituzioni pubbliche italiane la possibilità di procedere unilateralmente all'archiviazione di risorse web di interesse storico-culturale per ben determinate finalità. Risoluzioni se non valide in generale,<sup>31</sup> almeno in casi definiti di emergenza o urgenza, in cui l'intempestività

<sup>29</sup> Il progetto *SUCHO* ha avuto anche il sostegno e la collaborazione di importanti istituzioni e associazioni della cultura, come l'ULA – Ukrainian Library Association: <a href="https://www.sucho.org/partners">https://www.sucho.org/partners</a>.

<sup>&</sup>lt;sup>30</sup> La Convenzione per la protezione di beni culturali in caso di conflitto armato, firmata per la prima volta all'AJA nel 1954 e successivamente modificata ed ampliata, ad esempio, non prende ancora in alcuna considerazione il patrimonio culturale digitale: <a href="https://www.unesco.beniculturali.it/english-convenzione-dellaja-1954/">https://www.unesco.beniculturali.it/english-convenzione-dellaja-1954/</a>.

<sup>&</sup>lt;sup>31</sup> A livello internazionale il complesso di norme entro cui si collocano le attività di Web archiving "istituzionale" è variegato ma riconducibile sostanzialmente a due scenari: il Web e Social media è regolato all'interno del deposito legale oppure è considerato un'attività lecita nell'ambito delle eccezioni al diritto d'autore per finalità culturali (interesse scientifico e storico): <a href="https://netpreserve.org/web-archiving/legal-deposit/">https://netpreserve.org/web-archiving/legal-deposit/</a>.



e quindi l'incompletezza della raccolta arrecherebbe grave danno al patrimonio digitale costituito o in costituzione. Ferma restando la facoltà dei produttori di informazioni, laddove la "pubblicità" di queste ultime non fosse pacifica, di impedire o limitare preventivamente la raccolta, con un utilizzo corretto di strumenti quali i file robots.txt³²e le sitemap,³³ oppure di richiederne a posteriori una motivata cancellazione dagli archivi.

## Risorsa in archivio

Definire un modello di governance per la creazione e la tutela del patrimonio web, valutando adeguatamente le risorse necessarie ad attuarlo, è fondamentale dunque non solo per non perdere la nostra memoria recente ma anche perché come sceglieremo di farlo influenzerà in molti modi l'immagine che di noi avranno nel futuro. Il primo passo, come si è visto, dovrebbe essere quello di definire le responsabilità sull'intero processo, anche e soprattutto con appositi atti normativi: in breve, chi deve conservare e chi è tenuto ad assicurarsi che i propri contenuti siano conservati. L'individuazione delle istituzioni pubbliche deputate alle attività di Web a Social media archiving dovrebbe avvenire abbandonando le tradizionali categorie di "risorsa bibliografica" o "archivistica" o "museale", ridistribuendo gli ambiti di competenza in base alla tipologia di informazione veicolata più che alla forma delle risorse o a chi le ha prodotte. Non dimenticando di delineare gli ambiti in cui la cooperazione interistituzionale si configuri come la soluzione più efficace ed efficiente per il raggiungimento degli obiettivi di conservazione e accesso nel lungo periodo. Cooperazione interistituzionale non solo orizzontale, a copertura delle risorse di diversa natura, ma anche verticale centro-periferia tra istituti appartenenti allo stesso dominio, al fine di intercettare dati di interesse locale o specializzato (Storti 2023). La fase successiva riguarda il coinvolgimento delle comunità, intese sia come gruppi portatori di una storia o di un interesse comune, per individuare e preservare le "memorie particolari", sia come comunità di esperti e attivisti che insieme lavorano per salvaguardare l'infrastruttura cooperativa per eccellenza: il Web. Non trattandosi solo di stabilire una collaborazione, come già avviene da tempo, con enti privati come Internet Archive, ma anche di tentare di ricondurre l'azione di gruppi come l'Archive Team all'interno di una strategia comune. Certo potrebbero esistere dei limiti alla possibilità di stabilire rapporti con gruppi autocostituitisi, sia per ragioni legali che di pura volontà dipendenti da una differente visione di cosa e come dovrebbe essere conservato e reso accessibile, ma anche continuare a progettare infrastrutture che non tengano in alcuna considerazione l'esistenza di queste importanti iniziative non pare producente. Ineludibile, inoltre, l'attivazione di politiche di sensibilizzazione nei confronti dei produttori delle informazioni. Sensibilizzazione sia sul tema generale della archiviabilità<sup>34</sup> delle risorse sia, in particolare per i contenuti diffusi tramite Social media, sull'uso consapevole delle piattaforme terze. Queste ultime, infatti, quasi sempre detengono legalmente l'esclusiva facoltà di decidere se e quando cancellare contenuti, renderli inaccessibili o impedirne l'archiviazione. Per

<sup>32</sup> https://it.wikipedia.org/wiki/Protocollo di esclusione robot.

<sup>33 &</sup>lt;a href="https://it.wikipedia.org/wiki/Sitemap">https://it.wikipedia.org/wiki/Sitemap</a>. È bene precisare che entrambi gli strumenti appena citati possono imporre limiti "logici" ma non "fisici" alla copia, trattandosi di semplici informazioni di tipo testuale che i *crawler* possono o meno tenere in considerazione.

<sup>&</sup>lt;sup>34</sup> Per approfondimenti: https://www.bncf.firenze.sbn.it/biblioteca/archiviabilita-dei-siti-web/.



questo motivo, oltre che formare alle pratiche di auto-archiviazione dei profili personali ma anche di quelli istituzionali e aziendali, è necessario parlare dell'importanza della differenziazione dei canali di distribuzione dei contenuti, e della possibilità di pubblicare informazioni con licenze aperte e su piattaforme altrettanto "aperte". In ultimo non bisogna tralasciare una riflessione sulle piattaforme di accesso agli archivi web e l'educazione degli utenti finali al loro utilizzo, nella consapevolezza che la creazione di un patrimonio non è mai un processo del tutto neutrale ma, al contrario, condizionato da numerosi fattori come i vincoli legali ed etici, i limiti e le possibilità delle tecnologie, nonché la sensibilità storico-culturale di chi ne ha la responsabilità.

Nel 2021, in occasione del suo 25° anniversario, Internet Archive ha pubblicato una utopica *Wayforward Machine*<sup>36</sup>che mostra come potrebbe essere il Web del 2046. L'immagine restituita è quella di un mondo in cui avere accesso ad informazioni affidabili in rete potrebbe essere né semplice, né libero, né gratuito:

Imagining utopia: #EmpoweringLibraries

We can find a better way forward.

The Internet Archive is facing a lawsuit by a cartel of corporate publishers that threatens the age-old right of libraries to buy, preserve and lend materials to the public.

To fight for a world where libraries and learners are empowered through access to information, join our #EmpoweringLibraries campaign.<sup>37</sup>

Scongiurare un futuro che può apparire nel complesso certamente distopico, ma che per molti aspetti rispecchia scenari plausibili, è però nelle possibilità delle istituzioni della memoria purché siano in grado di agire in sinergia con tutti gli attori coinvolti nel processo e all'interno di una strategia complessiva di gestione del patrimonio culturale digitale.

<sup>&</sup>lt;sup>35</sup> Il caso recente dell'esodo da Twitter verso Mastodon, una piattaforma social decentralizzata, è emblematico in questo senso. Molti utenti, preoccupati dalla parzialità nel trattamento, sia a breve che a lungo termine, dei propri dati da parte dell'azienda recentemente acquisita dal magnate Elon Musk, hanno non solo iniziato a scaricare il proprio archivio ma anche a spostarsi su una piattaforma che sia a livello legale che tecnologico sia resiliente rispetto a singole iniziative potenzialmente dannose. A questo proposito: <a href="https://www.ilpost.it/2022/11/21/archivio-dati-twitter/">https://www.ilpost.it/2022/11/21/archivio-dati-twitter/</a>.

<sup>36</sup> https://wayforward.archive.org/.

<sup>37</sup> https://wayforward.archive.org/ia2046/.



## Riferimenti bibliografici

Allegrezza, Stefano. 2023. "Web e social media come nuove fonti per la storia." *Umanistica Digitale*, January, 137-162. <a href="https://doi.org/10.6092/ISSN.2532-8816/15665">https://doi.org/10.6092/ISSN.2532-8816/15665</a>.

Bergamin, Giovanni. 2006. "La raccolta dei siti web: un test per il dominio 'punto it." *DigItalia* 1 (2): 170-74. <a href="http://digitalia.sbn.it/article/view/306">http://digitalia.sbn.it/article/view/306</a>.

Il Post. 2022. "Come richiedere l'archivio di tutto ciò che avete fatto su Twitter." November 21, 2022. https://www.ilpost.it/2022/11/21/archivio-dati-twitter/.

Hegarty, Kieran. 2022. "The Invention of the Archived Web: Tracing the Influence of Library Frameworks on Web Archiving Infrastructure." *Internet Histories*, July, 1–20. <a href="https://doi.org/10.1">https://doi.org/10.1</a> 080/24701475.2022.2103988.

International Institute of Social History. n.d. "Occasio Digital Social History Archive." Text. Occasio Digital Social History Archive. Accessed January 6, 2023. https://iisg.nl/occasio/index.php.

"List of Web Archiving Initiatives." 2023. In *Wikipedia*. <a href="https://en.wikipedia.org/w/index.php?title=List">https://en.wikipedia.org/w/index.php?title=List</a> of Web archiving initiatives&oldid=1132919882.

National Library of Australia. 2023. *PANDORA*, *Australia's Web Archive*. Accessed January 6. <a href="https://pandora.nla.gov.au/">https://pandora.nla.gov.au/</a>.

OCLC – Web Archiving Metadata Working Group (WAM). 2018. 'Descriptive Metadata for Web Archiving'. <a href="https://www.oclc.org/research/publications/2018/oclcresearch-descriptive-metadata.html">https://www.oclc.org/research/publications/2018/oclcresearch-descriptive-metadata.html</a>.

Ogden, Jessica. 2022. "Everything on the Internet Can Be Saved': Archive Team, Tumblr and the Cultural Significance of Web Archiving." *Internet Histories* 6 (1–2): 113–32. <a href="https://doi.org/10.108">https://doi.org/10.108</a> 0/24701475.2021.1985835.

Parlangeli, Diletta. 2017. "Errore 404: la storia della 'Pagina non trovata' più frequente di Internet." *Wired* (blog), December 7, 2017. <a href="https://www.wired.it/internet/web/2017/12/07/errore-404-la-storia-not-found/">https://www.wired.it/internet/web/2017/12/07/errore-404-la-storia-not-found/</a>.

Signorelli, Andrea Daniele, and VOIS. 2022. *Errore 404, così il web sta perdendo i pezzi*. CRASH – La chiave per il digitale. https://open.spotify.com/episode/7iYIin0mYnzkCKporF9pjN.

Storti, Chiara. (In press). "Community Webs. Il Web Archiving per La Creazione e l'accesso Permanente Alle Collezioni Di Interesse Locale Nell'ecosistema Del Web." In *Atti Del Convegno* "Le Collezioni in Biblioteca: Nuovi Approcci per Un Elemento Di Importanza Strategica", Bolzano – Eurac Research, 21 Ottobre 2022. Collana Sezioni Regionali AIB. Trentino-Alto Adige. AIB – Associazione Italiana Biblioteche.

Tuttosport.com. 2021. "Ecco come i social network stanno proteggendo gli utenti afghani." August 23, 2021. <a href="https://tuttosport.com/news/attualit/cronaca/2021/08/23-84728825/ecco come i social network stanno proteggendo gli utenti afghani.">https://tuttosport.com/news/attualit/cronaca/2021/08/23-84728825/ecco come i social network stanno proteggendo gli utenti afghani.</a>

Wodinsky, Shoshana, and Dhruv Mehrotra. 2021. "We're Archiving Yahoo Answers So You'll Always Know How Babby Is Formed." *Gizmodo*, September 4, 2021. <a href="https://gizmodo.com/were-archiving-yahoo-answers-so-youll-always-know-how-b-1846643969">https://gizmodo.com/were-archiving-yahoo-answers-so-youll-always-know-how-b-1846643969</a>.



Yale University Library. n.d. "Cataloging Online Integrating Resources." Accessed January 10, 2023. <a href="https://web.library.yale.edu/cataloging/e-resources/updating-websites">https://web.library.yale.edu/cataloging/e-resources/updating-websites</a>.

Zannettou, Savvas, Jeremy Blackburn, Emiliano De Cristofaro, Michael Sirivianos, and Gianluca Stringhini. 2018. "Understanding Web Archiving Services and Their (Mis)Use on Social Media." *Proceedings of the International AAAI Conference on Web and Social Media* 12 (1). <a href="https://doi.org/10.1609/icwsm.v12i1.15018">https://doi.org/10.1609/icwsm.v12i1.15018</a>.



# Geodata. Spatial data in metadata creation

## Klaus Kempf (a)

a) Independent consultant – formerly Bayerische Staatsbibliothek, https://orcid.org/0000-0003-3674-5437

Contact: Klaus Kempf, <u>klauskempf@gmx.de</u>
Received: 22 December 2022; Accepted: 26 February 2023; First Published: 15 May 2023

#### **ABSTRACT**

In recent years the demand for geoinformation and the use of GIS (Geographic Information System) applications by different subject areas has increased enormously. Libraries are attempting to respond to these needs through the provision of adequately developed spatio-temporal information, the digitisation of historical cartographic and geographic collections, and the enrichment of the metadata associated with digital resources with geodata that allow for proper georeferencing. After defining some key concepts related to the assignment of terrestrial geocoordinates in contemporary times, the article illustrates the potential of their use in the bibliographical domain with particular emphasis on their use by German-speaking libraries.

#### **KEYWORDS**

Cartographic Resources; Terrestrial geocoordinates; Georeferencing; German Libraries; Metadata.

# Geodati. Dati spaziali nella metadatazione

#### **ABSTRACT**

Negli ultimi anni sono aumentate enormemente la richiesta di geoinformazioni e l'uso di applicazioni GIS (Geographic Information System) da parte dei diversi settori disciplinari. Le biblioteche stanno cercando di rispondere a queste esigenze attraverso un'offerta di informazioni spazio-temporali adeguatamente sviluppate, la digitalizzazione di fondi storici cartografici e geografici e l'arricchimento dei metadati associati alle risorse digitali con geodati che consentono una corretta georeferenziazione. L'articolo dopo aver definito alcuni concetti chiave per comprendere le modalità di assegnazione contemporanea delle geocoordinate terrestri, illustra le potenzialità derivanti dal loro impiego in ambito bibliografico con particolare attenzione al loro uso da parte delle biblioteche di lingua tedesca.

#### PAROLE CHIAVE

Risorse Cartografiche; Geocoordinate Terrestri; Georeferenziazione; Biblioteche tedesche; Metadatazione.

Desidero ringraziare Gabriele Messmer ed Esther Scheven per la collaborazione nella stesura del saggio; Mauro Guerrini e Laura Manzoni per la revisione del testo in italiano.

<sup>© 2023,</sup> The Author(s). This is an open access article, free of all copyright, that anyone can freely read, download, copy, distribute, print, search, or link to the full texts or use them for any other lawful purpose. This article is made available under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. JLIS.it is a journal of the SAGAS Department, University of Florence, Italy, published by EUM, Edizioni Università di Macerata, Italy, and FUP, Firenze University Press, Italy.



# Introduzione: aumento della domanda e dell'offerta di geodati

Negli ultimi anni è aumentata sensibilmente la richiesta, da parte di molte discipline accademiche, dell'implementazione di applicazioni per la geoinformazione e la ricerca con componenti GIS (Geographic Information System). La domanda, assai eterogenea in termini di dati e servizi richiesti, viene soddisfatta dalle biblioteche, che stanno completando la loro trasformazione verso il digitale, con un'offerta di informazioni spazio-temporali adeguatamente sviluppate che, però, solo parzialmente risponde alle richieste degli utenti (Manzoni 2022). La digitalizzazione di fondi storici con contenuto geografico o con rilevanza geografica, sotto forma di creazione e fornitura di copie digitali è insufficiente per applicare con successo le odierne possibilità della scienza dell'informazione, come i metodi di text mining o data mining. Durante il processo di digitalizzazione, infatti, i dati e i metadati della risorsa devono essere arricchiti con i geodati mediante la cosiddetta georeferenziazione, in modo da soddisfare con la massima precisione possibile i bisogni dell'utente. Le biblioteche, gli archivi e le altre istituzioni della memoria registrata hanno un grande interesse verso i geodati applicati ai propri servizi informativi (per esempio, portali culturali regionali come bavarikon.de (Müller und Schmunk 2019, 23-29). L'ampia gamma di geodati richiesti e realmente a disposizione, tuttavia, spesso porta a confondere i termini e, quindi, come ci ricorda il sempre valido brocardo ante omnia explicatio terminorum, è utile precisare il significato delle parole che verranno usate.

## Definizioni terminologiche

I geodati sono il risultato della georeferenziazione di una risorsa o di un oggetto. La relativa voce di *Wikipedia* elenca quattro fattispecie:<sup>1</sup>

- I. l'individuazione di un indirizzo postale (codifica dell'indirizzo postale);
- II. l'individuazione di una coordinata geografica (geocodifica, Geotagging, Geo-imaging);
- III. l'indicazione di una trasformazione (geocodifica implicita, calibrazione sulla mappa, orientamento interno ed esterno);
- IV. l'applicazione di una trasformazione (geocodifica esplicita, rettifica, rettificazione).

I paragrafi seguenti spiegano sinteticamente l'assegnazione e l'utilizzo delle cosiddette geo-coordinate terrestri.<sup>2</sup>

## Le coordinate geografiche<sup>3</sup>

"In geodesia le coordinate geografiche sono valori utili per individuare la posizione di un punto sulla superficie terrestre. Esse sono la latitudine, la longitudine e l'altitudine. La latitudine è la distanza angolare d'un punto dall'equatore, mentre la longitudine è la distanza angolare di un punto da un arbitrario meridiano di riferimento lungo lo stesso parallelo del luogo (misurati in gradi). Dal 1884

¹ http://de.wikipedia.org/wiki/Georeferenzierung, vedi anche https://it.wikipedia.org/wiki/Georeferenziazione.

<sup>&</sup>lt;sup>2</sup> L'articolo analizzerà la situazione nei paesi di lingua tedesca.

<sup>&</sup>lt;sup>3</sup> Le definizioni sono tratte da Wikipedia, https://it.wikipedia.org/wiki/Coordinate\_geografiche.



il meridiano fondamentale di riferimento è convenzionalmente fissato a Greenwich. La sua longitudine è quindi 0. L'altitudine è la distanza, misurata lungo la verticale del punto considerato sulla superficie terrestre, dal livello del mare". Le latitudini e le longitudini sono grandezze angolari e come tali sono misurate in gradi. Storicamente, l'ordine con cui si indicano le coordinate è sempre lo stesso, prima la latitudine e poi la longitudine, usando diversi formati per scrivere i gradi.

#### Gradi minuti secondi (DMS)

Ogni luogo viene espresso in base sessagesimale. Esempio: le coordinate del Colosseo sono N 41° 53' 24" E 12° 29' 32". Talvolta, per fornire informazioni più precise, pur utilizzando la notazione DMS (Degrees Minutes Seconds), i secondi vengono espressi in formato decimale. Per esempio, N 41°53'24.8280 E 12°29'32.0136.

#### Gradi minuti decimali (DM, Degrees Minutes)

Esempio: le suddette coordinate diventano 41°53.41380′, 12°29.53356′ oppure 41d 53.41380m, 12d 29.53356m.

#### Gradi decimali (DD Decimal Degrees)

Di solito da 4 a 6 cifre decimali. Esempio: le suddette coordinate diventano 41.8902300°, 12.4922260°.

L'indicazione degli emisferi N (Nord) / S (Sud) ed E (Est) / O (Ovest) può essere sostituita dal segno relativo alla direzione. In particolare, avremo valori negativi per latitudini nell'emisfero Sud e longitudini a Ovest del meridiano fondamentale. Si usa sempre più l'ordine longitudine-latitudine, per uniformarsi ai sistemi UTM (Universal Transverse Mercator) e MGRS (Military Grid Reference System).

Essendo la Terra un corpo irregolare (geoide), fornire una descrizione matematica della sua superficie è arduo, poiché non si hanno tutti i dati necessari. Solitamente si suppone che essa sia assimilabile a un ellissoide, in modo che esso approssimi bene la sua superficie (soprattutto per quanto riguarda le quote). L'estrema variabilità della superficie terrestre ha portato diversi studiosi a proporre varie forme di ellissoide, tra le quali quella più usata (perché si suppone approssimi meglio la superficie terrestre) è l'ellissoide di Hayford. Solitamente gli ellissoidi vengono orientati localmente per una superficie terrestre riconducibile a quella d'una regione, una nazione, un continente. Per questo quando si parla di coordinate geografiche si deve menzionare il *datum*, ovvero l'ellissoide di riferimento e l'orientamento. Oggi il più usato (anche da *Wikipedia*, nel campo coordinate delle località) è il sistema WGS84.

#### Geocoordinate nei files d'autorità ovvero nella Gemeinsame Normdatei - GND

Si rimane stupiti nel constatare come nei file d'autorità delle biblioteche le geocoordinate non siano state incluse, per molto tempo. lo stato attuale della qualità dei metadati è, pertanto, insufficiente per rendere possibile ricerche esplorative, come per uno Smart-City-Hub. Ne consegue che



gli authority record in cui non siano presenti le geo-coordinate hanno un'utilizzabilità circoscritta, come in presenza di nomi geografici semanticamente ambigui o con omonimi; per esempio, Granada in Nicaragua e Granada in Spagna. Proprio in questi casi, infatti, possono offrire una soluzione gli strumenti per l'indicizzazione automatica di opere, nonché di dati di ricerca correlati dalle informazioni geografiche, come le coordinate sulla base di file d'autorità. In questo contesto negli ultimi anni è stato condotto uno sforzo titanico per sviluppare authority file completi; nel mondo bibliotecario di lingua tedesca, per esempio, sono stati elaborati vari soggettari, tra cui Gemeinsame Normdatei (GND).<sup>5</sup> Le coordinate geografiche, invece, sono memorizzate in banche dati come GeoNames<sup>6</sup> o Wikidata.<sup>7</sup> Quest'ultima si è affermata negli ultimi anni come uno dei protagonisti principali nel collegare i vari pool di authority file esistenti. Per esempio, la Deutsche Nationalbibliothek (DNB) ha avviato una collaborazione con Wikimedia Deutschland per rafforzare la connessione tra gli authority file disponibili via Wikidata e i propri authority file. È evidente, infatti, che più i dati avranno una formattazione standard più sarà facile collegarli in rete e riutilizzarli (Mattamann und Regenass 2021, 313 n. 41).8 Di conseguenza, le informazioni geografiche registrate dovrebbero essere arricchite nei metadati il più possibile e segnatamente con le coordinate attraverso servizi di geocodificazione, per esempio GeoNames (Martins, Anastacio and Calado 2010). La fornitura di coordinate per le mappe e per le voci d'autorità, peraltro, consente opzioni di ricerca e selezione più efficienti e intuitive, grazie ad applicazioni cartografiche interattive nei cataloghi delle biblioteche, come si può vedere nel portale della Biblioteca nazionale scozzese.9 Un sistema di ricerca funzionale e utile valuta le voci d'autorità sulla base dei termini impiegati e tiene conto delle geo-coordinate registrate nelle voci d'autorità (Scheven 2016, 33). Last but not least, l'aggiunta delle coordinate fornisce, specificando il sistema di riferimento, un punto d'identificazione geografica unico (Müller und Schmunk 2019, 30).

## Ricerca geografica nei cataloghi delle biblioteche<sup>10</sup>

In quale situazione (usando un catalogo) si trova oggi un utente quando cerca una risorsa con un riferimento geografico, per esempio, una mappa o una carta? La ricerca inizia di solito con un nome, oppure con un termine geografico. Ogni OPAC prevede la possibilità di una ricerca libera benché la cosa migliore sia partire da un termine presente in un authority file. Un sistema di ricer-

<sup>&</sup>lt;sup>4</sup> Müller und Schmunk 2019, 30, nota: 94: Historisches OrtsNamensVerzeichnis (HONV) für MecklenburgVorpommern nell'ambito del progetto Virtuelles Kulturlandschaftslaboratorium (VKLandLab) vedi <u>www.vklandlab.uni-rostock.de</u>.

<sup>&</sup>lt;sup>5</sup> Sulla GND, la sua storia, sviluppo e contenuto vedi le informazioni sul sito web della DNB <a href="https://www.dnb.de/DE/Professionell/Standardisierung/GND/gnd\_node.html">https://www.dnb.de/DE/Professionell/Standardisierung/GND/gnd\_node.html</a>.

<sup>&</sup>lt;sup>6</sup> http://www.geonames.org.

<sup>&</sup>lt;sup>7</sup> https://www.wikidata.org/wiki/Wikidata:Main\_Page.

<sup>&</sup>lt;sup>8</sup> Da queste interconnessioni nasce una rete di informazioni che costituisce il punto di partenza ideale per l'arricchimento delle informazioni presenti nel proprio dataset. Il metacatalogo svizzero *Swissbib* lo ha implementato negli ultimi anni, da prototipo d'avanguardia. Le informazioni provenienti da vari insiemi di dati standard come il GND e *Wikidata* sono state collegate tra loro e il valore aggiunto su *Swissbib* è stato reso visibile e ricercabile <a href="https://swissbib.blogspot.com/2020/07/auf-wikipediawikidata-basierende-suchen.html">https://swissbib.blogspot.com/2020/07/auf-wikipediawikidata-basierende-suchen.html</a>.

<sup>9</sup> http://maps.nls.uk/geo/find/#.

<sup>&</sup>lt;sup>10</sup> Per approfondire i concetti dei paragrafi 4.1-4.3, vedi Scheven 2015, 40-42.



ca che funzioni solo con termini verbali, tuttavia, presenta svantaggi strutturali anche quando si utilizzano dati standard e precisamente:

- I. dipende dalla lingua che si usa: nel caso di voci d'autorità, infatti, potrebbero non essere elencate tutte le varianti linguistiche;
- II. qualora la ricerca non riguardasse un termine geografico, ma tutto ciò che fosse legato a un termine geografico specifico, le ricerche verbali mostrerebbero i loro limiti, nel senso che tutti i termini dell'area geografica desiderata dovrebbero essere conosciuti e inseriti come query. Se si volesse sapere tutto sull'area geografica "Baviera", per esempio, si dovrebbero inserire tutti i nomi di tutte le città e le entità amministrative bavaresi, nonché tutte le aree naturali;
- III. lo *zooming*, cioè la riduzione e l'allargamento dell'area di ricerca non è possibile con una ricerca solo verbale.

Se, però, oltre al termine verbale, nella ricerca si utilizzassero notazioni o codifiche, gli svantaggi di cui abbiamo appena parlato verrebbero attenuati in modo considerevole. Nella Classificazione decimale Dewey (DDC), per esempio, le notazioni relative all'area geografica sono in forma numerica e quindi indipendenti dalla lingua. Nella GND è possibile utilizzare il codice-Paese (Scheven 2013), formulato secondo lo standard DIN ISO 3166.<sup>11</sup> Le ricerche condotte mediante questo codice e selezionando l'opzione "Cerca tutto su (l'area geografica della) Baviera", restituiscono un numero di risultati maggiore rispetto a ricerche che utilizzino solo termini verbali (Baviera) (Wiesenmüller 2013). Il codice del Paese – convertito in una geofacetta – è abbastanza semplice da usare: i primi esempi l'hanno dimostrato già dieci anni fa in OPAC locali selezionati (già allora basati sulla tecnologia dei motori di ricerca), come quelli delle biblioteche universitarie di Heidelberg e Mannheim (Wiesenmüller, Maylein und Pfeffer 2011). Sappiamo, infatti, che un'accurata marcatura dei dati anche con i codici-Paese pertinenti e l'implementazione d'una ricerca per localizzazione geografica permettono d'ottenere un numero di risultati notevolmente maggiore (Müller und Schmunk 2019, 30).

## La ricerca geografica ideale

Una ricerca geografica ideale parte dall'utilizzo di un termine geografico, meglio se nella forma standardizzata in una voce d'autorità (Scheven 2015, 41-42). Le coordinate fungono da supporto arricchendo le voci d'autorità e svolgendo, in questo modo, la funzione di identificatori persistenti. Esse sono:

- a) indipendenti dalle denominazioni, che possono variare molto a seconda della lingua,
- b) possono cambiare nel corso del tempo,
- c) provengono da un sistema di riferimento internazionale e universalmente riconosciuto, nel senso che sono protette dal succedersi delle mode, nonché da cambiamenti arbitrari, sia della scienza pertinente (cioè la cartografia), sia degli utenti.

<sup>&</sup>lt;sup>11</sup> ISO – ISO 3166 — Country Codes.

<sup>&</sup>lt;sup>12</sup> Il codice-paese di Baviera è XA-GE-BY.



Per esempio: le coordinate del centro della città di Colonia, che si chiama *Cologne* in inglese e *Keulen* in olandese, sono sempre le stesse (E 6°57'00" N 50°56'00") indipendentemente dalla lingua e dagli eventuali cambiamenti politici che, nel corso dei secoli, possono modificare l'appartenenza della città a un determinato stato o nazione (anticamente l'Impero romano, oggi la Germania). Le coordinate sono importanti, perché consentono di compiere ricerche in spazi geografici, e possono essere utilizzate per ampliare o circoscrivere i risultati di una ricerca. A questo scopo, dunque, le carte geografiche debbono essere acquisite e georeferenziate singolarmente. Utilizzando le coordinate come punti di riferimento condivisi, inoltre, è possibile che esse vengano riutilizzate in altre applicazioni.

Nei cataloghi, una ricerca con termini verbali dovrebbe valutare automaticamente le voci d'autorità e tenere conto delle coordinate registrate in esse. Chi cerca una mappa può quindi inserire il termine di ricerca geografico a lui noto. Tramite le geo-coordinate del set di voci d'autorità, poi, vengono trovate tutte le carte geografiche relative a quella zona.

## Registrazione delle geo-coordinate da parte della DNB

Le coordinate geografiche vengono registrate tra i dati del catalogo della DNB e per le voci d'autorità nella GND (Scheven 2015, 42-44). Le nuove mappe da catalogare contengono i relativi dati di coordinate dal 2010, mentre dal gennaio 2014 la registrazione avviene tramite i formati MARC21 e RDF. L'inserimento avviene nel campo 034, che corrisponde al MARC21Authority<sup>13</sup> con alcune estensioni: le coordinate vengono registrate e visualizzate analogicamente con gradi, minuti e secondi. Al fine di consentire la condivisione dei dati da parte di altre applicazioni, nonché per il successivo utilizzo ed elaborazione da parte del computer, tuttavia, le informazioni devono essere disponibili anche in forma decimale; pertanto, le coordinate vengono automaticamente convertite in rappresentazione decimale à momento della loro registrazione nella GND. La rappresentazione analogica e decimale è stabilita da un valore in un sottocampo definito dal catalogatore. È, inoltre, possibile specificare se i valori siano stati solo determinati o stimati; ciò è utile quando si registrano le coordinate di mappe storiche: queste, infatti, adottano sistemi di registrazione delle coordinate che non sono più compatibili con i dati standard di oggi, perciò dovrebbero essere convertiti con costi elevati (Scheven 2015, 43).<sup>14</sup> Una soluzione a ciò, può essere individuata nel cosiddetto Bounding Box,<sup>15</sup> nel senso che lo spazio geografico rappresentato sulla mappa storica

<sup>13</sup> http://www.loc.gov/marc/authority/ad034.html.

<sup>&</sup>lt;sup>14</sup> Coordinate basate e definite su un modello ellissoide di riferimento geodetico differente da quello usato oggi (GPS o WGS84; mentre in passato si sono usati sistemi differenti, a partire dal sistema Maupertuis, del 1738, fino ai più recenti sistema ellissoide internazionale, dal 1924 fino al 1971; poi il GRS-67, dal 1971 al 1984, poi l'attuale WGS84), oppure prendendo come meridiano di riferimento uno diverso da quello di Greenwich (che è ufficialmente stato adottato come meridiano zero solo nel 1884, alla conferenza di Washington, mentre sino ad allora si usavano il meridiano che passava per l'Osservatorio di Parigi, quello che passava per l'isola di Ferro, nelle Canarie, quello che passava alle Azzorre, quello che passava per l'attuale linea del cambiamento di data, e quello di Greenwich: fino al 1884, dunque, quando si parla di meridiano zero' bisogna anche precisare quale scala di riferimento si usa): queste problematiche sono oggi superate grazie all'adozione di sistemi universali di riferimento per le misurazioni della superficie della Terra, ma vanno tenute presenti quando si lavora con documenti antichi.

<sup>15</sup> http://boundingbox.klokantech.com.



può essere disegnato in una *bounding box*, cioè una mappa interattiva con la georeferenziazione odierna aggiornata. Le coordinate dell'area raffigurata nella mappa sono determinate tramite il *bounding box*; così facendo si ottengono valori approssimativi molto buoni, ma occorre ricordarsi di contrassegnarli come tali. L'origine delle coordinate è obbligatoria; se possibile, si dovrebbe sempre designarle tramite un URI.<sup>16</sup>

Il formato della GND è, per quanto riguarda le coordinate, molto aperto e flessibile: esse possono essere registrate come *coordinate del centro* con due parametri, <sup>17</sup> oppure come *coordinate dei bordi* con 4 parametri<sup>18</sup> o, ancora, come poligoni con molti valori. <sup>19</sup> Le coordinate sono convertite in MARC 21 e in RDF per le applicazioni del web semantico. In quest'ultimo caso, viene utilizzato il formato standard GeoSPARQL<sup>20</sup> dell'Open Geospatial Consortium (OGC). <sup>21</sup> GeoSPARQL applica le regole del WKT (Well-Known Text) per la rappresentazione delle coordinate, esso fa parte dello standard Simple Features Access pubblicato da OGC e ISO. <sup>22</sup> Nel 2014 sono state importate nella GND circa 40.000 coordinate del centro dal database GeoNames, e in seguito altre 20.000 coordinate del medesimo tipo sono state acquisite tramite gli uffici statali di rilevamento, oppure sono state fornite dalla DNB stessa o da altre biblioteche o reti di biblioteche. Alla fine del 2021, infatti, la GND conteneva ben 62.000 voci d'autorità circa, registrate con coordinate del centro.

## Scenari e applicazioni

Al momento le geo-coordinate possono essere trovate solo nel sottocampo dell'indice di ricerca "koo", attraverso la cosiddetta 'ricerca avanzata' nel portale DNB (per esempio, koo=E0\*); la funzione 'ricerca avanzata', però, cerca solo nei record e nelle voci d'autorità, ma, purtroppo, non nelle voci d'autorità collegati ai record.<sup>23</sup> Quest'ultimo tipo di ricerca, invece, sarebbe indispensabile per un uso più ampio. Già la necessità di dover passare attraverso la 'ricerca avanzata' rappresenta un filtro, una barriera, perché impedisce agli utenti inesperti di utilizzare con successo questo strumento di ricerca. Mancando ancora, tuttavia, l'indicizzazione di questo campo, che pertanto non può venir usato dall'utente, la possibilità di fare ricerche sulla base dei dati geografici rimane aleatoria, anche quando essi sono presenti. Se le geo-coordinate sono incluse nei metadati, cioè nelle voci d'autorità, ed è possibile trovarle grazie a un'indicizzazione appropriata, ne risulterà un'ampia gamma di scenari di utilizzo. Ecco alcuni esempi:

1. per le richieste da parte di utenti con interessi turistici, in primis, l'integrazione e la valutazione delle geo-coordinate offrono un grande valore aggiunto: mediante una ricerca per area dei luoghi nelle vicinanze di un punto di riferimento, infatti, potrebbero venir trovate

<sup>&</sup>lt;sup>16</sup> Nelle istruzioni per l'inserimento della GND nel campo 034, ci sono i dettagli esatti per l'inserimento delle coordinate nella GND.

<sup>&</sup>lt;sup>17</sup> Le coordinate di un punto centrale indicano il punto centrale determinato nello spazio geografico.

<sup>18</sup> L'area geografica in questione viene completamente racchiusa in un quadrato che la individua con precisione.

<sup>&</sup>lt;sup>19</sup> Le aree geografiche hanno, di solito, contorni irregolari, poligonali, spigolosi, con numerosi punti in corrispondenza degli angoli che potrebbero venir registrati come valori identificativi.

<sup>20</sup> http://www.geosparql.org/.

<sup>21</sup> http://www.opengeospatial.org/.

<sup>&</sup>lt;sup>22</sup> http://www.opengeospatial.org/standards/sfa.

<sup>&</sup>lt;sup>23</sup> Comunicazione di Scheven/DNB in una mail del 14 giugno 2022.



- tutte le voci d'autorità geografiche della GND relativi a quest'area; inoltre, sarebbero collegate altre risorse indicizzate, che possano guidare l'utente-turista e accompagnarlo a letture di suo interesse, come mappe di città o guide turistiche (Scheven 2015, 45);
- 2. la creazione di geo-coordinate per le mappe e per le voci d'autorità consente opzioni di ricerca e selezione più efficienti e intuitive, mediante un'applicazione cartografica interattiva nei cataloghi delle biblioteche: così è stato fatto, per esempio, nel portale cartografico della Biblioteca nazionale scozzese.<sup>24</sup> Una ricerca funzionale e ben congegnata, infatti, non valuta solo le voci d'autorità sulla base dei termini presentati, ma tiene conto delle coordinate registrate nelle voci d'autorità (Scheven 2016, 33).
- 3. Le coordinate possono essere utilizzate per collegare diversi insiemi di dati relativi agli stessi elementi, di cui evidenziano angolature diverse: come frontend, cioè come strumento unificante, le mappe interattive sono spesso o prevalentemente utilizzate in questo modo. Un buon esempio è Fachinformationsdienst (FID), la biblioteca virtuale per l'archeologia, storia antica ed egittologia, o la piattaforma Propylaeum (Kempf 2021, 230). Nel quadro generale d'un progetto portato avanti da diversi anni, infatti, si sta cercando di collegare gli indici dei nomi dei luoghi di diversa provenienza attraverso le rispettive geo-coordinate, raggiungendo così il risultato di consolidare i relativi riferimenti bibliografici. La ricerca della posizione e la visualizzazione dei risultati vengono tramite una mappa interattiva. L'idea è frutto della cooperazione tra il FID Propylaeum e il Deutsches Archeologisches Institut (DAI) di Roma, con l'obiettivo di collegare i termini di localizzazione della GND utilizzati da Propylaeum con le voci di localizzazione del gazetteer del DAI, ossia un vocabolario standard di localizzazione archeologica completo che è stato costruito nel corso di decenni. Il gazetteer è stato sviluppato per le applicazioni bibliotecarie-bibliografiche del DAI ed è stato inizialmente utilizzato come soluzione locale. Ora, tuttavia, ci si è resi conto che il collegamento alla GND e la sincronizzazione dei dati aprono nuove possibilità di applicazione, di portata immensamente maggiore, come la GeoSuche in Propylaeum-SEARCH, con la quale è possibile cercare la letteratura scientifica relativa (traendo i dati dal DAI e dal Propylaeum)<sup>25</sup> ai singoli luoghi archeologicamente rilevanti – utilizzando una mappa interattiva che copre la regione mediterranea. Nel corso del progetto è previsto che altri "vocabolari d'autorità" del DAI verranno collegati alla GND.
- 4. Il notevole potenziale delle informazioni spazio-temporali nelle biblioteche e in altre istituzioni della memoria registrata, che finora è stato utilizzato in misura rudimentale, può essere adeguatamente sfruttato tramite maggiori e innovative possibilità d'esplorazione digitale, come i modelli di visualizzazione dei dati per le rappresentazioni e le valutazioni cartografiche. Il progetto di *Smart City "metro-basel digital*" (Mattmann und Regenass 2021, 310-311) ne è una dimostrazione: numerose istituzioni, tra cui la Biblioteca universitaria di Basilea, permettono di sperimentare lo sviluppo urbano e del traffico della città di ieri, oggi e domani su un hub interattivo con un modello 3D; la Biblioteca, inoltre, alimenta lo Smart City Hub con i propri dati bibliografici su vari argomenti, come le aree di sviluppo o le infrastrutture di trasporto, le quali a loro volta possono essere consultate in

<sup>24</sup> http://maps.nls.uk/geo/find/#.

<sup>&</sup>lt;sup>25</sup> https://www.propylaeum.de/en/searching/propylaeumsearch/.



modo georeferenziato in un luogo specifico. Così facendo, dunque, il portale di ricerca o il catalogo della biblioteca non diventeranno mai obsoleti grazie a questa offerta di servizi. Tutte le risorse ricercabili tramite lo Smart City Hub possono essere trovate anche al suo interno: con questa forma di visualizzazione dei dati, infatti, la ricerca non sarà più limitata a un solo campo del catalogo, ma i dati potranno essere cercati in parallelo attraverso una mappa interattiva di Basilea in 3D, venendo allo stesso tempo adeguatamente ancorati geograficamente. E non è finita: per il servizio Smart City, infatti, non sono interessanti solo i dati bibliografici della letteratura attuale, ma anche, e anzi forse soprattutto, le collezioni speciali, come manoscritti, ritratti, stampe antiche o fondi archivistici di biblioteche accademiche, oggi accessibili tramite portali online specifici per argomento o per tipologia di risorsa (si vedano di seguito gli esempi dalla Svizzera), come "e-rara," 26 "e-manuscripta,"27 "e-codices",28 o "Memobase"29 per cui le Smart Cities offrono un'ulteriore opzione di rappresentazione tramite mappe interattive. Pensiamo al sito che ha messo online in 3D la mappa di Basilea, di Matthäus Merian, che risale al 1617, e l'ha integrata nello Smart-City-Hub: ecco un esempio perfetto d'enorme potenziale.<sup>30</sup> Lo stesso vale per la cosiddetta "Haus zur Luft", dove si trovava la famosa stamperia di Hieronymus Froben (1501-1563).<sup>31</sup> le copie digitali delle sue stampe, i suoi ritratti e i dati bibliografici correlati possono essere richiamati in modo differenziato, una possibilità che, purtroppo, il catalogo della biblioteca non può (ancora) offrire. Uno Smart-City-Hub, dunque, offre la possibilità di condurre una ricerca più approfondita rispetto a quella consentita da un classico portale di ricerca o un catalogo, dando l'opportunità di mettere i dati (anche esterni alla biblioteca) in rete con una Regione e vari altri insiemi di dati (per esempio, dell'Istituto di statistica). In definitiva, i modelli di visualizzazione dei dati, come quelli esemplificati da uno Smart-City-Hub, possono non solo integrare un portale di ricerca, ma anche offrire agli utenti un orientamento completamente nuovo nella giungla dei dati (Mattmann und Regenass 2021, 311).

- 5. Nei portali dei giornali si potrebbero geovisualizzare i luoghi della loro distribuzione e le sedi delle istituzioni proprietarie di loro azioni. La base è l'inclusione preventiva delle geo-coordinate nei campi delle voci d'autorità.
- 6. I giochi per computer e online multigiocatore su internet sono molto popolari e grazie alla tecnologia Linked Open Data è possibile creare ambientazioni storicamente accurate: scenario e dettagli costruiti in modo interattivo, grazie al fatto che i dati delle geo-coordinate vengono presi in considerazione quando viene visualizzato uno spazio geografico (Scheven 2015, 45 n. 34).
- 7. Un altro campo è il *crowdsourcing*, ossia utenti o terze parti interessate si occupano dell'indicizzazione tramite il cosiddetto *geotagging*, per esempio, d'immagini aeree, cioè la loro localizzazione su una mappa interattiva, come quelle di Google Maps, o quelle degli uf-

<sup>&</sup>lt;sup>26</sup> https://www.e-rara.ch/.

<sup>&</sup>lt;sup>27</sup> https://www.e-manuscripta.ch/.

<sup>28</sup> https://www.ecodices.ch/.

<sup>&</sup>lt;sup>29</sup> http://memobase.ch/.

<sup>30</sup> https://ub.unibas.ch/de/ub-als-kantonsbibliothek/.

<sup>&</sup>lt;sup>31</sup> Fu lo stampatore delle opere di Erasmo da Rotterdam.



fici statistici statali, o altra tipologia di risorse cartografiche prodotte, cioè disegnata e programmata internamente con le istituzioni geografiche e geoscientifiche statali. Alcuni esempi sono il Politecnico di Zurigo (Müller und Schmunk 2019, 11) o lo Smapshot della Biblioteca regionale del Vorarlberg.<sup>32</sup>

### Conclusione

Gli esempi hanno illustrato le enormi potenzialità di impiego delle geo-coordinate; la possibilità di utilizzarle per supportare approcci esplorativi nei cataloghi delle biblioteche e per collegare insiemi di dati indipendenti attraverso la loro inclusione negli authority file (parola chiave: open linked data), in particolar modo, prospetta un uso più ampio. I loro vantaggi tradizionali – l'indipendenza linguistica e il fatto che non siano esposte al rischio di modifiche dei nomi geografici, solitamente motivate da ragioni politiche - sono sufficienti a renderle accessibili e ricercabili nei cataloghi delle biblioteche, ovunque sia possibile. E tuttavia, purtroppo, è proprio questo che ancora manca. Nei cataloghi delle grandi reti regionali e interregionali del mondo bibliotecario di lingua tedesca, come quello della rete di Baviera-Berlino-Brandenburgo, cioè il BVB3Kat, oppure negli OPAC di grandi biblioteche, come la Bayerische Staatsbibliothek, i dati della geolocalizzazione non sono né descritti sistematicamente, né resi disponibili per la ricerca degli utenti. La DNB, che – seppur con certe restrizioni – pratica già dal 2014 la metadatazione con le coordinate, ha trovato finora solo pochi imitatori.<sup>33</sup> Vi è comunque motivo di essere cautamente ottimisti. Le maggiori possibilità di presentazione e diffusione digitale, come l'utilizzo della visualizzazione dei dati per contestualizzare dati molto diversi tra loro, che nel caso dei dati geografici significa soprattutto lo sviluppo e l'utilizzo di visualizzazioni interattive delle mappe, giustificano la speranza che la vera svolta nell'utilizzo dei geodati debba ancora avvenire, e certamente avverrà in un futuro non troppo lontano.

<sup>&</sup>lt;sup>32</sup> Uno dei più grandi protagonisti e pionieri del geocrowdsourcing è la British Library. Il relativo progetto partiva nel 2011. La biblioteca mise in rete copie digitali del suo enorme fondo di mappe storiche (con circa 4,5 milioni pezzi) invitando gli interessati a localizzarle. Nel frattempo, circa 95 % delle mappe storiche digitalizzate (più di 72.000 pezzi) sono state arricchite con geodati attraverso il crowdsourcing. https://www.bl.uk/projects/georeferencer.

<sup>&</sup>lt;sup>33</sup> Per la tematica vedi Guerrini, Mauro. (2022). *Metadatazione: la catalogazione in era digitale*, prefazione di Diego Maltese; introduzione di Paola Castellucci e Gino Roncaglia. Milano: Editrice Bibliografica.



## Riferimenti bibliografici

Guerrini, Mauro. 2022. *Metadatazione: la catalogazione in era digitale*. Prefazione di Diego Maltese, introduzione di Paola Castellucci e Gino Roncaglia. Milano: Editrice Bibliografica.

Kempf, Klaus. 2021. "I nuovi sviluppi nel campo delle piattaforme tematiche". In *La Biblioteca* – *Piattaforma della Conoscenza: collaborativa, inclusiva, reticolare*, Convegno delle Stelline 2021, 223-231. Milano: Editrice Bibliografica.

Manzoni, Laura. 2022. Le risorse cartografiche: definizione, storia, catalogazione. Milano: Editrice Bibliografica.

Martin, Bruno, Ivo Anastacio, e Pavel Calado. 2010. "A machine learning approach for resolving place references in text". *Geospatial thinking* 221–236.

Mattmann, Beat, e Noah Regenass. 2021. "Eine neue Form der Recherche in Bibliotheken. Suchschlitz contra Exploration – Reduktion statt Orientierung?". Bibliothek – Forschung und Praxis 45(2): 304-316. https://doi.org/10.1515/bfp-2021-0010.

Müller, Friedrich, e Stefan Schmunk. 2019. "Bedeutung und Potenzial von Geoinformationen und deren Anwendungen im Kontext von Bibliotheken und digitalen Sammlungen". *Bibliothek – Forschung und Praxis* 43 (1): 21-34.

Scheven, Esther. 2013. Der Ländercode nach ISO 3166 und seine Nutzung. http://www.opus-bayern.de/bib-info/volltexte//2013/1513/.

Scheven, Esther. 2015. "Geokoordinaten in Bibliotheksdaten. Grundlage für innovative Nachnutzung". *O-bib. Das offene Bibliotheksjournal* 2(1): 35-46. <a href="https://www.o-bib.de/bib/article/view/2015H1S35-46/1881">https://www.o-bib.de/bib/article/view/2015H1S35-46/1881</a>.

Scheven, Esther. 2016. "Geografische Koordinaten und Bibliotheksdaten – Regelwerke und Anwendungen". *Dialog mit Bibliotheken* 28(1): 30-38.

Wiesenmüller, Heidrun, Leonhard Maylein, e Magnus Pfeffer. 2011. "Mehr aus der Schlagwortnormdatei herausholen – Implementierung einer geographischen Facette in den Katalogen der UB Heidelberg und der UB Mannheim". *B.I.T. online* 14(3): 245-252. <a href="http://www.ub.uni-heidelberg.de/archiv/12555">http://www.ub.uni-heidelberg.de/archiv/12555</a>.

Wiesenmüller, Heidrun. 2012. *Die geographische Facette. Eine neuartige Rechercheoption für Suchmaschinenkataloge*. <a href="https://tinyurl.com/4ckf5m5p">https://tinyurl.com/4ckf5m5p</a>.



# Organizational climate in the indonesian public library

## Dyah Srirahayu<sup>(a)</sup>

a) Universitas Airlangga

Contact: Dyah Srirahayu, <u>dyah.srirahayu@vokasi.unair.ac.id</u>
Received: 17 December 2022; Accepted: 15 March 2023; First Published: 15 May 2023

#### **ABSTRACT**

To maintain the existence and organization's innovation capability, it is necessary to pay attention to the development of the work environment. This work environment can be a benchmark in seeing the performance of employees or the organization itself. Therefore, the researchers conducted a study to see how the climate or work environment in public libraries in Indonesia. The method used is a qualitative approach with data collection through a Google forms questionnaire. The sample used is a public library librarian in Indonesia. The data analysis technique is through frequency distribution to see respondents' answers tend to explain what kind of work environment climate. The result of this study is that public libraries in Indonesia are categorized as having a high category and prioritize aspects of innovation and flexibility.

#### KEYWORDS

Public Library; Work Environment; Librarian; Library Innovation.

<sup>© 2023,</sup> The Author(s). This is an open access article, free of all copyright, that anyone can freely read, download, copy, distribute, print, search, or link to the full texts or use them for any other lawful purpose. This article is made available under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. JLIS.it is a journal of the SAGAS Department, University of Florence, Italy, published by EUM, Edizioni Università di Macerata, Italy, and FUP, Firenze University Press, Italy.



#### 1. Introduction

currently, the development of information technology has become a life for the community (Purnamayanti and Oktaria 2022). The reason is, information technology is developing to change civilization to be easier with the output of its products. Due to its rapid development, almost every aspect of social life strives to utilize information technology. This development is also followed by libraries where almost all libraries utilize information technology to optimize work, to use information sources, to disseminate information (Purnamayanti and Oktaria 2022). Research conducted by (Okunlaya, Syed Abdullah, and Alias 2022) illustrates research by developing (AI-LSICF) an innovative conceptual framework of artificial intelligence library services. The purpose of the study is to encourage library professionalism and improve service delivery. So that the current library has undergone many changes from the previous library. This provides the value and functionality of the library. If previously, the function of the library was only as a learning center, today the library shifts its function to a recreational function through its collections and services.

Development in the field of library, indirectly change the technical or procedural. For example, drive-thru services were recently developed by several libraries due to the impact of COVID-19. All staff and technical change work activities, in the past lending and return services, were carried out by physical contact, today changing to use different techniques or procedures. These changes result in a new organizational culture or work environment. The work environment is considered important because it affects the performance results of the organization (Kucharska 2021)transformational leadership and innovativeness are also included in the investigation as ascendants and consequences of the focal relation of intellectual capital and knowledge processes. Design/methodology/approach: Based on a sample of 1,418 Polish knowledge workers from the construction, healthcare, higher education and information technology (IT. The good and bad of the work environment will affect the final result of the performance. In addition to improving employee performance outcomes, the work environment can also create trust and innovative employee behavior. After that will be followed by an innovative work environment. Yu's research (2018) found that the existence of trust in organizations and innovative behaviors results in an innovative work environment.

The importance of the work environment is also stated by (Yu et al. 2018) that the work environment must have its characteristics. From that type, it will affect and become an important factor for the employees themselves because they will be considered capable of encouraging innovation or new ideas for the organization. Thus, today's organizations need to improve to create a work environment that is more distinctive, innovative, and has a climate of trust. It is proven, from research conducted (Ekvall and Ryhammar 1999) that if employees have a sense of trust in the organization and organizational climate, employees will involve themselves in making a more beneficial contribution to the organization.

Research related to this in various organizations is carried out to find out whether there is indeed a positive influence of the work environment on its employees. Research conducted by (Nurhasanah 2010) looked at whether there was an influence between the physical and non-physical work environment on employee performance. The results show that there is a real influence relationship between the physical and non-physical work environment on employee work performance partially or simultaneously. Another research conducted by (Prabowo 2018) wanted to find out whether the influence on employee performance. The results found that the existing work environment of



the Malang City Library and Archives Office did not have a significant effect on the performance of its employees. This is evident from the results of the regression analysis test which states that there is no influence of the work environment on the device because it is considered that the work environment is well-formed. Even when compared to other variables of motivation and stress, leaving results that influence the results of regression tests.

Further research on the work environment within the scope of the library. Research conducted by (Romadhoni, Asmony, and Survatni 2015) on the influence of the work environment, workload, and stress on librarian burnout in Mataram City. As a result, it has a positive influence on librarian burnout. This study used a linear multiple regression test where the test results stated F count > F table. Further research on the influence of the physical work environment on the work productivity of librarians at the Pacitan Regency Public Library. This research was conducted by (Farida 2012) who used physical work environment variables; lighting, sound, cleanliness, air, and workspace. The method in this study uses qualitative data collection with in-depth interviews. The results of this study found that four of the five factors influence work productivity. Because these four factors have not given comfort to librarians. Only the hygiene factor is the most superior among other factors. Research in 2022 conducted by Bahrudin is also related to building the resilience of the library environment during the COVID-19 pandemic. Due to the prolonged effects of the pandemic, this is also felt by libraries engaged in services. Bahrudin in this study interviewed 6 people, 4 of whom were librarians and 2 others were library users. Regarding the resilience of the library environment, libraries are expected to adopt the SNI ISO/45005 policy which is a practical solution to these conditions. Although the government has established occupational health and safety (K3) guidelines, it must be acknowledged that this has been a good step and agreed upon internationally. This means that the work environment of organizations including libraries is something that needs to be monitored, especially during a pandemic. Because it becomes an icon or image for the organization or company itself.

From several previous studies, a conclusion can be drawn that this work environment can directly or indirectly affect. If in other organizations, a corporate environment can affect its employees, then within the scope of the library can also affect librarians.

#### 2. Methods

This study uses descriptive quantitative using numbers and statistical tests. This study tested the theory using data that had been collected from 42 respondents. Data collection was carried out by distributing Google Forms to librarians in public libraries in Indonesia, namely the Malang City Public Library, the Yogyakarta Special Region Public Library, and the Gunungkidul Public Library. The questionnaire used to measure organizational culture in public libraries uses measurements from Patterson et al (2005). Organizational culture that is measured is formalization, tradition, innovation & flexibility.

The population of this study is all public library librarians in Indonesia who were selected as research locations. This criterion is determined to ensure that the respondent can provide information that is in line with the existing conditions. The distribution of questionnaires through Google Forms uses the Likert 5 scale method, namely Strongly Disagree, Disagree, Simply Agree, Agree, and Strongly Agree.



The Likert scale for respondents' answers is as follows:

Scale	Category				
1	Very Disagree				
2	Disagree				
3	Neutral				
4	Agree				
5	Very Agree				

The data from the respondent's answer results will be processed into a frequency distribution of each question item. Each item of the statement has a score from a predefined Likert scale. After that, it is grouped based on the number of respondents who answered each question item score. When each of them is grouped, it will be calculated on average for each aspect of the work environment by calculating the average of the known Likert scales.

How to calculate:

Likert scale 1 gained an average of 4.67

The likert scale 2 gained an average of <u>2.31</u> etc...

 $\bar{x}$  Formalization = 1 + 2 + 3 + 4 + 5/5 =  $\bar{x}$   $\bar{x}$   $\bar{x}$   $\bar{x}$ 

 $\bar{x}$  Tradition = 1 + 2 + 3 + 4 + 5/5 =  $\bar{x} \, \bar{x} \, \bar{x} \, \bar{x}$ 

 $\bar{x}$  Innovation & Flexibility = 1 + 2 + 3 + 4 +  $5/5 = \bar{x} \bar{x} \bar{x} \bar{x} \bar{x}$ 

After finding the average in each aspect of the work environment, it remains only to be included in the category that has been determined for each average. The average categories are as follows: Interval = (5-1)/3 = 1.33

Low: 1 – 2.33 Medium: 2.34 – 3.66 High: 3.67 – 5

## 3. Results and finding analysis

Table 1. Respondent Demographics

Gender	Sum	Percentage						
Man	13	31%						
Woman	29	69%						
Total	42	100%						
Origin of Respondents								
Hapless	14	33%						
Jogja	14	33%						
Gunung Kidul	14	33%						
Total	42	100%						
Length of Service								
<1 year	4	10%						
1-2 Years	8	19%						
3-4 Years	6	14%						
>5 years	24	57%						
Total	42	100%						



From the demographics of respondents listed in the table above, it can be seen that the number of respondents from Malang is 33% of the total number of 42 respondents, namely 14 people. The group of respondents from Jogja Kidul obtained 33% of the total number of respondents and Gunung Kidul obtained 33% of the total number of respondents. Each of the three locations obtained 14 people who were respondents to the study. It can also be taken from 69% of respondents, 29 are female respondents, while the other 31% are 13 male respondents. The quantity of females is greater than that of males. Meanwhile, the difference in the length of work from each respondent was obtained, namely, 57%, namely 24 respondents who had worked for >5 years, 19% that is 8 respondents worked for 1-2 years, 14% that is 6 respondents worked for 3-4 years, and 10% namely 4 respondents worked for <1 year in a public library work environment in Indonesia. This means that librarians who work >5 years have a greater quantity than librarians who are still working for <5 years and have the possibility that the answers given by respondents with longer experience can describe valid answers.

Table 2. Section/field of work

Field of Work	Sum	Percentage	
Collections Processing/Development	23	55%	
Service	12	29%	
Coaching	3	7%	
Public relations	1	2%	
IT	3	7%	
Total	42	100%	

The target respondents taken were librarians in three public libraries in Indonesia, this is the Malang City Library and Archives Office, the Gunungkidul Library and Archives Office, and the Yogyakarta Special Region Library and Archives Office. Several respondents are spread across several different fields of work, namely 55% of the total number of respondents as many as 23 respondents are in the field of processing and collection development work, 29% of which 12 respondents are in the field of service work, 7% are in the IT field as many as 3 respondents and the coaching field by 7% as many as 3 people, and the last 2% several 1 person is in the field of Public Relations. This diversity of work fields is certainly very natural to occur in the work environment of public libraries because there are indeed many fields of work and the division must be adjusted to the needs of the library. It should not be too little or too much in one area of work. For example, in the collection processing department where there are indeed many work activities that must be done, the officers who work must also be adjusted to their workload so that library work activities can be completed optimally.



Table 3. Organization climate

Organizational Climate	Frequency				Means	Information	
	1	2	3	4	5		
Formalization	14	18	33	49	54	3.66	Currently
Traditions	28	42	48	20	27	2.84	Currently
Innovation & Flexibility	3	7	43	81	76	4.48	Tall
Average						3.66	Tall

Table 3 above is a frequency distribution table of the Organizational Climate. Climate or organizational culture consists of several 3 dimensions, namely Formalization, Tradition, Innovation & Flexibility. This dimension of Formalization or formalization gets an average of the overall total of 3.66 and belongs to the medium category. The Tradition dimension gets an average of the overall total of 2.84 and belongs to the medium category. Finally, the Innovation & Flexibility dimension gets an average of the overall total of 3.38 and is included in the High category.

The innovation aspect of flexibility obtained the highest average among the other dimensions, at 4.48. This means that innovation and flexibility in organizational culture are approved by respondents with a High average score. Several respondents agreed on innovation and flexibility in organizations because the ideas of librarians are very welcome, library management also responds quickly when there are changes, and is responsive. When librarians are encouraged to propose several ideas or innovations then the organization must be able to accommodate or pay attention to the proposed idea. So that the librarian feels that the organization or management of the library is open to new ideas or developments.

In addition, several respondents agreed that flexibility in organization public library management is flexible in dealing with changes and quickly adapts to conditions last but not least, public library management assists in the development of ideas for librarians. The response from librarians shows the innovation and flexibility found in their work environment, namely the work environment of public libraries in Indonesia running properly. The journey of an organization is not always easy, there are even some challenges that demand a quick movement to change a situation. Only organizations that have good flexibility can read the situation to drive change. Innovation and flexibility will be one element or aspect that synergizes with each other if an organization or library can run both.

Meanwhile, the results of the formalization aspect of the climate or work environment obtained an average of 3.66 and were included in the medium category. This means that not a few and not many still agree with the formalization aspect of an organization. Indeed, formalization is still needed by organizations and even employees to regulate a procedure for work activities and their rules. For example, just as librarians are required to abide by rules, librarians can ignore procedures if they assist in the completion of public library work, all work must comply with the rules, and violations made will leave a response. The reality is that there are organizations that are formal and informal, all back to the needs of each individual, whether to enter and contribute to a formal or informal organization. However, in the end, everything that has clear procedures and rules will certainly be far from uncertainty. In a public library work environment, respondents do not turn a blind eye to formalization or rules that exist in an organization. The average from



the formalization aspect obtained an average of 3.66 and because it is included in the medium category, it means that formalization is still found in the work environment of public libraries in Indonesia.

Finally, the Tradition aspect of the organizational climate averaged 2.84 and was included in the Moderate category. It can be concluded that tradition in public libraries in Indonesia has begun to be abandoned. Because there are responses from respondents who disagree that tradition is still inherent in the work environment in Indonesian public libraries. For example, just as senior librarians like to maintain the traditional way of doing a job in a library, public library management is not interested in trying ideas from librarians, and public libraries are slow in doing a job. This does make the work environment boring with a traditional way of working as well as a senior librarian who is still attached to the culture of the past. Although the average tradition is still found in the work environment of public libraries in Indonesia, the average results show the lowest number compared to the average of the other two aspects. It can be concluded that the traditional aspects of the work environment of public libraries in Indonesia have begun to be abandoned because they are considered incompatible with the rapid development of information technology. Even some work systems can be easier to do because there is information technology.

#### 4. Discussion

judging from the results of research on the climate or work environment of public libraries in Indonesia using a qualitative approach method, a diverse picture of the work environment is obtained, such as a work environment that obeys the rules, a modern one that does not maintain traditional procedures, and a flexible, innovative work environment. From the results of observations, it can be seen that research that adopts 3 (three) aspects, namely Formalization, Tradition, and Innovation Flexibility, tends to describe more on the Innovation & Flexibility aspects. However, not a few also describe the formalization that is still well adhered to and obeyed by the people within the organization. In contrast to the Tradition aspect where each organization has developed and abandoned the old way of using the latest techniques.

#### 1) Innovation & Flexibility

The innovation and flexibility aspects get the highest value from other aspects where innovation and flexibility are more approved in the organizational environment in public libraries in Indonesia. Innovation and flexibility in organizations are preferred by librarians because a library not only innovates but also has flexibility in that regard. Looking at the response to the Innovation & Flexibility dimension which has obtained a high average, it can be seen that the work environment of public libraries has developed in such a way that it prioritizes innovation and flexibility in the organization. Research conducted by (Shahzad, Xiu, and Shahbaz 2017) Is it only one? Two different? It can't be verified because it is not in the references also shows positive results that the performance of this organizational innovation depends on how the culture of the organization. Does it have a positive influence such as being open to ideas/innovations from librarians, being flexible in changing procedures and adjusting to circumstances, and the speed of responding in work?



#### 2) Formalization

Formalization specs are still needed and used in work environments in public libraries as a matter of regulating work rules and procedures. The response to the formalization aspect ranks second only to the aspect of innovation and flexibility which acquires a high average value. We see that formalization is still happening in the working environment of public libraries in Indonesia. (García-Sánchez, García-Morales, and Martín-Rojas 2018)where other internal variables-such as technological assets, employee training, coordination of new management capabilities, and new flexible human resources and more adaptable organizational designs-must be encouraged to create value and competitive advantage. The purpose of our research is to analyse whether technological assets influence absorptive capacity (potential and realized absorptive capacity argues that formalization is considered a binding rule, policy, and procedure in an organization. When this formalization occurs conducively, it can produce important things for the company, but when it occurs rigidly, it will reduce flexibility so that it is considered an unnecessary obstacle (Gibson, Dunlop, and Cordery 2019). Organizational formalization in this case is related to well-defined work as well as rulemaking, decision-making, and policy implementation (Victor 2020). In addition, the focus of formalization is not on the relationship between superiors and subordinates but on workers and jobs. Formalization is likely to affect the ability of employees because it is considered limiting but management standards are to increase formalization in the sense of establishing and enforcing appropriate procedures and actions (Fischer et al. 2019). Because of this, formalization still occurs naturally in the work environment of public libraries in Indonesia.

#### 3) Tradition

The response from the aspect of tradition in the work environment in public libraries in Indonesia obtained the lowest average. Boedov argues that tradition is considered a conservative structure passed down from previous generations and understood as an inherent behavior in the community of an organization (2017). This tradition is usually characterized by the intensity of communication with which there are usually certain behaviors and types of communication. In a more modern context, Boedov also states that this tradition is a creative process that creates a worldview, and forms attitudes and programs of life. Unfortunately, research on the work environment in public libraries in Indonesia does not tend to describe aspects of tradition by obtaining low average figures.

#### 4. Conclusion

Result of the study it can be concluded that of the three organizational cultures the most highly challenged is a culture of innovation and flexibility in carrying out duties as a librarian. The results also reveal that in the library a culture of formality still exists, this is because in the library a lot of work is routine in nature and also because it is located in a public organization with a lot of bureaucracy to go through. But this is not a barrier for libraries to update because in completing their work they are flexible. Libraries need to maintain a culture of innovation and flexibility in work so that librarians can be improved to complete their work in routine jobs. Future research uses other variables to measure factors that influence cultural innovation in public libraries.



#### References

Bahrudin, Muhammad. 2022. "Membangun Ketahanan Lingkungan Perpustakaan Di Masa Pandemi COVID-19 Dengan Penerapan SNI ISO/PAS 45005:2020." *Pertemuan Dan Presentasi Ilmiah Standardisasi* 2021: 83–92. https://doi.org/10.31153/ppis.2021.117.

Boedov, Andrey A. 2017. "Tradition as a Factor of Self-Organization of Society." *Tradition as a Factor of Self-Organization of Society* 2: 1–7. https://doi.org/10.47451/phi2020-08-001.

Ekvall, Goran, and Lars Ryhammar. 1999. "The Creative Climate: Its Determinants and Effects at a Swedish University." *Creativity Research Journal* 12 (4): 303–10. <a href="https://doi.org/https://doi.org/10.1207/S15326934CRJ1204\_8">https://doi.org/https://doi.org/https://doi.org/https://doi.org/10.1207/S15326934CRJ1204\_8</a>.

Farida, Nur. 2012. "Pengaruh Lingkungan Kerja Fisik Perpustakaan Terhadap Produktivitas Kerja Petugas Perpustakaan Di Perpustakaan Umum Kabupaten Pacitan." *Journal of Library Science* 1 (1): 157–62.

Ronald Fischer, Maria Cristina Ferreira, Nathalie Van Meurs, Kubilay Gok, Ding-Yu Jiang, Johnny R J Fontaine, Charles Harb, Jan Cieciuch, Mustapha Achoui, Ma Socorro D Mendoza, Arif Hassan, Donna Achmadi, Andrew A Mogaji & Amina Abubakar. 2019. "Does Organizational Formalization Facilitate Voice and Helping Organizational Citizenship Behaviors? It Depends on (National) Uncertainty Norms." *Journal of International Business Studies* 50 (1): 125–34. <a href="https://doi.org/10.1057/s41267-017-0132-6">https://doi.org/10.1057/s41267-017-0132-6</a>.

García-Sánchez, Encarnación, Víctor J. García-Morales, and Rodrigo Martín-Rojas. 2018. "Influence of Technological Assets on Organizational Performance through Absorptive Capacity, Organizational Innovation and Internal Labour Flexibility." *Sustainability (Switzerland)* 10 (3). <a href="https://doi.org/10.3390/su10030770">https://doi.org/10.3390/su10030770</a>.

Gibson, Cristina B., Patrick D. Dunlop, and John L. Cordery. 2019. "Managing Formalization to Increase Global Team Effectiveness and Meaningfulness of Work in Multinational Organizations." *Journal of International Business Studies* 50 (6): 1021–52. <a href="https://doi.org/10.1057/s41267-019-00226-8">https://doi.org/10.1057/s41267-019-00226-8</a>.

Kucharska, Wioleta. 2021. "Leadership, Culture, Intellectual Capital and Knowledge Processes for Organizational Innovativeness across Industries: The Case of Poland." *Journal of Intellectual Capital* 22 (7): 121–41. https://doi.org/10.1108/JIC-02-2021-0047.

Nurhasanah, Andi. 2010. "Pengaruh Lingkungan Kerja Karyawan Pada Bank Indonesia Cabang Samarinda". *Jurnal Eksis* 6 (1): 1349–56.

Okunlaya, Rifqah Olufunmilayo, Norris Syed Abdullah, and Rose Alinda Alias. 2022. "Artificial Intelligence (AI) Library Services Innovative Conceptual Framework for the Digital Transformation of University Education." *Library Hi Tech* 40 (6): 1869–92. https://doi.org/10.1108/LHT-07-2021-0242.

Patterson, M.G. et al. 2005. "Validating the Organizational Culture Measure: Links to Managerial Practices, Productivity and Innovation." *Journal of Organizational Behavior*. 26: 379-408. https://doi.org/10.1002/job.312.



Prabowo, Lindu. 2018. "Pengaruh Lingkungan Kerja, Motivasi Kerja Dan Stres Kerja Terhadap Kinerja Pegawai." WIGA: Jurnal Penelitian Ilmu Ekonomi 8 (1): 24–31.

Purnamayanti, A, and N Oktaria. 2022. "Kesiapan Perpustakaan SMA Muhammdiyah 2 Bandar Lampung Dalam Pengembangan Website E-Perpustakaan." *Informatio: Journal of Library and Information Science* 2 (1): 19–28. <a href="http://jurnal.unpad.ac.id/informatio/article/view/36938">http://jurnal.unpad.ac.id/informatio/article/view/36938</a>.

Romadhoni, Lalu Ciptadi, Thatok Asmony, and Mukmin Suryatni. 2015. "Pengaruh Beban Kerja, Lingkungan Kerja, Dan Dukungan Sosial Terhadap Burnout Pustakawan Di Kota Mataram." *Khizanah Al-Hikmah : Jurnal Ilmu Perpustakaan, Informasi, Dan Kearsipan 3* (2): 124–45. <a href="https://doi.org/10.24252/kah.v3i2a3">https://doi.org/10.24252/kah.v3i2a3</a>.

Shahzad, Fakhar, Guoyi Xiu, and Muhammad Shahbaz. 2017. "Organizational Culture and Innovation Performance in Pakistan's Software Industry." *Technology in Society* 51: 66–73. <a href="https://doi.org/10.1016/J.TECHSOC.2017.08.002">https://doi.org/10.1016/J.TECHSOC.2017.08.002</a>.

Victer, Rogerio S. 2020. "Connectivity Knowledge and the Degree of Structural Formalization: A Contribution to a Contingency Theory of Organizational Capability." *Journal of Organization Design* 9 (1). <a href="https://doi.org/10.1186/s41469-020-0068-3">https://doi.org/10.1186/s41469-020-0068-3</a>.

Yu, Ming Chuan, Qiang Mai, Sang Bing Tsai, and Yi Dai. 2018. "An Empirical Study on the Organizational Trust, Employee-Organization Relationship and Innovative Behavior from the Integrated Perspective of Social Exchange and Organizational Sustainability." *Sustainability (Switzerland)* 10 (3). https://doi.org/10.3390/su10030864.



# Questionnaire

Gender:

Нр:

Originating Agency:

Field of Work:

Length of Work:

Innovations that exist in the workplace:

Innovations once proposed:

Is Making a Proposal every year? Yes/No

No	Statement	STS	TS	N	S	SS
1	Librarians are required to comply with established regulations					
2	Librarians can ignore formal procedures and rules if it helps get work done in public libraries					
3	All work performed in the library must comply with the established rules					
4	No one is too angry if there is a librarian who violates the rules in the public library					
5	Senior librarians like to maintain the traditional/conventional way of doing things in public libraries					
6	The way public libraries do their jobs hasn't changed much					
7	Library management is not interested in trying new ideas from librarians					
8	Changes in public libraries in the way they do work happen very slowly					
9	The ideas of librarians are very accepted by the public library					
10	Public library management is quick to respond when changes are needed to be made in the work					
11	Public library management quickly recognizes the need to do things differently					
12	Public library management is very flexible, it can quickly change procedures to adapt to new conditions and solve problems that arise					
13	Public library management provides assistance in developing new ideas from librarians					



# Books in a bubble: assessing the OAPEN Library collection

# Ronald Snijder(a)

a) OAPEN Foundation, Netherlands, https://orcid.org/0000-0001-9260-4941

Contact: Ronald Snijder, <u>R.snijder@oapen.org</u>
Received: 12 August 2022; Accepted: 21 December 2022; First Published: 15 May 2023

#### **ABSTRACT**

Open access infrastructure for books is becoming more mature, and it is being used by an increasing number of people. The growing importance of open access infrastructure leads to more interest in sustainability, governance and impact assessment. The assessment of the OAPEN Library – containing over 20,000 titles – fits within this trend. How well does the collections meets the needs of its stakeholders: readers, libraries, funders and publishers?

The composition of the collection is measured using subject and language. Both dissemination and the content-related aspects are paired to the number of publications. The average number of downloads per title is relatively similar for all subjects. However, the mean downloads of titles in English is roughly twice as much compared to German and the other languages. Combining subjects and languages shows that the dissemination of books in languages other than English is less predictable. This assessment has illustrated the composition of the collection and how its readers make use of it. The visualisation helped to tell a complicated story in a simple way; a powerful instrument to guide the further development of this open access infrastructure.

#### **KEYWORDS**

Open access; OAPEN Library; Library assessment; Books; Book impact.

<sup>© 2023,</sup> The Author(s). This is an open access article, free of all copyright, that anyone can freely read, download, copy, distribute, print, search, or link to the full texts or use them for any other lawful purpose. This article is made available under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. JLIS.it is a journal of the SAGAS Department, University of Florence, Italy, published by EUM, Edizioni Università di Macerata, Italy, and FUP, Firenze University Press, Italy.



## 1. Introduction

Open access infrastructure for books is becoming more mature, and it is being used by an increasing number of people. The growing importance of open access infrastructure leads to more interest in sustainability, governance and impact assessment. For this paper, we will assess the OAPEN Library. In the spring of 2022, it passed the milestone of 20,000 titles. This was a good moment to evaluate the core asset of the OAPEN Library: its collection.

The OAPEN Library contains freely accessible books and chapters, all of which have undergone external peer review. In other words, it functions as an academic library and in our assessment we should treat it as such. However, it is an online library and the limitations of physical books do not play a part. Shelf space is not to be considered.

More important is the question of how well the collection meets the needs of its users. The OAP-EN Library sees global usage; the collection reflects this by offering titles in over 50 languages. The collection is not focused on a specific subject area, but the choice of medium – books and chapters, not journals and articles – is more strongly associated with the humanities and social sciences. It does not track its users, but the supporters of the OAPEN Libraries are globally distributed academic institutions, scientific and scholarly funders and publishers. An assessment of the OAPEN Library should therefore take into account the diversity of languages, subjects and stakeholders.

The following section describes the history of the OAPEN Library and its closely related platform: the Directory of Open Access Books.

# 2. The OAPEN Library and DOAB

The OAPEN Library was officially launched in 2010 and has been set up to host and disseminate open access books and chapters (Ogg 2010). Together with publishers OAPEN is building a quality-controlled collection; OAPEN also provides services for publishers, libraries, and research funders in the areas of dissemination, quality assurance, reporting and digital preservation. Furthermore, OAPEN collaborates with research funders, allowing the deposition of books that have received financial support from several funders. These deposited titles also must have been subject to an external quality control. In 2013 OAPEN improved its reporting by providing COUNT-ER-conformant usage statistics in collaboration with IRUS-UK (IRUS-UK 2018).

The OAPEN Library is built using open-source software and all its metadata has been made available under a CC0 licence. It is not optimized to keep visitors as long as possible on the site to show them advertisements or sell them as many things as possible. Instead, it maximises the dissemination of books and chapters using as many venues as possible. Books might be discovered on the OAPEN Library, or through the Directory of Open Access Books, library catalogues, Google Scholar, OpenAIRE, Twitter, a blog or by any other means.

When OAPEN was in its project phase, discussions on setting up an additional services were taking place. Based on the Directory of Open Access Journals (DOAJ) – which was launched in 2003 – the idea of an directory listing open access books was developed. This lead to the development of the Directory of Open Access Books (DOAB). It would differ from the OAPEN Library in several aspects. Firstly, it would not host the full text of books but provide a searchable index



with links to the publications at the publisher's website or repository. Secondly, while OAPEN allows books that may be downloaded for private use only, all books listed IN DOAB should be published under an open licence, allowing re-use of the publications. And lastly, the service would be free for publishers that were approved to submit their publications.

DOAB was officially launched in 2013, as a joint service of OAPEN, OpenEdition, CNRS and Aix-Marseille Université, provided by the DOAB Foundation (Whitford 2014). Since then it has become a global focal point for open access books. The growth of the collection – in December 2022 it contained over 63,000 title descriptions – was also aided by other organisations that have become part of the "DOAB Trusted Platform Network": OpenEdition; Project MUSE; SciELO Books (DOAB 2021). The fourth member of the DOAB Trusted Platform Network is the OAPEN Library. In 2020 and 2021, both the OAPEN Library and the Directory of Open Access Books have been migrated to the DSpace 6 platform.

In the next section, we will discuss the literature on library and collection assessments.

## 3. Literature review

Assessments of libraries and their collections are taking place within a certain context. For instance, much of the research literature discusses the distinction between print holdings and electronic holdings. In libraries containing curated print collections and a more general digital collection it is important to measure the usage of both collections, which is called "book impact". In several cases, print holdings are more strongly correlated with book impact, while this correlation is less clear for electronic holdings. (Borin and Yi 2008) describe the changing environment in which libraries find themselves, where electronic titles play a more important role, leading to different demands on how the library must be operated. An assessment model must be flexible enough to function. In a subsequent article, (Borin and Yi 2011) describe the deployment of their model. Measuring the impact of a digital collection is challenging. (Perrin et al. 2017) describe the complexities of understanding the data generated by Google Analytics of a DSpace environment.

Other authors focus more on the stronger and weaker points of the library collection; an example is the work by (Hyödynmaa, Ahlholm-Kannisto, and Nurminen 2010) on collection mapping in Tampere University Library. At the George Mason University, (Kelly 2014) combines a multi-tool analysis with a systematic, subject-by-subject approach to the collection. And finally, (Lantzy, Matlin, and Opdahl 2020) describe a continuous library assessment with recurring cycles.

What is the result of these library assessments? According to (Folb, Wessel, and Czechowski 2011), the patrons of the University of Pittsburgh are accepting e-books as an alternative to print books. (Lamothe 2010) correlates usage of electronic books to the amount of curation applied to the collection and the collection size. The next step is to make the research results easy to understand and communicate. Visualisation can be a useful tool to convey complex ideas. (Chen 2017) describes several visualisation techniques, and (Finch and Flenner 2016) introduce the bubble charts that will be used in this paper.

We could argue that libraries are much more than their collections, and that the level of openness is an important precursor. If the library has a function beyond its parent organisation, it might



serve an even larger community. (Brookes-Kenworthy et al. 2019) looked into the relationship between academic library access policies and institutional positions on open access or open science publishing. Another aspect is policies towards unaffiliated user access, the so-called "walk-in users", as observed by (Stregger and McLean 2021). In a similar vein, infrastructures like the OAP-EN Library could also be assessed based on guidelines such as the Principles of Open Scholarly Infrastructure (Bilder, Lin, and Neylon 2020).

# 4. Data

In the previous section we have seen that library and collection assessments are complex and have to take into account the context of that specific library. Thus, what are the relevant factors for the OAPEN Library? Bibliodiversity¹ might be the most prominent aspect: the collection consists of publications from many academic publishers in multiple languages and is used by a global audience. Furthermore, the collection does not have a specific subject-oriented focus but the publication forms – books and chapters – are more used in the humanities and social sciences and less in the disciplines of science, technology and medicine.

Three types of stakeholders support the OAPEN Library: publishers, funders and libraries. Both publishers and funders contribute to the collection by making publications available. They will be interested in the dissemination of the books and chapters. For libraries, the composition of the collection will be paramount. How do the titles on offer fit within the information needs of their patrons?

The assessment of the OAPEN collection should consider these two aspects. The dissemination of books and chapters is measured through the number of downloads – based on COUNTER R5 conformant data. The composition of the collection is measured among two axes: subject and language. Both dissemination and the content-related aspects will be paired to the number of publications. The number of downloads are measured during a period of 12 months, ending March 2022.

The data set consists of 20,517 titles. For this assessment, the subject description and language of each title has been reduced to one major term. OAPEN uses the BIC Standard Subject Categories (Book Industry Communication 2010). The following subject listings have been used: "The arts"; "Economics, finance, business & management"; "Humanities"; "Language"; "Literature & literary studies"; "Society & social sciences" and all other topics are grouped under "Other subjects". Languages have been grouped into: English; German; Other. During the period April 2021 until March 2022, the total usage of the collection consisted of 11,876,706 downloads.

The data is freely accessible at <a href="https://doi.org/10.5281/zenodo.6985305">https://doi.org/10.5281/zenodo.6985305</a>.

The next section discusses the assessment of the OAPEN collection in more detail.

<sup>&</sup>lt;sup>1</sup> "[T]he diversity of academic content, both at the national and international level is essential for preserving research in a wide range of global and local topics, studied from different epistemic and methodological approaches, inspired by various schools of thought and expressed in a variety of languages."(Giménez Toledo et al. 2019)



# 5. Assessment of the OAPEN collection

In this paper the assessment of the OAPEN collection is based on several charts. All charts list three dimensions: the number of titles on the X-axis; the number of downloads on the Y-axis; and the size of the bubble's circumference displays the average number of downloads.

It is clear that the number of titles directly affects the number of downloads; larger groups of books and chapters will also accumulate more usage. Comparing the average number of downloads for the separate groups gives an indication of dissemination patterns: are certain sets of publications more 'popular' than others?

All charts are based on the distribution of titles within the complete collection. When we list the subjects and the languages in one table, several things are immediately visible. First of all, the dominance of English: 80% of the collection is written in this language. Secondly, the six subjects 'encompass' 76% of the collection, which means there is a considerable 'long tail' of titles discussing other topics, among them are science, technology and medicine.

Table 1 Percentages of subjects and languages

	English	German	Other	Total
Society & social sciences	23%	4%	2%	29%
Humanities	13%	2%	1%	16%
Economics, finance, business & management	7%	1%	0%	8%
Literature & literary studies	4%	1%	1%	7%
Language	3%	1%	1%	6%
The arts	5%	1%	0%	6%
Other subjects	24%	3%	1%	28%
Total	80%	12%	8%	100%

In the next subsections, the relations between these aspects and the dissemination are described in more detail. Dissemination is measured as the average number of downloads for a group of titles. Just as it is a simplification to reduce a book's subject to one main category, calculating the average number of downloads might hide large differences.



## **5.1** Collection by subject

Figure 1 displays the complete collection, sorted by subject. The large difference between "Social sciences" and "Humanities" compared to the other subjects is clearly visible, as is the large number of titles that are categorized under "Other subjects". What is interesting to note is that the average number of downloads per title is relatively similar for all subjects: between 496 and 638. From a dissemination perspective we can assume that subject is not a large factor. To put it differently, the users of the OAPEN Library seem to be equally interested in all the subjects the collection has to offer.

However, language is another important factor affecting usage, which is depicted in the next subsection.

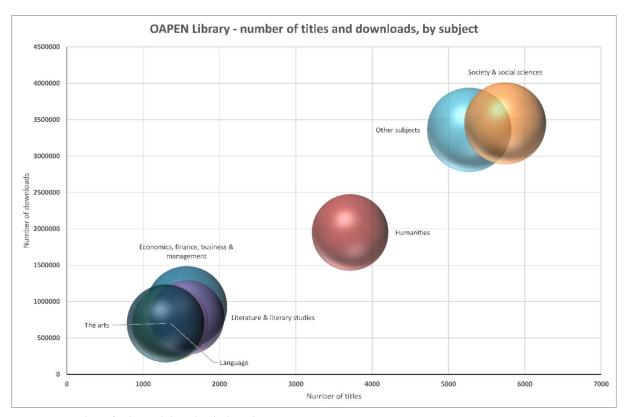


Figure 1. Number of titles and downloads, by subject.



## 5.2 Collection by language

The averages in Figure 2 differ quite dramatically. The mean downloads of titles in English is roughly twice as much compared to German and the other languages. And it is obvious when we look at the number of publications that English is the dominant language in the collection.

Here we also see that a larger number of titles does not always lead to higher average use. While the number of German-language books and chapters is twice as high as the number of titles in other languages, the average number of downloads of German titles is only 75% of the average for other languages.

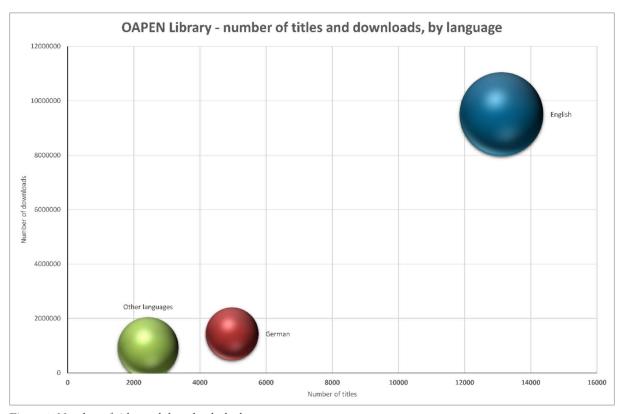


Figure 2. Number of titles and downloads, by language.



## 5.3 Subject and language combined

In this subsection, all combinations of subject and languages are illustrated. Within each of these sets, the number of books in English is always higher than the number of German-language titles or titles in other languages. Nevertheless, the dissemination of books in languages other than English tends to be less predictable. The average number of downloaded publications in other languages is much higher for "Humanities", "Literature" and "the Arts". In the other subjects, there is no such large difference.

### 5.3.1 Social sciences

The collection contains 5,743 publications on social sciences.

Language	Number of titles	Total downloads	Average downloads
English	3,500	2,763,290	790
German	1,448	468,247	323
Other	795	218,161	274

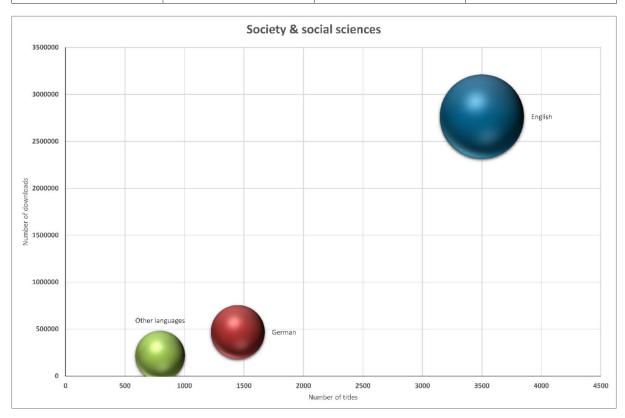


Figure 3. Social sciences.



# 5.3.2 Humanities

The collection contains 3,711 publications on humanities.

Language	Number of titles	Total downloads	Average downloads
English	2,443	1,583,925	648
German	787	199,119	253
Other	481	167,592	348

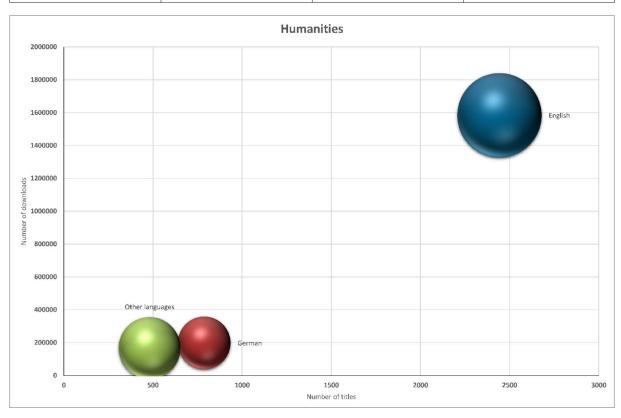


Figure 4. Humanities.



# 5.3.3 Economics

The collection contains 1,567 titles on economics.

Language	Number of titles	Total downloads	Average downloads
English	969	787,851	813
German	476	119,411	251
Other	122	21,437	176

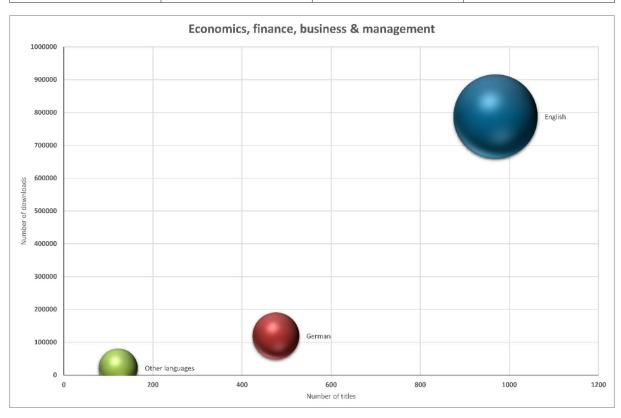


Figure 5. Economics, finance, business and management.



# 5.3.4 Literature

The collection contains 1,575 titles on literature.

Language	Number of titles	Total downloads	Average downloads
English	823	520,149	632
German	492	113,699	231
Other	260	147,371	567

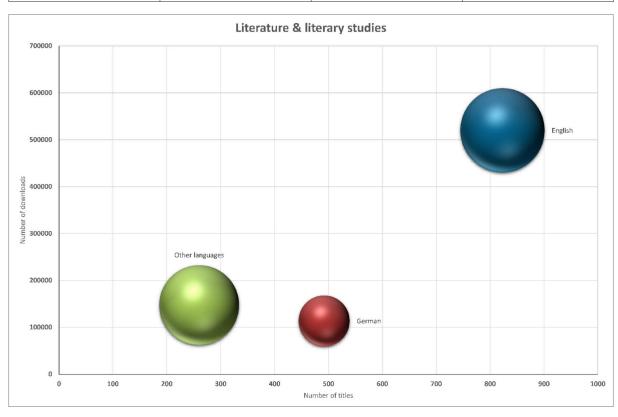


Figure 6. Literarture and literary studies.



# 5.3.5 Language

The collection contains 1,356 titles on language.

Language	Number of titles	Total downloads	Average downloads
English	656	392,272	598
German	488	144,530	296
Other	212	167,841	792

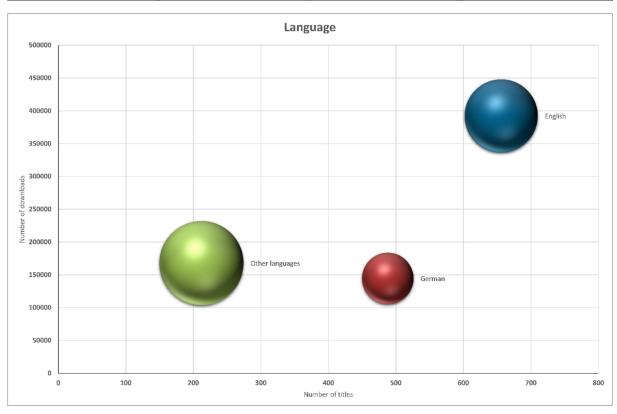


Figure 7. Language.



# **5.3.6** The Arts

The collection contains 1,294 titles on the arts.

Language	Number of titles	Total downloads	Average downloads
English	855	576,126	674
German	301	72,615	241
Other	138	51,814	375

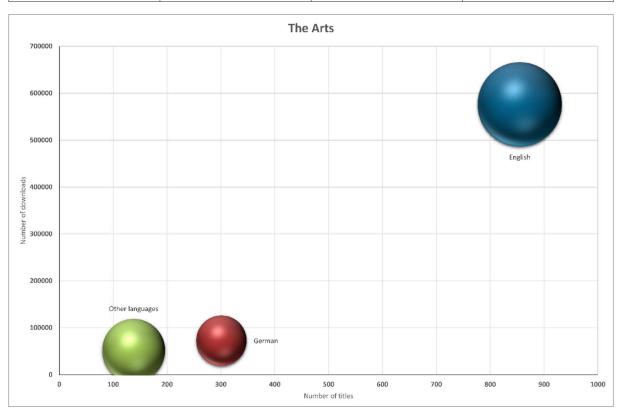


Figure 8. The arts.



# 5.3.7 Other subjects

The collection contains 5,271 titles on other subjects.

Language	Number of titles	Total downloads	Average downloads
English	3,868	2,887,631	747
German	979	317,828	325
Other	424	155,797	367

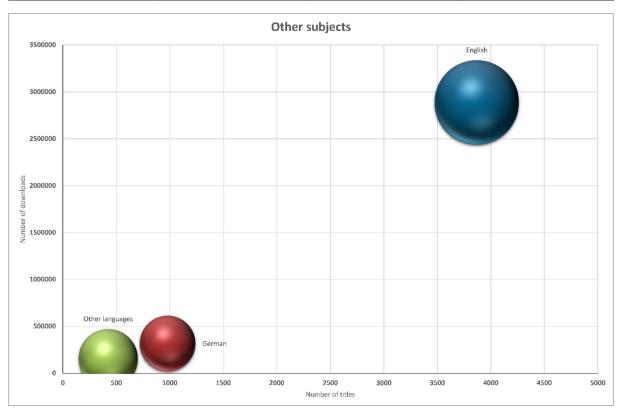


Figure 9. Other subjects.



## 6. Discussion

As we have seen in the literature review section, the assessment performed by a library should fit within its environment. Different circumstances lead to different questions. The OAPEN library operates in a different way from a more traditional library; it makes no distinction between members and visitors, as OAPEN Library content is freely available to everyone, but also functions as a supplier of an online collection for other libraries. This collection can be added completely or partially, leading to the question of the composition of the set of titles on offer. The assessment reflects this by focusing on subjects. However, within the humanities and social sciences, publication languages other than English are of more importance compared to the sciences. While most of the books and chapters are written in English, over 7,400 titles have been published in German or other languages.

Within the context of OAPEN, both publishers and funders are also important stakeholders. Publishers provide OAPEN with peer reviewed content. The collection is not subject based; any books and chapters can be added, provided that they have undergone an external quality assessment and are freely available to read and download. Both the publishers and the funders have an interest in how well the books and chapters are disseminated, and this is measures using COUNTER conformant download figures. Those numbers are affected by the language and the subject of the publications. That is why the combination of subject and language provides a better indication than the averages per subject alone.

The previous section displayed several patterns. Firstly, the dominance of English is visible in the number of titles available in the collection, and in the usage data. A larger number of titles will almost certainly lead to more downloads in total, but when we look at the average downloads per title, it is clear that English is more read by the users of the OAPEN Library. However, the impact of other languages varies strongly per subject. German is quite clearly the second largest language when we look at the number of publications. The total number of downloads for books and chapters in this language is consistently larger than the number of downloads for titles in other languages. Nevertheless, the average usage per title fluctuates. In the case of "Humanities", "the Arts" and especially "Language" and "Literature", the average number of downloads is higher for books in other languages. For these subjects, the collection of publications in other languages than English or German is smaller, but its book impact is larger.

This analysis did not take into account another important aspect of the usage: the regional diversity of the readers. We have seen that language has a large influence on the average number of downloads, especially when we compare English to other languages. Previous research on the international usage of the OAPEN Library has shown that readers from many countries have a strong preference for books written in their native or widely used languages (Snijder 2022). For instance, the ten most downloaded books in Chile were written in Spanish, while ten French language titles are most widely read in Cote d'Ivoire.

However, this also adds another layer of complexity to the assessment. Currently, the titles have been simplified by assigning them just one language and one subject. If we are to condense the geographical aspects of the readers, what choices need to be made? A possible option is to divide by the largest use per country. However, this would put much emphasis on large countries with a well-established internet infrastructure. Another possibility could be to create groups based on



continents. Of course, this will also hide away much of the diversity. Further discussion would be required to determine whether regional diversity of the readers should be added to these types of assessments and if so, how this should be implemented.

## 7. Conclusion

As can be seen from the results, even when the collection has been simplified to titles with one subject and one language, the usage patterns are not altogether straightforward. Comparing absolute download numbers are useless when the context – the subject and language – are not taken into account. Creating benchmarks to compare the book impact of a specific title is certainly a possibility, but only when the context has been taken into account.

Furthermore – and this is an important reason behind the continuous evaluation described by (Lantzy, Matlin, and Opdahl 2020) – this assessment is a "snapshot" depicting the situation at a certain time. Since the spring of 2022, the collection of the OAPEN Library has grown, not in the least with the addition of over 1,400 Italian-language titles. Obviously, a larger and more diverse collection will lead to different assessment results.

In conclusion, this assessment has vividly illustrated the composition of the collection and how its readers make use of it. The visualisation helped to tell a complicated story in a simple way; a powerful instrument to guide the further development of this open access infrastructure.

# 8. Acknowledgements

The author would like to thank the OAPEN colleagues for commenting on a previous version of this paper.



## References

Bilder, Geoffrey, Jennifer Lin, and Cameron Neylon. 2020. 'The Principles of Open Scholarly Infrastructure'. <a href="https://doi.org/10.24343/C34W2H">https://doi.org/10.24343/C34W2H</a>.

Book Industry Communication. 2010. 'BIC Standard Subject Categories – an Overview'. 2010. <a href="http://www.bic.org.uk/7/BIC-Standard-Subject-Categories/">http://www.bic.org.uk/7/BIC-Standard-Subject-Categories/</a>.

Borin, Jacqueline, and Hua Yi. 2008. 'Indicators for Collection Evaluation: A New Dimensional Framework'. *Collection Building* 27 (4): 136–43. https://doi.org/10.1108/01604950810913698.

——. 2011. 'Assessing an Academic Library Collection through Capacity and Usage Indicators: Testing a Multi-dimensional Model'. *Collection Building* 30 (3): 120–25. <a href="https://doi.org/10.1108/01604951111146956">https://doi.org/10.1108/01604951111146956</a>.

Brookes-Kenworthy, Chloe, Richard Hosking, Chun-Kai (Karl) Huang, Lucy Montgomery, Cameron Neylon, Alkim Ozaygen, and Katie Wilson. 2019. "Is the Library Open?": Correlating Unaffiliated Access to Academic Libraries with Open Access Support', September. <a href="https://hcommons.org/deposits/item/hc:26879/">https://hcommons.org/deposits/item/hc:26879/</a>.

Chen, Hsuanwei Michelle. 2017. 'Chapter 4. Real-World Uses for Information Visualization in Libraries'. *Library Technology Reports* 53 (3): 21–25. <a href="https://www.journals.ala.org/index.php/ltr/article/view/6291">https://www.journals.ala.org/index.php/ltr/article/view/6291</a>.

DOAB. 2021. 'DOAB Trusted Platform Network'. 2021. <a href="https://www.doabooks.org/en/publish-ers/doab-trusted-platform-network">https://www.doabooks.org/en/publish-ers/doab-trusted-platform-network</a>.

Finch, Jannette L., and Angela R. Flenner. 2016. 'Using Data Visualization to Examine an Academic Library Collection'. *College & Research Libraries* 77 (6): 765–78. https://doi.org/10.5860/crl.77.6.765.

Folb, Barbara L., Charles B. Wessel, and Leslie J. Czechowski. 2011. 'Clinical and Academic Use of Electronic and Print Books: The Health Sciences Library System e-Book Study at the University of Pittsburgh.' *Journal of the Medical Library Association : JMLA* 99 (3): 218–28. <a href="https://doi.org/10.3163/1536-5050.99.3.009">https://doi.org/10.3163/1536-5050.99.3.009</a>.

Giménez Toledo, Elea, Emanuel Kulczycki, Janne Pölönen, and Gunnar Sivertsen. 2019. 'Bibliodiversity-What It Is and Why It Is Essential to Creating Situated Knowledge'. *Impact of Social Sciences Blog.* http://eprints.lse.ac.uk/103440/1/impactofsocialsciences 2019 12 05 bibliodiversity what it is and why.pdf.

Hyödynmaa, Merja, Aniita Ahlholm-Kannisto, and Hannele Nurminen. 2010. 'How to Evaluate Library Collections: A Case Study of Collection Mapping'. *Collection Building* 29 (2): 43–49. https://doi.org/10.1108/01604951011040125.

IRUS-UK. 2018. 'International Collaboration and Value: Working with OAPEN'. IRUS-UK Case Studies. <a href="https://irus.jisc.ac.uk/documents/IRUS-UK">https://irus.jisc.ac.uk/documents/IRUS-UK</a> working with OAPEN.pdf.

Kelly, Madeline. 2014. 'Applying the Tiers of Assessment: A Holistic and Systematic Approach to Assessing Library Collections'. *The Journal of Academic Librarianship* 40 (6): 585–91. <a href="https://doi.org/10.1016/j.acalib.2014.10.002">https://doi.org/10.1016/j.acalib.2014.10.002</a>.



Lamothe, Alain. 2010. 'Electronic Book Usage Patterns as Observed at an Academic Library: Searches and Viewings'. *Partnership: The Canadian Journal of Library and Information Practice and Research* 5 (1): 1–16. http://journal.lib.uoguelph.ca/index.php/perj/article/viewArticle/1071.

Lantzy, Tricia, Talitha Matlin, and Judy Opdahl. 2020. 'Creating a Library-Wide Collection Management Cycle: One Academic Library's Approach to Continuous Collection Assessment'. *Journal of Library Administration* 60 (2): 155–66. https://doi.org/10.1080/01930826.2019.1677092.

Ogg, Lesley. 2010. 'Launch of the OAPEN Library at the Frankfurt Book Fair'. 20 September 2010. https://blog.alpsp.org/2010/09/launch-of-oapen-library-at-frankfurt.html.

Perrin, Joy M., Le Yang, Shelley Barba, and Heidi Winkler. 2017. 'All That Glitters Isn't Gold: The Complexities of Use Statistics as an Assessment Tool for Digital Libraries'. *The Electronic Library* 35 (1): 185–97. https://doi.org/10.1108/EL-09-2015-0179.

Snijder, Ronald. 2022. 'Big in Japan, Zimbabwe or Brazil – Global Reach and National Preferences for Open Access Books'. *Insights* 35 (June): 11. <a href="https://doi.org/10.1629/uksg.580">https://doi.org/10.1629/uksg.580</a>.

Stregger, Elizabeth, and Jaclyn McLean. 2021. 'Location-Based Trickery: Librarian Efforts to Provide Access to Unaffiliated Users'. *Weave: Journal of Library User Experience* 4 (2). <a href="https://doi.org/10.3998/weaveux.222">https://doi.org/10.3998/weaveux.222</a>.

Whitford, Leslie. 2014. 'DOAB: Directory of Open Access Books'. *Reference Reviews* 28 (3): 12–13. https://doi.org/10.1108/RR-10-2013-0269.



# 'Definite cataloguing rules set down in writing': Antonio Panizzi's *Rules* and the catalogue's manifestations

# Mauro Guerrini<sup>(a)</sup>, Stefano Gambari<sup>(b)</sup>

a) Università degli Studi di Firenze, Dipartimento di Storia, Archeologia, Geografia, Arte e Spettacolo (SAGAS), https://orcid.org/0000-0002-1941-4575 b) https://orcid.org/0000-0002-2910-9654

Contact: Mauro Guerrini, <u>mauro.guerrini@unifi.it</u>; Stefano Gambari, <u>s.gambari@gmail.com</u> Received: 21 December 2022; Accepted: 30 January 2023; First Published: 15 May 2023

#### **ABSTRACT**

The authors evaluate the symbolic value of the 91 rules published in 1841, as incipit of the first volume of the *Catalogue of printed books of the British Museum*, declaring its founding basis, and legitimizing the growing reputation of Antonio Panizzi at the same time. Monument of the élites and Trustees of the British Museum Library, the 91 were preceded by 73 rules, printed and distributed only among the cataloguers of the library in March 1839. The essay explains the existence of a normative stratification and evolution of the rules (oral tradition, limited circulation, revision and wider diffusion) and therefore the need for an archeology of knowledge's approach. The novelties of the first cataloging code reside in the passage from an empirical to a conscious approach to the functions of the catalogue, while the printing of the catalogue is considered by Panizzi an antieconomical manifestation. The catalogue had not yet found its expressive paradigm, consolidated only today in the format of an open digital work, constantly updated.

## **KEYWORDS**

Cataloguing; Antonio Panizzi; 91 Cataloguing Rules; British Museum.

# 'Definite cataloguing rules set down in writing': le *Rules* di Antonio Panizzi e le manifestazioni del catalogo

#### ABSTRACT

Il saggio analizza il valore simbolico delle 91 regole edite nel 1841 quale incipit del primo volume del *Catalogue of printed books of British Museum*, a dichiararne le basi fondative, e insieme legittimando la crescente reputazione di Antonio Panizzi. Monumento delle élites e dei Trustees della British Museum Library, le 91 furono precedute da 73 regole, stampate e diffuse solo tra i catalogatori della biblioteca nel marzo 1839. Il saggio esplicita l'esistenza di una stratificazione normativa e di una evoluzione delle norme (tradizione orale, stampa in circolazione limitata, revisione e diffusione più ampia) e dunque la necessità di un approccio di archeologia del sapere. Le novità del primo codice di catalogazione risiedono nel passaggio da un criterio empirico a un approccio consapevole delle funzioni del catalogo, mentre la stampa del catalogo è considerata da Panizzi una sua manifestazione antieconomica. Il catalogo non aveva ancora trovato il suo paradigma espressivo, consolidatosi solo oggi nel formato di un'opera digitale aperta, in continuo aggiornamento.

#### PAROLE CHIAVE

Catalogazione; Antonio Panizzi; 91 Regole di catalogazione; British Museum.

<sup>© 2023,</sup> The Author(s). This is an open access article, free of all copyright, that anyone can freely read, download, copy, distribute, print, search, or link to the full texts or use them for any other lawful purpose. This article is made available under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. JLIS.it is a journal of the SAGAS Department, University of Florence, Italy, published by EUM, Edizioni Università di Macerata, Italy, and FUP, Firenze University Press, Italy.



Le 91 regole di Antonio Panizzi, che furono edite nel 1841 nell'*incipit* del primo volume del *Catalogue of printed books of British Museum*, a dichiararne le basi fondative, erano state precedute da 73 regole, stampate e diffuse solo internamente tra i catalogatori della biblioteca del British Museum nel marzo 1839. Esse furono presto dimenticate sia perché superate dalle 91 regole, che riscontrarono immediatamente un grande interesse nella comunità bibliotecaria europea e statunitense, sia per la loro natura di strumento a uso interno che le rendeva "limitatamente" pubbliche o, *de facto*, "non pubbliche". Le 73 regole uscirono dall'oblio nel 2002 quando vennero descritte sommariamente in un articolo su "Cataloging & Classification Quarterly" da Michael Carpenter,¹ che tuttavia non ne riportava il testo. Oggi sono rese pubbliche "universalmente" grazie alla liberalità della British Library, che ringraziamo infinitamente, la quale ha risposto positivamente alla nostra richiesta di ottenerne la riproduzione digitale e all'autorizzazione della loro pubblicazione in questo fascicolo di "JLIS.it".

Le 73 regole costituiscono un passaggio necessario nel processo di verifica e riflessione che porterà alle 91 regole. Il 13 luglio 1839 i Trustees, approvate le 91 Rules, disposero che Panizzi le desse alle stampe nel numero di 50 copie «for the use of the establishment, and should send a copy to every officer and assistant, whether supernumerary or otherwise in his department for their guidance».<sup>2</sup> La stampa delle Regole, quale *ouverture* del primo volume del Catalogue of printed books of British Museum (1841)<sup>3</sup> testimonia, infatti, l'evoluzione di un percorso redazionale che inizia almeno due anni prima e che segna il risultato di un'elaborazione complessa avvenuta nel tempo; le Rules sono

frutto di una stratificazione normativa che ha visto quale base le sedici regole di Henry Hervey Baber, un ampliamento determinato dal lavoro di Panizzi, che ne ha portato il numero a settantatré, infine due revisioni a opera dei Trustees, la prima in seguito alla quale sono state introdotte altre sei norme, per un totale di settantanove, la seconda che ha fissato il numero alle novantuno regole oggi conosciute.<sup>4</sup>

Le *Regole* costituiscono una pietra miliare, un 'monumento' sia per le élites intellettuali e politiche inglesi sia per i Trustees della British Museum Library, che si avvia in tal modo a essere riconosciuta come la più importante biblioteca al mondo e una delle biblioteche con i migliori servizi al pubblico. Il valore simbolico è rilevante anche per Panizzi. Giunto a Londra nel maggio 1823, si afferma nella società inglese grazie alla sua rete di relazioni e alla risonanza della sua fama nei circoli letterari e politici dell'aristocrazia cosmopolita *whig* e di altre matrici politiche. Nel 1831 egli viene proposto alla carica di *extra assistant librarian* del Dipartimento dei libri a stampa del British Museum; la sua nomina è approvata all'unanimità dai Trustees, per le sue doti, ampiamente riconosciute, di studioso e per l'ottima presentazione di Henry Brougham<sup>5</sup> e di Thomas Grenville.<sup>6</sup> Nel 1832, a

<sup>2</sup> Papers 1847, xvii-xviii.

<sup>&</sup>lt;sup>1</sup> Carpenter 2002.

<sup>&</sup>lt;sup>3</sup> Le 91 Rules furono tradotte in francese da Olivier Barbier nel 1845; cfr. per una traduzione italiana Tavoni 1982.

<sup>&</sup>lt;sup>4</sup> Guerrini e Neri 2020, 153.

<sup>&</sup>lt;sup>5</sup> Henry Peter Brougham (1778-1868), barone di Brougham e Vaux, avvocato di origini scozzesi, fondò con Francis Jeffrey, Sydney Smith, e Francis Horner l'Edinburgh Review nel 1802 pubblicandovi vari articoli. Deputato whig, sostenitore delle cause liberali, membro del Parlamento per alcuni collegi elettorali, contribuì all'approvazione di numerose riforme civili, quali il Reform Act (1832) e il Slavery Abolition Act (1833). Fu tra i fondatori della Society for the Diffusion of Useful Knowledge (1825) e della London University (1828), poi University College of London (UCL).

<sup>6</sup> Thomas Grenville (1755–1846), membro del Parlamento dal 1780 al 1810, con un intervallo; colto bibliofilo, collezionò



meno di dieci anni dall'abbandono dell'Italia, Panizzi acquisisce la cittadinanza inglese e in dieci anni di lavoro, integrandosi perfettamente nella sua seconda Patria, riversa le sue energie nel lavoro quotidiano di biblioteca, continuando a insegnare italiano alla London University e coltivando gli amati studi letterari.<sup>7</sup>

In un periodo brevissimo Panizzi analizza i punti di forza e di debolezza dei servizi dell'istituzione di cui è un funzionario, pronto a pianificare gli interventi necessari a rendere i servizi più adeguati alle necessità contemporanee del pubblico della borghesia e degli studiosi londinesi. L'altro obiettivo che persegue il 'Gran Pan', come verrà chiamato, è riedificare nel suo complesso la British Museum Library, l'istituzione a cui sta dedicando, con modestia e determinazione, la sua attività professionale. Egli immagina le funzioni e i servizi di una grande Biblioteca nazionale, autonoma rispetto al Museo, e prefigura nel dettaglio le finalità che vanno definendosi tramite il continuo confronto con le analoghe realtà bibliotecarie presenti a livello internazionale; è una 'ricostruzione' alla quale egli sa che potrà mettervi mano quando gli accadrà di divenire quasi 'naturalmente' *Keeper* nel 1837 e *Principal Librarian* nel 1856.

Nel contesto della sua rapida ascesa, le Regole consolidano l'autorevolezza di Panizzi, ne legittimano la crescente posizione, la competenza riconosciuta da tutti e forse temuta, accreditandone l'elevata professionalità e il rigore nella gestione dell'intero processo di elaborazione, condivisione, verifica e comunicazione delle norme. Panizzi introduce l'uso del metodo nella pratica catalografica, tramite un codice di norme puntuali e innovative, basate sui principi di uniformità e consistenza, completezza e accuratezza che costituiranno le basi della prassi catalografica inglese sino al 1880. Il codice del British Museum è il punto di partenza e il modello di tutte le successive norme di catalogazione, seppure il suo ruolo storico vada giudicato senza alcuna enfatizzazione. Panizzi si occupa dell'organizzazione del catalogo per autore: stabilisce l'opportunità di riunire le opere di uno stesso autore sotto un'unica forma del suo nome, prevede la possibilità di accessi plurimi alle registrazioni catalografiche in presenza di possibili alternative di ricerca, propone il concetto di unità letteraria, precisa l'ordine di presentazione degli elementi descrittivi, fornisce solide motivazioni in favore della fedeltà alle informazioni riportate sul frontespizio. La novità delle sue norme consiste nel passaggio da un criterio empirico a un approccio consapevole, critico, teorico degli scopi e delle funzioni del catalogo e della sua moderna organizzazione strutturale, portando così a maturazione una tradizione sperimentata nel corso dei secoli. Le norme del British Museum nascono, infatti, dalla necessità di codificare la pratica dell'istituto e sono il risultato dell'esperienza di molti anni di lavoro, del riferimento a codici redatti da numerosi bibliotecari, a cominciare dal domenicano Giovanni Battista Audiffredi. Panizzi riconosce la funzione della stampa del codice:

libri sin da ventenne. La sua biblioteca, di oltre ventimila volumi, preziosa per i manoscritti e i numerosi rari incunaboli, venne utilizzata da Panizzi che vi consultò le edizioni cinquecentesche di Boiardo e Ariosto, per la sua edizione critica dell'*Orlando innamorato* e dell'*Orlando furioso* (London, William Pickering, 1830-1834, in 9 volumi). Panizzi si prodigò affinché la British Museum Library potesse acquisire in donazione la collezione, e Grenville – per l'amicizia e la fiducia che riponeva nel keeper – la donò con il vincolo che rimanesse unita in un fondo di cui si sarebbe dovuto pubblicare il catalogo. Panizzi curò in ogni dettaglio il trasferimento e la disposizione dei libri nella nuova sede: era il più prestigioso fondo di edizioni antiche donato al Museum dopo la biblioteca di Giorgio III.

<sup>&</sup>lt;sup>7</sup> Cfr. Gambari e Guerrini 2020a, 2020b.

<sup>&</sup>lt;sup>8</sup> Giovanni Battista Audiffredi (1714–1794) fu incaricato nel 1765 della Biblioteca Casanatense, pubblicò *Bibliotheca Casanatensis librorum typis impressionorum*, 1761–1790 e altri repertori bibliografici; studioso di matematica e astronomia,



produrre quella uniformità nelle pratiche catalografiche che, con le parole di Albert Predeek<sup>9</sup> «was only to be reached by definite cataloguing rules set down in writing, and not by verbal and variable agreements».<sup>10</sup>

Cosa significava stampare il catalogo di un istituto che voleva assurgere a modello di Biblioteca nazionale? Nell'Inghilterra dell'età vittoriana la stampa del catalogo alfabetico della British Museum Library costituì un obiettivo a cui contribuirono dal 1830 al Novecento generazioni di bibliotecari, di funzionari e rappresentanti delle istituzioni culturali e governative, con l'intenzione di curare una pubblicazione che non fosse solo, come scrive Arthur Hugh Chaplin,<sup>11</sup> «a piece of internal library apparatus», quanto «the product of an organised collective activity».

Nell'epoca attuale è evidente che nessuno proporrebbe mai di stampare il catalogo, tipica risorsa in continuazione che documenta il patrimonio descritto di una biblioteca, continuamente 'in movimento' per le nuove acquisizioni e revisioni delle collezioni. La stampa è solo una forma di pubblicazione, di 'render pubbliche' informazioni di un insieme di risorse; gli stessi schedari di una biblioteca costituiscono una 'pubblicazione' in quanto sono 'rese pubbliche' agli utenti in quello spazio magico e cruciale che era la 'Sala cataloghi'.

Il catalogo si è manifestato nel corso del tempo in diverse forme: volumi manoscritti, volumi a stampa, schedari a schede mobili, elettronico, forme-catalogo che non sono neutrali rispetto all'utente, alle funzioni che svolgono, alla diversa utilità, 'reticolarità' e 'profondità' della struttura sindetica, alla conservazione e al controllo delle informazioni che possono assicurare al pubblico dei lettori.

Panizzi e i suoi collaboratori nutrivano numerose perplessità sulla proposta di stampare il catalogo, sostenendo che le attività del suo aggiornamento in forma manoscritta non potevano essere sospese o ridotte, e che la disponibilità di un solo catalogo a stampa per il pubblico sarebbe stata inutile e svantaggiosa. La complessità delle attività di catalogazione volte a produrre un catalogo a stampa richiedeva un approccio di gestione del progetto, diretto al controllo dei tempi e della qualità delle registrazioni e al coordinamento del personale, delle singole attività e dei prodotti. Nell'audizione del 23 febbraio 1836, Panizzi proponeva di creare una copia del catalogo

transcribing titles on half-sheets of an uniform size, kept togheter by *reliures mobiles* (which I have seen advantageously used in great foreign libraries, and specimens of which have been ordered by Mr.

-

osservò le posizioni della cometa di Halley pubblicando i risultati nella "Dimostrazione della stazione della cometa, 1769" (1770).

<sup>&</sup>lt;sup>9</sup> Albert Predeek (1883-1956), bibliotecario tedesco, rappresentante della disciplina della documentazione (*Dokumentationswissenschaft* o *Dokumentationsbewegung*) derivata da Paul Otlet, direttore dal 1929 della Bibliothek der Technischen Hochschule di Berlino, vi istituì dal 1933 un centro d'informazione per la letteratura tecnica e scientifica. Dal 1947 al 1951 fu direttore della Biblioteca dell'Università di Jena.

<sup>&</sup>lt;sup>10</sup> Predeek 1937, 516-517.

<sup>&</sup>lt;sup>11</sup> Arthur Hugh Chaplin (1905–1996), bibliotecario inglese. *Principal keeper of printed books* della British Library dal 1966 al 1970, autore di numerosi saggi di biblioteconomia, partecipò alle più importanti conferenze internazionali sugli standard catalografici. Nel 1963 ricevette il riconoscimento 'Margaret Mann Citation in Cataloging and Classification' per 'his masterly preparation of the *Draft Statement of Principles* which showed creative insight into universal conditions of bibliographical entry', che costituì un riferimento rilevante per i Principi di Parigi dell'IFLA (1961).

<sup>12</sup> Chaplin 1987, vii.



Baber) so that copies of the Catalogue, both for the reading-room and for the use of the house, would be completed simultaneously with the original Catalogue on slips.<sup>13</sup>

Nella lettera che scrisse all'Earl of Ellesmere<sup>14</sup> il 29 gennaio 1848 Panizzi approfondirà il tema delle schede, con una comparazione costi/benefici delle strisce di catalogo (*slips*) rispetto alle schede (*sheets*).<sup>15</sup> Le *slips* erano un semplice espediente dai notevoli inconvenienti, istituito per rendere ancora attuale il catalogo a stampa di Henry Ellis e Henry Hervey Baber pubblicato in 7 volumi dal 1813 al 1819. Il loro contenuto poteva essere trasposto nelle schede, e nuove schede potevano essere prodotte per le recenti accessioni, permettendo così di dar vita a una nuova forma o manifestazione del catalogo che avrebbe consentito diversi ordinamenti funzionali e un'indubbia utilità per il pubblico degli studiosi.

In questo periodo storico, in cui il catalogo a schede mobili non è ancora divenuto un consolidato 'paradigma biblioteconomico', Panizzi ipotizza delle modifiche alla forma fisica del catalogo (*schede* che affiancano e sostituiscono le *strisce*) che avranno ripercussioni sulla stessa forma concettuale espressa dalle future norme di catalogazione.

La transizione al catalogo a schede sarebbe stata

'epocale' per la storia delle biblioteche e per l'evoluzione stessa della teoria catalografica [...]. Le conseguenze di questa innovazione sono di grande rilievo per le procedure di catalogazione, in quanto le schede mobili consentono uno sviluppo ottimale della struttura sindetica del catalogo e ne rendono più agevole l'aggiornamento in 'tempo reale', rispettando il principio del rigoroso ordine alfabetico degli autori e dei titoli. <sup>16</sup>

Le 91 *Rules* segnano un punto saldo di partenza: come ogni codice, sarà applicato e 'testato', ma le regole saranno anche 'interpretate'. Il processo applicativo è insieme adattivo, poiché il codice subisce un'evoluzione continua che conduce a periodiche revisioni rispondenti alle esigenze reali di una rappresentazione sempre più funzionale delle risorse bibliografiche. Le mutazioni del codice, le aggiunte o le modifiche, un certo insieme di pratiche accettate saranno, nel periodo successivo a Panizzi, a volte registrate e approvate formalmente, ma più spesso trasmesse solo oralmente; come scrive Arthur H. Chaplin: «A body of accepted practice [...] transmitted orally by revisers to new cataloguers».<sup>17</sup>

Alla codifica delle nuove regole che si aggiungevano formalmente o informalmente, come sapere implicito, lavorò Russell Martineau, <sup>18</sup> per darle a stampa nel 1887 in una pubblicazione di sole 20 copie! Di esse sopravvisse solo una copia, trovata intorno al 1950, in uno scaffale del suo ufficio, dal Principal Keeper C.B. Oldman.<sup>19</sup>

<sup>&</sup>lt;sup>13</sup> Papers 1847, 54-55; la riproduzione digitale è del progetto Google libri su una copia della Bodleian Library proveniente dal British Museum.

<sup>&</sup>lt;sup>14</sup> Francis Leveson-Gower (1800 – 1857), politico, scrittore, viaggiatore e mecenate delle arti, primo Earl of Ellesmere, presidente della Royal Commission Appointed to Inquire into the Constitution and Government of the British Museum.

<sup>15</sup> Panizzi1848, trad. it., 35.

Rizzo 2001, 30. In Italia i nuovi schedari sono impiegati dal 1882, quando la Biblioteca Vittorio Emanuele II di Roma affiderà le forniture ad Aristide Staderini, ideatore dei due modelli, a cassetta e a volume.

<sup>17</sup> Chaplin 1987, 77.

<sup>&</sup>lt;sup>18</sup> Russell Martineau (1831–1898), insegnante, esperto orientalista, operò presso la British Library dal 1857 al 1896, per gli ultimi dodici anni come *assistant Keeper*.

<sup>19</sup> Rules 1887.



Siamo dinanzi quasi a un 'ricorso storico' che accosta il caso di questo testo a uso interno a quello delle 73 *Regole*, pur nella differenza di genesi e di contenuto: la letteratura a diffusione interna delle regole di catalogazione è quasi sempre di difficile conservazione, conoscenza e accesso, ma vitale per comprendere la nascita, l'evoluzione e la trasformazione del codice.

Il confronto analitico tra le 73 e le 91 regole compiuto da Franco Neri è insieme un saggio storico e metodologico, a volte quasi 'indiziario', sempre accorto all'ingresso di nuovi concetti e nuovi termini e agli slittamenti semantici che subiscono quelli già in uso nei lessici catalografici e bibliografici, mutamenti che a volte occorrono durante l'arco di un decennio. La ricerca storica volta a spiegare il processo di formazione delle *Regole* basandosi su modelli, dati e presupposti è condizionata da fatti ed espressioni di opinioni, che sono documentate 'performativamente' ossia espresse nel modo più funzionale agli obiettivi che i diversi attori perseguono.

Panizzi intendeva il nuovo catalogo quale "great national undertaking", impresa "worthy of the Nation" che doveva essere all'altezza di una grande e moderna biblioteca nazionale: nasce un nuovo genere in ambito biblioteconomico, il codice di catalogazione, composto da un insieme definito di 'Rules to be observed in preparing and entering titles' (le 73), e in forma più compiuta 'Rules for the compilation of the catalogue' (le 91), o ancora più tardi 'Rules to be observed in the compilation of the Catalogue' (1887).



# Riferimenti bibliografici

Carpenter, Michael. 2002. "The original 73 Rules of the British Museum: a preliminary analysis". *Cataloging & Classification Quarterly* 35 (1-2): 23-36.

Chaplin, Arthur Hugh. 1987. GK: 150 years of the General Catalogue of printed books in the British Museum. Aldershot: Scolar Press.

Gambari, Stefano, e Mauro Guerrini. 2020a. "Voi sarete più ricco, ma dubito moltissimo se sarete più felice". A proposito del periodo d'insegnamento di Antonio Panizzi a Londra: 1828-1831". In La Rilliana e il Casentino. Percorsi di impegno civile e culturale. Studi in ricordo di Alessandro Brezzi, a cura di A. Busi, L. Conigliello, P. Scapecchi, 151-69. Firenze: Consiglio regionale della Toscana.

Gambari, Stefano, e Mauro Guerrini. 2020b. "Antonio Panizzi e le sue due antologie di letteratura italiana: Extracts from the Italian Prose Writers e Stories from Italian Writers". In *Libri, biblioteche e società. Studi per Rosa Marisa Borraccini,* a cura di Alberto Petrucciani, Valentina Sestini e Federico Valacchi, 245-59. Macerata: Editrice Eum.

Guerrini, Mauro, e Franco Neri. 2020. "La tormentata formulazione delle Regole del British Museum del 1839". In *Scaffali come segmenti di storia: studi in onore di Vincenzo Trombetta*, a cura di Rosa Parlavecchia e Paola Zito, 153-65. Roma: Edizioni Quasar.

Panizzi, Antonio. 1848. "Mr. Panizzi to the Right Hon. The Earl of Ellesmere". In Carpenter and Svenonius, eds. 1985: 18-47. Traduzione italiana: "Mr. Panizzi al Right Hon. Earl of Ellesmere". Traduzione di Luigi Crocetti. *Biblioteche oggi*, 18, n. 6 (lug.-ago. 2000): 33-48.

Papers 1847 = Papers relating to the Alphabetical Catalogue of printed books. Private and confidential. 1847. London: George Woodfall and Son (<a href="http://dbooks.bodleian.ox.ac.uk/books/PDFs/590119436.pdf">http://dbooks.bodleian.ox.ac.uk/books/PDFs/590119436.pdf</a>).

Predeek, Albert. 1937. "Panizzi and the British Museum Catalogue". *The Library Association Record*", 1937: 515-20; 579-82; 622-26. Traduzione inglese del saggio, edito nel medesimo anno: "Antonio Panizzi und der alphabetischen Katalog des Britischen Museums.", in *Festschrift Georg Leyh*, 257-82. Leipzig: Otto Harrassowitz.

Rizzo, Alberto. 2001, "Aristide Staderini e il catalogo a schede mobile. Profilo di un pioniere". *Biblioteche oggi* 19, n. 3 (aprile): 30-32. <a href="http://www.bibliotecheoggi.it/2001/20010303001.pdf">http://www.bibliotecheoggi.it/2001/20010303001.pdf</a>.

Rules 1887 = Rules to be observed in the compilation of the Catalogue of Printed Books in the British Museum. Embracing the original ninety-one rules, with additions, corrections and explanations. 1887. HMSO: London.

Tavoni, Maria Gioia. 1982. "Regole per la compilazione del catalogo: Appendice". In *Atti del Convegno di studi su Antonio Panizzi, Roma, 21-22 aprile 1980*, a cura di Enzo Esposito, 225-241. Galatina (Lecce): Editrice Salentina.



# 73 vs 91 Rules: drafting a code for new alphabetical catalog of the British Museum

# Franco Neri

Contact: Franco Neri, <u>franconeri50@gmail.com</u>
Received: 18 December 2022; Accepted: 30 January 2023; First Published: 15 May 2023

#### ABSTRACT

The paper, starting from the draft of 73 Rules Museum proposed by Panizzi to the Trustees in March 1839, aims at bringing out Panizzi's underlying design for the new alphabetical catalog of the British Museum. Two temporal perspectives are crossed in contextualizing the process of drafting the rules: one of shorter duration, with a focus on items and actors of a confrontation that became intense between 1834 and 1839; and another, of a wider time span, aimed at reconstructing the conflicts of bibliographic cultures, the lines of tradition and the originality of Panizzi's approach.

At the bottom of the draft, there is an intellectual wager: although in the context constraints, to go beyond the boundaries of the English bibliographic traditions of the time by opening up to other bibliographic cultures. The most stimulating perspective comes from Audiffredi's lesson mediated by a principle of *exactitude* which, if in the most immediate references has roots in the highest peaks of the contemporary French bibliographic culture (Barbier and Querard), draws its sap from the erudite tradition of the Italian eighteenth century (Tiraboschi). The paper discusses whether and what changes were made to this design in the passage from 73 to 91.

#### **KEYWORDS**

Antonio Panizzi; British Museum; Descriptive Cataloguing; 91 Rules; 73 Rules (Draft).

# 73 vs. 91: alle origini delle Regole per il nuovo catalogo alfabetico del British Museum

#### **ABSTRACT**

Il saggio, a partire dal draft di 73 regole proposto da Panizzi ai Trustees nel marzo 1839 per il nuovo catalogo alfabetico del British Museum, intende ricostruire il modello di catalogo che sottostà al testo delle 73. Nel contestualizzare il processo di elaborazione sono incrociate due linee temporali: una di più breve durata, ricostruendo il momento e gli attori di un confronto che si fa intenso fra il 1834 e il 1839; e un'altra, di più ampio arco temporale, tesa a far emergere il conflitto di culture bibliografiche, le linee di tradizione e l'originalità dell'approccio di Panizzi. Al fondo delle 73 *Regole* c'è una scommessa intellettuale: nei vincoli dati, andare oltre i confini della tradizione inglese dell'epoca con l'apertura ad altre culture bibliografiche. Il corto circuito più stimolante proviene dalla lezione di Audiffredi mediata da un principio di *exactitude* che, se nei suoi riferimenti più immediati ha radici nelle punte più alte della coeva cultura bibliografica francese (Barbier e Querard), trae la sua linfa dalla tradizione erudita del Settecento italiano (Tiraboschi). Il saggio discute se e quali modifiche vi siano state a tale disegno nel passaggio da 73 a 91.

#### PAROLE CHIAVE

Antonio Panizzi; British Museum; Catalogazione descrittiva; 91 regole; 73 regole.

<sup>© 2023,</sup> The Author(s). This is an open access article, free of all copyright, that anyone can freely read, download, copy, distribute, print, search, or link to the full texts or use them for any other lawful purpose. This article is made available under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. JLIS.it is a journal of the SAGAS Department, University of Florence, Italy, published by EUM, Edizioni Università di Macerata, Italy, and FUP, Firenze University Press, Italy.



«We do not make catalogues for Mr. Bolton Corney: this is a great national undertaking». Antonio Panizzi, Audizione del 14 maggio 1849 dinanzi alla Royal Commission, Q. 9751

# A) 73 vs. 91 Regole: problematiche di una transizione

È il 13 luglio 1848. La Royal Commission, incaricata di verificare le modalità di gestione delle attività del British Museum e in particolare della sua biblioteca, si avvia a completare la seconda sessione di audizioni.¹ È la sesta audizione di Antonio Panizzi. All'esame dei Commissari sono le 91 Regole elaborate per il nuovo catalogo alfabetico del British Museum. Panizzi sta illustrando il processo di revisione attraverso il quale dal 20 marzo al 13 luglio 1839, a seguito delle riunioni con i Trustees, si è passati da un draft originario di 73 Regole da lui proposto a 91, con un testo intermedio di 79 regole. Il resoconto è puntuale, con tanto di esibizione dei testi e delle modifiche apportate.

The rules now before the Commissioners are the result of the meetings of the Trustees. I will show the Commissioners the rules as they were submitted to the Trustees. will show how they altered them. Each Trustee, the Secretary, and I, had, every one of us, a copy of these rules. I was not able to prepare the whole at first. They were not 91 originally; they were only 73 [...]. This is my own copy, and the Commissioners will see here that Mr. Baber's rules are printed in italics, and the proposed additions in Roman. They were read one by one; every Trustee made his observations [...]. I took, on the spot, memoranda, in pencil, of those alterations (and here they are in my copy), and then I embodied and drew them up in the shape of rules. In order that the Trustees should well understand the effect of their rules, my original rules were printed as they stand here, and opposite to each rule, in a column originally left blank on purpose, were printed the alterations in Sub-Committee, and this was done all through [...]. Here is a copy of the rules I proposed, and here are the alterations in Sub-Committee [...]. The rules then were 79, and they came to be 91, by all the additions and alterations made in Sub-Committee [...]. The present rules are the rules of the Trustees, such as they have been altered in Sub-Committee and in Committee (*Report* 1850, Q. 4117, 240-241).

È il primo riferimento pubblico esplicito all'esistenza di un draft di 73 regole. Michael Carpenter (2002, 24), che per primo ha studiato il tema in un importante saggio, così descrisse la fonte:

These 73 rules, printed without date, and titled *Alphabetical Catalogue of Printed Books. Rules to be observed in preparing and entering titles*, were mislaid through having been incorrectly marked in the archives of the British Library's Department of Printed Books, until Mr. P. R. Harris, a then retired Deputy Keeper of the Department of Printed Books, found them for the present writer on July 4, 1988 [...]. An intermediate printing, containing 79 rules and some of the revisions of the Trustees, is yet to be found.

-

<sup>&</sup>lt;sup>1</sup> La Royal Commission venne nominata il 17 giugno 1847. La prima sessione si svolse dal 10 al 13 luglio 1847; la seconda dal 1 febbraio 1848 al 21 luglio 1848; la terza ed ultima dal 6 febbraio al 21 giugno 1849.

<sup>&</sup>lt;sup>2</sup> Report of the Commissioners appointed to inquire into the constitution and government of the British Museum: with minutes of evidence. London: William Clowes, 1850. D'ora in poi Report 1850. La citazione degli interventi è nella forma: numero della "Question" (Q.) di riferimento o di un gruppo consecutivo di "Questions", di norma senza indicazione di pagina. La pagina è aggiunta qualora l'argomentazione in una medesima Q. si sviluppi su più pagine.



Il testo è costituito dalle 73 regole (1-8), dagli esempi illustrativi (9-17, es. n. 1-284) e dalla prova della voce complessa Horace (n. 285-604).

Da due report di Panizzi del 18 marzo 1839 e del 7 gennaio 1847 ricaviamo ulteriori informazioni. Sono documenti interni al dibattito nel British Museum.<sup>3</sup> L'8 marzo 1839 Panizzi presenta ai Trustees una proposta che si può considerare una svolta nella sua prospettiva di costruzione di un insieme di norme adeguato alla sfida del nuovo catalogo delle collezioni a stampa. Sin dal *Rapporto Baber* (1834) la sua posizione era rimasta coerente. Egli l'aveva espressa con compiutezza per la prima volta nel report ai Trustees del 23 febbraio 1836 (*Papers* 1847, 53-57):

- condividere un set di principi sufficientemente coerente da configurare un catalogo "full and accurate";
- valutare i casi bibliografici nuovi per verificare se la loro soluzione fosse possibile all'interno dei principi e dei criteri individuati, o elaborarne altri più comprensivi;
- affidare il coordinamento ad un funzionario del British Museum in grado di governare con una presenza competente, costante ed esclusiva l'intero processo. I Trustees avrebbero dovuto approvare solo i principi del catalogo, ma lo sviluppo delle norme, la loro eventuale revisione e le modalità tecniche di esecuzione avrebbero dovuto essere lasciate all'autonomia del bibliotecario responsabile del progetto.

Negli ultimi mesi del 1838 due eventi in particolare influiscono sul cambiamento della strategia. Innanzitutto, la reiterata richiesta del Segretario del British Museum, Josiah Forshall, affinché Panizzi sottoscrivesse le *16 Regole di Baber*, redatte nel maggio 1834 da estrapolazione dall'omonimo Rapporto come riferimento per il lavoro di catalogazione dello staff.<sup>4</sup>

In secondo luogo, la lunga *querelle* con Forshall (*Papers* 1847, 75-78) sull'indicazione data da Panizzi a Richard Garnett (1789-1850) di intestare le opere anonime – in caso di difficoltà a individuare "some prominent or leading word in the title as the heading to be prefixed to the title" (regola XIII Baber) – alla prima parola del titolo che non fosse preposizione o articolo. Si comprende la scelta di Panizzi (*Papers* 1847, 90), per evitare lo stillicidio di un confronto isolato sull'una o l'altra integrazione e garantire coerenza e uniformità al nuovo catalogo, di proporre l'8 marzo 1839 ai Trustees l'elaborazione di un insieme strutturato di norme. Il giorno successivo i Trustees approvano la proposta. Il 18 marzo 1839 Panizzi, nel presentarne ai Trustees il testo, fa presente la momentanea assenza di regole relative a pubblicazioni "of so peculiar a nature as to form exceptions ... Academies, Almanacks, Calendars or Ephemerides, Biblia, including the Old and New Testament, Liturgies, and probably Councils and periodicals in general" (*Papers* 1847, 91), impegnandosi a produrle successivamente.

Sono le 73 regole.

Al momento della presentazione il draft non copre dunque tutto quanto sarà definito nelle 91 regole. Tematicamente le 73 corrispondono alle prime 78 delle 91. È un punto essenziale. Una lettura frettolosa dell'audizione di Panizzi, in particolare dell'inciso "The rules then were 79, and they came to be 91, by all the additions and alterations made in Sub-Committee", può fare ritenere

<sup>&</sup>lt;sup>3</sup> La ricostruzione documentale del dibattito interno al British Museum dalle decisioni dei Trustees negli anni Settanta del secolo XVIII sino al 1847 in *Papers relating to the Alphabetical Catalogue of printed books*. Private and confidential. London: George Woodfall and Son, 1847 (<a href="http://dbooks.bodleian.ox.ac.uk/books/PDFs/590119436.pdf">http://dbooks.bodleian.ox.ac.uk/books/PDFs/590119436.pdf</a>). D'ora in poi: *Papers* 1847.

<sup>&</sup>lt;sup>4</sup> Cfr. Papers 1847, 53: doc. n. 8.



erroneamente che le 73 fossero complete e il processo di analisi e revisione del draft avesse portato a modificazione e aggiunte tali da condurre a 91 regole.

Panizzi più volte rivendica l'elaborazione del draft, per quanto già in questa fase molto concentrata (8-18 marzo) si sia avvalso dell'aiuto e del confronto con i più stretti collaboratori. Afferma a tale proposito (Panizzi 1847a, 9) alcuni anni dopo: "[The Rules] included all Mr. Baber's rules in distinct type [...] those which I had drawn up myself were but a development of the principles laid down by M. Baber in his Report of April, 1834, already well known to the Trustees and approved by them. In drawing up these rules I had asked, obtained and profited by the suggestions of all the gentlemen in the Museum generally, most conversant with the subject". Ricostruzione confermata nella citata audizione del 13 luglio 1848: "When they had been printed, in March, I gave a copy to every gentleman in the house who understood the business, and who had the goodness to assist me, to have the advantage of their suggestions before submitting the rules to the Trustees" (*Report* 1850, Q. 4117). Il testo delle 73, tuttavia, non è soltanto il frutto di una fase particolarmente densa di elaborazione: esso riflette, come vedremo, materiali e talora incorpora testi preesistenti già proposti in report da Panizzi.

## A1) La struttura testuale: verso un nuovo genere

Panizzi rende visibili nel testo delle 73 le 16 Regole di Baber: la doppia numerazione (n. regola / n. Baber) presente in caso di coincidenza o correlazione testuale (anche parziale) fra i due testi permette di evidenziarne il legame, talora l'identità di formulazione, oppure le differenze espositive e di sviluppo argomentativo. La scelta permette di presentare la proposta come approfondimento necessario di un set troppo limitato di norme. Il testo non appare così come completamente nuovo e, in quanto tale, potenzialmente estraneo, mantenendo una connessione con nuclei normativi ritenuti dai Trustees intoccabili sino a poco tempo prima.

Il testo delle 73 non segue la progressione delle 16 Regole. Queste, scisse dal contesto di provenienza, si presentavano in una sequenzialità tutt'altro che trasparente. La loro destrutturazione nel nuovo testo proposto da Panizzi era inevitabile proprio perché esso potesse manifestare da subito un'articolazione logica leggibile. Siamo all'inizio della costruzione di un codice di regole modernamente inteso, un nuovo autonomo genere testuale, oggetto di progressiva riconoscibilità sociale proprio a partire dalla sfida di Panizzi. La nascita di questo nuovo genere coincide non casualmente con l'inizio di grandi progetti di catalogazione di rilevanza nazionale o internazionale, come quello del British Museum o, alcuni anni dopo, il progetto di catalogo cooperativo ipotizzato da Charles Jewett.

Il problema storico-critico che tentiamo di affrontare è dunque molteplice e procederemo per differenza e tracce, talora in modo non lineare. Dovremo quindi:

contestualizzare il processo di elaborazione delle regole incrociando due prospettive temporali: una che potremmo definire di breve durata, ricostruendo il momento e gli attori di un confronto che si fa intenso fra il 1834 e il 1839; e un'altra, di più ampio arco temporale, tesa a ricostruire il conflitto, spesso non esplicito, di culture bibliografiche, le linee di tradizione e a far emergere l'originalità dell'approccio di Panizzi, con una differenza fra i modelli di riferimento della sua riflessione e quelli che invece definiamo i presupposti, il background culturale e metodologico che ha contribuito a strutturare in Panizzi habitus interpretativi;



- tentare di ricostruire il modello di catalogo che sottostà al testo delle 73 regole, stante i vincoli posti dai Trustees. Il modello non coincide con la concezione del catalogo di Panizzi. Le 73 sono un testo di mediazione, anche se, in quanto draft, di una mediazione meno segnata di quella che troveremo nelle 91.
- verificare nel passaggio da 73 a 91 se e quali modifiche vi siano state a tale disegno.

Panizzi durante le audizioni definì le 91 regole più volte come Rules of the Trustees. È così? È solamente così? Possiamo assumere senza mediazioni, in sede storiografica, la valutazione di un solo attore, fosse anche il principale? Non è l'interrogativo meno importante.

# B) Il processo di formazione: storia, modelli e presupposti

## B1) Le parole del catalogo

Non possiamo parlare di una tradizione catalografica unitaria inglese fra il secolo XVIII e il XIX quasi che vi fosse sul tema del catalogo e della sua funzione una condivisione di principi o riferimenti. Una tradizione, nel senso di insieme consapevole di valori e criteri per una grande biblioteca di ricerca, inizia a formarsi ad Oxford presso la Bodleiana sin da 1605. Se la mission della biblioteca è alta e riconosciuta come tale, innanzitutto dal fondatore, Sir Thomas Bodley, lo strumento di mediazione, il catalogo, è all'epoca assai meno evoluto. Il vero cambiamento è rappresentato dal catalogo della Bodleiana di Thomas Hyde (1674). Con la *Praefatio ad lectorem* di Hyde si può parlare dell'avvio di una riflessione catalografica capace di delineare precisi riferimenti teorici e metodologici e di ricercare strutture nuove di organizzazione del catalogo tali da corrispondere ai bisogni di ricerca dei lettori-studiosi. Il successivo catalogo di Fysher (1738), evoluzione del modello hydeano, è esemplare da questo punto di vista.

Potremmo parlare per l'Inghilterra del tardo Settecento di *linee di tradizione plurime*, dando alla parola *tradizione* un'accezione composita, fondata innanzitutto sulla "attribuzione di valore a qualche tratto tramandato" (Alessandro Cavalli). Essa si manifesta nello sviluppo di un policentrismo, sia pure inizialmente molto limitato, databile all'inizio degli anni Settanta del Settecento, circa un decennio dopo l'apertura del British Museum. A fine anni Sessanta si pone il problema del catalogo delle collezioni del British Museum, e primariamente delle collezioni bibliografiche (Guerrini e Neri 2020). La parola che domina il dibattito e le decisioni sulla preparazione del catalogo è *compendious*. Il 23 febbraio 1771 i Trustees decidono di orientare l'impegno del British Museum verso un "useful and compendious Catalogue of Printed Books" (*Papers* 1847, ii). Una decisione del 6 dicembre 1777 chiarisce i confini del catalogo. In merito alla catalogazione di opuscoli e pamphlet della King's Library si stabilisce che "only the general title of every collection of works be entered in the Catalogue, without specifying the particular authors or tracts therein contained" (*Papers* 1847, iii). Sono descritte solo le unità fisiche considerate bibliograficamente e gestionalmente indipendenti. È una soluzione in netta controtendenza rispetto alla tradizione della Bodleiana.

<sup>&</sup>lt;sup>5</sup> Alessandro Cavalli, *Tradizione* (s.v.), *Enciclopedia delle scienze sociali* (<a href="https://www.treccani.it/enciclopedia/tradizione">https://www.treccani.it/enciclopedia/tradizione</a> (<a href="https://www.treccani.it/enciclopedia/tradizione</a> (<a href="https://www.treccani.it/enciclopedia/tradizione</a> (<a href="https://www.treccani.it/enciclopedia/tradizione</a> (<

<sup>&</sup>lt;sup>6</sup> Per l'analisi di contesto Willison 1989 e Willison 1996; Goldgar 2000. Per lo sviluppo del tema del catalogo delle collezioni MacNeil 2016 e MacNeil 2017.



Thomas Hyde aveva orientato il catalogo a obiettivi di ampia copertura bibliografica, con la scelta di segnalare, in presenza di volumi miscellanei, anche tutti gli opuscoli individualmente e i loro autori, "ne forte aliqui authorunculi, una cum aliis compacti, ibidem latitarent". L'insoddisfazione per la valenza descrittiva del catalogo in folio porta agli inizi del 1800 alla necessità di un nuovo catalogo, pubblicato da Ellis e Baber fra 1813 e 1819.

Negli anni Trenta emergono coppie semantiche oppositive, come *full/short; long/short* o sinonimiche, come *compendious/short*. *Long* e *full* non si equivalgono: *long* è correlato, anzitutto, alla maggiore ampiezza di trascrizione del titolo. *Long titles* possono convivere in un catalogo *compendious* perché la loro lunghezza è selettiva, riconducibile spesso a pubblicazioni di particolare rarità o a esemplari nobilitati dal pregio del donatore. L'accezione del termine, non a caso utilizzata dagli oppositori di Panizzi, è quasi unicamente quantitativa.

Per Panizzi le parole del decennio 1830-1839 sono, in relazioni oppositive e di affinità:

- *full* vs. *short*: il primo termine esprime la ricchezza descrittiva e connettiva coerente con la mission della biblioteca e con il suo intento di segnalare *works* (e responsabilità intellettuali correlate) e non solo *books*. *Short* denota il catalogo come finding list.
- *accurate* è qualificazione sia del linguaggio di trascrizione che della precisione e qualità delle informazioni. Esprime attenzione all'analisi e alla registrazione degli elementi informativi delle fonti; rispetto rigoroso del linguaggio originario della pubblicazione; cura dell'esattezza nelle informazioni aggiuntive con cui un catalogatore, se necessario, arricchisce la registrazione catalografica.

L'accoppiata *full* and *accurate* diviene la bandiera dell'innovatore Panizzi proprio perché il catalogo è il prodotto di analisi e ricerca intellettuale.

Le ultime parole del decennio sono *uniformity, consistency* e *principles. Consistency* e *uniformity* esprimono la qualità di un catalogo bibliografico vista come coerenza del suo linguaggio interno nella relazione con il pubblico. La coerenza e l'uniformità costruiscono quella anticipazione positiva, quella convinzione di elevata possibilità di una risposta soddisfacente che è il modo con cui si reinventa la grande tradizione del *commodum lectoris* di Hyde e Audiffredi. Un catalogo siffatto non è possibile se non basato su un insieme di principi e criteri condivisi e su una rigorosa coerenza e prassi interpretativa: e *principles* nell'accezione dell'epoca copre sia l'ambito dei criteri costitutivi che quello delle norme ritenute di carattere più generale in quanto correlate alle più frequenti situazioni bibliografiche. Infine nel dibattito degli anni 1845-1850 una parola, *remodelling*, le riassume tutte nell'orizzonte innovativo: "A Catalogue is remodelled when the plan and principles on which it was formerly compiled are so altered as to make it necessary to alter all the titles" (Panizzi 1847b, 22).

# **B2)** Worthy of the Nation

Fra il 1833 e il 1836, il British Museum è oggetto di riflessione in sede parlamentare e sulla stampa. Con le elezioni del 1833, successive al Reform Act del 1832, entrano alla Camera dei Comuni, oltre a una molto più forte rappresentanza whig, nuclei sia pure limitati di parlamentari radicali. Fra questi, William Cobbett (1763-1835). Fra le sue prime battaglie parlamentari, quella contro la gestione del British Museum e i suoi presunti sprechi. Il frequente uso del termine *curiosities* 

\_

<sup>&</sup>lt;sup>7</sup> Cfr. Hansard, v. 16, 25 marzo 1833, col. 1003-1005 (interventi di Cobbett e altri).



nel dibattito parlamentare (marzo-agosto 1833) può illuminare la discussione. Per i difensori del ruolo e delle potenzialità del Museo come grande istituzione culturale *curiosities* erano gli oggetti contrassegnati da ricercatezza e originalità in misura non usuale. La grandezza del Museo stava nel rappresentare una collezione non ripetitiva di materiali di tipologia varia (libri, oggettistica d'arte e scientifica, storia naturale, manoscritti). Per Cobbett queste erano invece *fineries*, oggetti di una eleganza affettata e autocelebrativa, di nessuna utilità rispetto ai bisogni di una popolazione che versava in condizioni economiche e sociali di crescente disagio. Nel dibattito erano state poste questioni di fondo e si ritenne che dovessero essere oggetto di indagine approfondita aspetti quali l'efficienza dei servizi al pubblico, gli orari di apertura, il ruolo stesso del British Museum. Temi che saranno ripresi nel 1835 con la proposta di istituire una Commissione d'Inchiesta (*Select Committee*) che opererà nel biennio 1835-1836.

All'inizio del 1836 i giochi sembrano riaprirsi rispetto al confronto di due anni prima. Scrive Panizzi (1847a, 4): "The subject was again considered in the beginning of the year 1836. Mr. Baber made a Report thereon, which I have not seen; but he informed me that he had deprecated the departure from his first plan of execution". <sup>8</sup> Al ritorno dal viaggio di studio nelle più importanti biblioteche del continente, Panizzi incontra i Trustees. Dietro loro richiesta, presenta il già citato report del 23 febbraio 1836. Sia questo che la successiva audizione del 7 giugno 1836 davanti al Select Committe potrebbero essere definiti spectacular, <sup>9</sup> per la capacità di visione strategica, l'esattezza del quadro comparativo, la cura del particolare. Nel 1836 ha modo di emergere il metodo di Panizzi: la capacità di connettere nell'esposizione critica e in ogni passaggio espositivo il quadro generale e la verifica del dettaglio. Un metodo e un processo argomentativo sostanzialmente antiretorici. L'orizzonte della fullness / accuracy del catalogo è collocato in un contesto di interesse nazionale.

The library of the British Museum does not possess such an alphabetical catalog as the public have a right to expect in such an institution. The work has never been attended on a scale worthy of the nation [...] The library of the British Museum [...] as a national library, in London, [...] is not worthy of the nation (*Papers* 1847, 53 e 55).

Questa parola, *worthy*, nel significato di "corresponding to the worth of", attraversa rapporto e audizione in una duplice correlazione: il catalogo deve essere *worthy* dell'istituzione, e questa deve essere *worthy* della nazione britannica. Panizzi, forzando il *Rapporto Baber*, espone per la prima volta la propria concezione del catalogo:

The titles of books should be written if not in full at least nearly so; no alteration should ever be made in the words of the title itself; all useful information as to the writers of prefaces, biographers, translators, editors, and annotators should be added (in brackets) when omitted in the title of the book; the same should be done with regard to names and titles of honour; anonymous publications should be entered according to the first word of the title, and pseudonymous inserted under the assumed name; no care should be spared to discover the real names which ought to be added (in brackets); collections of every

<sup>&</sup>lt;sup>8</sup> Sul tema Miller 1967, 116.

<sup>&</sup>lt;sup>9</sup> È la definizione lubetzkiana per le audizioni di Panizzi davanti alla Royal Commission nel biennio 1848-1849 (Lubetzky 1953, 1 e 55).



description should be entered either under the collector's name or as anonymous works, and their contents entered as distinct works; and the same should be done with respect to transactions and acts of societies. There should be, if I may so express myself, a profusion of cross-references. (*Papers* 1847, 54)

Tuttavia né le sollecitazioni di Baber, né la loro intensa riproposizione da parte di Panizzi, e neppure l'apertura del Select Committee a prospettive di rinnovamento, vedono da parte dei Trustees un ripensamento delle scelte compiute. Paradossalmente nel 1837 gli spazi anziché aprirsi, sembrano chiudersi.

# B3) L'invenzione delle 16 Regole di Baber

Per Panizzi, che aveva collaborato strettamente con Baber nella elaborazione del *Report*, la trasposizione di un nucleo a più definito contenuto normativo in regole poteva rappresentare niente più che un utile riferimento per il lavoro dello staff: non le riteneva un testo isolato e autosufficiente. L'audizione di John Humphreys Parry<sup>10</sup> dell'8 marzo 1849 (*Report* 1850, Q. 7319) conferma quanto affermato a suo tempo da Panizzi nel report dell'8 marzo 1839 sulla natura inevitabilmente limitata dell'estrapolazione normativa.

When I first came there were some written rules, and Mr. Edwards and I, who worked together at the same table in the library, had those written rules which we had to apply. Those written rules consisted, I believe, at that time, of not more than 12 or 13: they were very imperfect.

Erano le Regole di Baber, e *imperfect* ha qui il significato di "carente", un'inadeguatezza di copertura rispetto alla varietà e complessità delle situazioni bibliografiche. È nel 1837, poco dopo il suo pensionamento, che le 16 Regole di Baber, sino ad allora nell'arco di tre anni mai nominate in quanto tali, acquisiscono la sacralità di testo fondante e autosufficiente. Ciò è possibile, anzitutto, perché acquisiscono un nome, e ciò si verifica poco dopo la nomina di Panizzi a Keeper del Department of Printed Books (luglio 1837). Nelle comunicazioni ufficiali del Segretario a Panizzi o nei verbali delle riunioni dei Trustees il testo elaborato da Baber viene nominato come Rules circulated by Mr. Baber for the formation of the new catalog, oppure (e diverrà prevalente) Rules laid down/established by Mr. Baber for the formation of the new catalog. Assistiamo così alla comparsa di una entità nuova, le 16 Regole di Baber, e alla scomparsa, invece, di un'altra entità, la sola in realtà "generally approved" da parte dei Trustees: il Baber Report. Con la scomparsa del Rapporto dalla memoria ufficiale (sarà solo Panizzi a farvi continuamente e convintamente riferimento) l'estrapolazione normativa cambia natura a invarianza di testo: diventa le 16 Regole di Baber. Per questo, adattando una celebre formulazione di Eric Hobsbawm e Terence Ranger (1983), l'"invenzione della tradizione", possiamo legittimamente parlare di invenzione delle 16 Regole. Una invenzione culturale: il conferimento della valenza di testo definitorio a un sottoprodotto di un testo di origine ben più complesso, con la contemporanea rimozione di quest'ultimo. L'invenzione di una tradizione può corrispondere a esigenze differenziate che vanno dalla costruzione di una comunità specifica, al suo rafforzamento o ad azioni di mantenimento e difesa. Dietro l'invenzione delle

\_

<sup>&</sup>lt;sup>10</sup> Parry aveva lavorato al Department of Printed Books da gennaio 1839 a luglio 1843.



16 Regole vi è un presupposto implicito: la riduzione della complessità dell'orizzonte delineato da Baber prima e poi, con più nettezza, da Panizzi nel 1836. Essa diviene lo strumento costantemente riproposto nel biennio 1837-1839 per affermare la continuità di una presunta tradizione contro una temuta discontinuità. Ma, dobbiamo interrogarci, era difesa di una tradizione o di una prassi? O, meglio ancora, di una linea di tradizione consolidata in prassi?

## B5) Modelli e Presupposti

#### B5.1) I modelli

Quando nel report del 23 febbraio 1836 Panizzi presenta ai Trustees il suo progetto di costruzione del nuovo catalogo alfabetico i suoi riferimenti sono Hyde, "but still more Audiffredi" (*Papers* 1847, 54).

## Giovan Battista Audiffredi

Nel Catalogo della Casanatense (1761-1797) Panizzi vedeva realizzato, pur nell'incompiutezza dell'opera, il sogno di un catalogo che fosse al tempo stesso repertorio bibliografico, strumento di correlazione e individuazione di opere e edizioni, rappresentazione dell'universo della produzione letteraria di un autore. La complessità del catalogo di Audiffredi e la straordinaria ricchezza descrittiva e di analisi bibliografica non erano probabilmente riproducibili considerando l'ampiezza e la varietà delle collezioni del British Museum. Il riferimento ad Audiffredi è costante, esplicito, puntuale dal 1836 sino al 1850: il catalogo (e la prefazione) del grande domenicano operano inoltre come *modello motore*, anche nella diversità di talune soluzioni tecniche. Un modello inoltre, secondo Panizzi, in grado di corrispondere – data la struttura articolata di correlazione e rinvii – al dilemma divenuto ormai pressante del trattamento delle pubblicazioni anonime: la soluzione di Audiffredi consistendo nella scelta quale intestazione della prima parola del titolo che non fosse preposizione o articolo. Un modello tuttavia che, oltre a confliggere con la tradizione inglese degli ultimi sessanta anni, era italiano e sostenuto in successione dai tre grandi bibliografi francesi della prima metà del secolo: Antoine-Alexander Barbier, Jacques-Charles Brunet e Joseph-Marie Querard.

Brunet (1810) nel *Manuel du libraire et de l'amateur de livres* aveva definito il Catalogo della Casanatense "chef-d'oeuvre bibliographique". Valutazione condivisa da Ebert, che lo ritenne "unubertrussne Muster eines alphabetisches Nominalkat". Barbier (1822, vol. 1: xxii) nella seconda edizione del *Dictionnaire*, affrontando il problema del trattamento citazionale delle opere anonime, illustra – condividendolo – il metodo scelto da Audiffredi "[...] dans le magnifique catalogue de la bibliothéque Casanate".

Panizzi nei confronti dei suoi interlocutori a partire dal 1836 usa con intelligente accortezza e con variazione di accenti sia richiami alla tradizione di Hyde, Fysher e Audiffredi che a quella di Brunet, Barbier e Querard. Alla difficoltà di recepimento, segnatamente fra 1836 e 1844, corrisponde dal 1845 una capacità di ascolto più aperta, in un periodo in cui dinamiche di cultura colta e di

-

<sup>&</sup>lt;sup>11</sup> Friedrich Adolf Ebert, *Die Bieldung des Bibliothecars*, 1: A-L (Leipzig, 1820): n. 1354, p. 115. Nel 1837 era apparsa la traduzione inglese dell'opera (*A general bibliographical dictionary*, vol.1: A-E. Oxford: At the University Press).



aggiornamento sul dibattito culturale europeo si confrontano con la più tradizionale cultura bibliofila e bibliografica.

## Thomas Hyde

La prima novità della *Praefatio ad lectorem* sta nella definizione della funzione stessa del catalogo: non mera segnalazione di semplici titoli, ma strumento di ricerca che accompagna nella ricerca gli studiosi con dispositivi indicali adeguati. La produzione di un catalogo è opera complessa, irta di difficoltà, che nascono dalla natura stessa dell'impresa: la rappresentazione della ricchezza culturale, tematica e linguistica di una collezione. La necessità di una rigorosa e preventiva analisi bibliografica e letteraria è la premessa della riflessione di Hyde. Catalogare non è trascrivere dati da una fonte, né è processo meccanico come molti ingenuamente pensano: "Quid enim inquiunt facilius est quam inspecta librorum fronte, eorundem titulos excribere?" (p. i).<sup>12</sup> Il tema della coerenza metodologica e della ricerca di uniformità diviene centrale per la prima volta nella storia della catalogazione. L'intento è dichiarato quasi subito dal punto di vista del metodo e dell'obiettivo. La polemica col predecessore Thomas James non potrebbe essere più netta: "In Jamesiano catalogo nulla certa lege aut ordine, sed ut fors tulit, recensentur" (p. vi).

Il primo elemento di complessità è dato dalla varietà dei nomi di un medesimo autore nella concretezza della produzione bibliografica e dalla omonimia di molti autori. Compito del catalogatore – ed è funzione di rara complessità tecnica e culturale – è quello di *rappresentare* come di *riconciliare* le differenze. Il suo è un lavoro di ricerca soggetto a inevitabili errori di correlazione fra un possibile autore e un titolo ("proclive est hujus auctoris librum illius nomini ascribere", p. ii), ma è troppo facile nell'incertezza l'adozione di un'intestazione di raggruppamento come *anonymi*.

Analoga innovazione nell'analisi di Hyde sulla varietà dei titoli. Non solo una medesima opera, in altra espressione linguistica, cambia titolo, ma il titolo dell'opera – a parità di espressione linguistica – può differire da una edizione all'altra e addirittura risultarne profondamente mutata l'indicazione stessa dei temi trattati. Né si dovrà rimproverare il catalogatore per eventuali asprezze nella formulazione dei titoli: il catalogatore deve rappresentare il linguaggio dell'edizione, non adattarlo. Come il richiamo al metodo ha introdotto la *Praefatio*, così ne sostanzia la parte finale, insieme all'interesse-beneficio (*commodum*) del lettore-studioso che ne ha rappresentato il filo conduttore. La *Praefatio* si chiude con un saluto che è anche un invito: "Vale Lector, fruere!".

Negli interventi di Panizzi del 1836 non sono ancora citati i temi "caldi" della *Praefatio* di Hyde. Si può ragionevolmente ipotizzare che all'epoca non ne avesse fatto oggetto di una lettura approfondita, concentrandosi innanzitutto sul "prodotto", il catalogo del 1674. Poco dopo la nomina a Keeper, nel report del 12 gennaio 1838 (*Papers* 1847, 61-67) questi temi saranno invece compiutamente recuperati, con un meditato sviluppo di correlazioni fra le sfide del presente e la ricchezza dell'eredità intellettuale e metodologica di Hyde, Fysher e Audiffredi.

#### **B5.2**) I presupposti

Girolamo Tiraboschi

Panizzi arriva a Londra nel maggio 1823. Qui ha subito modo di conoscere alcuni delle figure più

<sup>&</sup>lt;sup>12</sup> Le pagine della *Praefatio* non sono numerate.



rappresentative fra gli esuli italiani: Ugo Foscolo, Santorre di Santarosa, Giuseppe Pecchio. Si trasferisce a Liverpool nell'estate del 1823. Il 5 settembre 1823 Santarosa scrive a Panizzi da Londra una lettera affettuosa e densa di consigli, comprensiva del momento di difficoltà dell'amico. E lo invita a studiare:

Se vi sono librerie a Liverpool o se ella può accattar libri, la conforto pure a rimettersi bene nella memoria la nostra istoria politica e letteraria. Muratori e Tiraboschi vogliono essere letti e studiati un poco, massime per certe epoche più rilevanti. Mi pare che il farsi perfettamente famigliare la nostra duplice storia sia il miglior modo di manifestare la differenza che passa fra maestro e maestro (Fagan 1880b, 12).

Non era un consiglio scontato. Nella galassia degli esuli la lezione di Tiraboschi era oggetto di interpretazioni talora contrapposte. Santarosa, Pecchio e Panizzi lo considerano una lettura fondamentale anche, forse, in nome di quella serietà dell'impegno dell'esule che essi rivendicano, a maggior ragione nella prospettiva di un esilio di lunga durata. Al polo opposto della valutazione critica su Tiraboschi, Camillo Ugoni (1822, vol. 3: 350-377): pur riconoscendone i profondi meriti di erudizione critica e la vastità delle conoscenze, vede il limite principale in una mancanza di "filosofia", di una visione d'insieme necessaria a connettere e interpretare fatti diversi. Foscolo (1826) in *On the antiquarians and critics of Italian history*, attenuando le critiche avanzate alla scuola erudita nell'orazione *Della origine e dell'uffizio della letteratura* (1809), recuperava il valore della ricerca storico-filologica di Tiraboschi definendolo con Muratori *cold genius*, genio freddo. Michele Mari (1990 e 1999) ha messo in evidenza il legame di Tiraboschi con la storiografia muratoriana e le grandi novità metodologico-critiche della *Storia della letteratura italiana*:

- il principio di *accertamento*: l'insofferenza verso ogni forma di impostura, auctoritas e tradizione non fondata e la verifica dell'esattezza dei dati nella correlazione.
- una concezione lata della *letteratura*: essa coincide con la produzione intellettuale e artistica nel suo insieme. Un criterio di inclusione recupera, come elemento strutturale della narrazione e della partizione stessa della *Storia della letteratura italiana*, il ruolo delle *istituzioni*: culturali, scientifiche, letterarie.

Tiraboschi è Prefetto (1770-1794) della Biblioteca Ducale di Modena, precedentemente diretta (1700-1750) da Ludovico Antonio Muratori. A Parma il primo bibliotecario della Palatina, Paolo Maria Paciaudi (1762-1771; 1772-74) ha relazioni assai strette con la cultura bibliografica francese dell'epoca; il successore Ireneo Arfò fu per oltre un ventennio in rapporti strettissimi con Tiraboschi. Un elemento di comunanza fra tutte queste personalità è rappresentato dal tema della "de bibliothecarum utilitate". Angelo Pezzana, direttore della Biblioteca Palatina dal 1810 e conosciuto da Panizzi negli anni della sua frequenza all'Università di Parma, condivide questo orizzonte metodologico e culturale: l'attenzione alle collezioni ed alla loro valutazione bibliografica e culturale; la coltivazione della cultura bibliografica, disciplina che egli insegna all'Università; la ricerca storica-erudita; l'utilità sociale delle biblioteche per una comunità.

Il consiglio di Santarosa trova dunque in Panizzi un interlocutore attento e già ben disposto, in un ambiente culturale nuovo in cui la lezione di Tiraboschi ha da tempo significativi estimatori. William Roscoe (1806, v. 1: xvii-xviii), storico e mentore di Panizzi negli anni di Liverpool, aveva a

<sup>13</sup> De bibliothecarum utilitate è il titolo di una orazione giovanile inedita (1762) di Tiraboschi: cfr. Tinti 2014.



suo tempo definito la *Storia della letteratura italiana* "the immortal work of Tiraboschi; the noblest specimen of that species of composition which any age or country has produced". Valutazione condivisa anni dopo da Henry Hallam (1777-1859): "In one country of Europe, and only in one, we find a national history so comprehensive as to leave uncommemorated no part of literary labour. This was first executed by Tiraboschi" (1837, vol. 1: viii).

Negli intensi anni di studi di italianistica di Panizzi, prima a Liverpool poi a Londra, questa dimensione antiretorica e antiromantica (senza essere classicistica) si rafforza ulteriormente.

# Intermezzo fra anonimia e pseudonimia

Nel 1802 viene fondata la *Edinburgh Review*, quadrimestrale di area culturale e politica whig. La sua influenza è difficilmente sottovalutabile. Ne sono segnate le riviste rivali, la tory *Quarterly Review*, sorta nel 1809, e la *Westminster Review*, nata nel 1824 per opera dei philosophical radicals Jeremy Bentham e James Mill.

La *Edinburgh Review* fonda lo stile della review-saggio critico: un contributo molto ampio, sino a 45-50 pagine, solo numerato, privo di titolo d'insieme formale (il titolo è presente solo nella forma di head title) e che, muovendo da un numero variabile (dall'unità a 14-15) di libri immediatamente citati in premessa, tesse la review. È un modello di rivista colta, pensato per una classe dirigente cui vuole offrire un orizzonte non locale: politica nazionale ed estera; politica economica; relazioni internazionali; storia. Per la dimensione di literary review l'accezione è molto estesa: letteratura, istituzioni culturali, arte, filosofia.

Un elemento accomuna tutta la stampa periodica inglese, dai bimestrali in su, nei primi sessanta anni del secolo: il rigoroso rispetto dell'*anonimato* di tutti i contributi. Una scelta, culturale e politica, condivisa da schieramenti pure fortemente oppositivi: la rivendicazione di un diritto di critica che è del periodico prima ancora che del singolo. La *Edinburgh Review* manterrà la scelta dell'anonimato sino al 1912; le altre grandi riviste la attenueranno a partire dagli anni Sessanta sino ad abbandonarla nel corso degli anni Ottanta. Da un punto di vista storico-critico, ciò pone rilevanti problemi di identificazione degli autori. È un tema che necessita approfondimenti ulteriori per comprendere le possibili connessioni culturali con le prassi catalografiche prevalenti all'epoca in materia di anonimia e di pseudonimia.

#### Brother in bibliography

Sir, I dedicate this notice of the life and works of one of the greatest of bibliographers to you, without asking your permission, for different reasons: you might not care to have your name placed on a pamphlet thus; and I wish to place it side by side with that of J.-M. Quérard [...] I dedicate this to you from a feeling of respect, not only for yourself, but for your departed brother bibliographer (Thomas 1867, 5).

È la dedica di edizione a Panizzi del bibliografo inglese Ralph Thomas (1840-1909?), che spesso si firmava con il *nom de plume* di Olphar Hamst negli scritti di natura bibliografica. Il mito di Panizzi all'epoca è consolidato da tempo. Querard era già ben noto ed apprezzato dal Keeper, quando – racconta Thomas – incontrò Panizzi nel 1844. E questi lo propose "as a proper person for a post in the Library of the British Museum". Proposta che purtroppo non ebbe esito positivo. L'efficace qualificazione coniata da Thomas, "brother in bibliography", coglie un aspetto essenziale nella



correlazione fra Panizzi e la scuola bibliografica francese. Panizzi, che durante la preparazione (1834-1836) del Catalogo dei libri della Royal Society (Biagetti 2001, 27) aveva ampiamente utilizzato *La France littéraire* di Querard (1827-1839), in *A Letter to His Royal Highness, The President of the Royal Society* così identifica le principali verifiche necessarie nella revisione di un catalogo prima della stampa:

Before, however, such a work could be called *ready for press*, it required at least four *revises*. The title of a book, its size, date, &c., are all positive *facts*, which admit of no fanciful correction: the same is to be said of the names, surnames, &c., of the author; the peculiar or antiquated ortography, &c. Add to this, all the difficulties of classification (Panizzi 1837, 27).

Nel repertorio di Quérard trovano compiuta espressione le quattro funzioni fondanti di una registrazione bibliografica e catalografica così care a Panizzi nella sua non formalistica interpretazione del principio di *accuracy*:

- una scrupuleuse exactitude<sup>14</sup> è il primo dei criteri costruttivi e valutativi adottati da Querard nei confronti della produzione bibliografica propria e altrui. Un principio di accertamento che lo porterà ad affrontare con più rigore metodologico di Barbier il tema delle opere anonime, con una chiara distinzione fra quelle che lo risultano effettivamente alla ricognizione critico-bibliografica, e quelle che invece tali si presentano solo in specifiche edizioni (Querard 1827, vol.1, xxii).
- la congruità e precisione nella conservazione del linguaggio peculiare dell'edizione.
- la coerenza metodologica;
- la funzione connettiva.

# C) 73 vs 91: un confronto

La momentanea assenza nelle 73 di norme riguardanti la Bibbia e specifiche categorie di pubblicazioni non era dovuta *solo* ad una carenza di tempo nell'elaborare un testo compiuto. Due erano i nuclei tematici di snodo di più complessa problematicità: la definizione di *corporate body* a fini catalografici e il trattamento delle *anonymous works/anonymous publications*. L'eventuale accoglimento oppure l'alterazione del testo proposto da Panizzi su tali aspetti avrebbe avuto un effetto inevitabile sulle regole ancora non prodotte e avrebbe contribuito a definire la fisionomia complessiva del nuovo catalogo.

Nel confronto sono comparati alcuni dei blocchi testuali principali, evidenziando i cambiamenti e le dislocazioni. Nella citazione e comparazione delle norme, seguiremo la sequenza: n. romano (norma) / n. arabo (73 o 91) per contestualizzare la singola regola o gruppo di regole. *I titoli dei nuclei sono nostri*.

<sup>&</sup>lt;sup>14</sup> Espressione già presente in Barbier; diverrà uno stilema della produzione bibliografica di Querard.



# Regola IX/73: corporate bodies

Assemblee o enti in quanto autori. Per la prima volta una norma catalografica, sia pure un draft, riconosce la responsabilità autoriale a un ente collettivo. È la prima di una serie di *norme di snodo:* 

Assemblies or corporate bodies (with the exception of Academies, Universities and Learned Societies, respecting which special rules are to be framed) to be considered as authors of any act, vote, resolution or other document purporting to be agreed upon, authorized or issued by them.

#### mentre IX/91:

Any act, resolution, or other document purporting to be agreed upon, authorized, or issued by assemblies, boards, or corporate bodies, (with the exception of academies, universities, learned societies, and religious orders, respecting which special rules are to be followed,) to be entered in distinct alphabetical series, under the name of the country or place from which they derive their denomination, or, for want of such denomination, under the name of the place whence their acts are issued.

La differenza è radicale. Nel testo IX/73, con le eccezioni di "Academies, Universities and Learned Societies", agli enti è riconosciuto – a determinate condizioni – lo status di autore. L'intestazione all'ente non ha una funzione di *caption* (etichetta) strumentale al recupero di un titolo in una collezione, come in una prassi assai diffusa, ma è il riconoscimento di un legame assimilabile a quello che nelle regole II-VIII/73 connette un titolo a un autore personale. Il richiamo, a fianco della norma, agli esempi n. 26, 76-77, 79, esplicita ancora più la portata dell'innovazione rispetto alla tradizionale prassi catalografica inglese che prevedeva l'intestazione non nella forma diretta al nome dell'istituzione/ente, ma al nome dello stato o della città seguito dal nome dell'ente/autorità che aveva emanato gli atti o le risoluzioni. Con la soluzione panizziana l'intestazione era costituita dalla denominazione specifica dell'assemblea o dell'istituzione. La regola IX/91 segna invece una conferma della prassi storica sia nella forma che nel ruolo attribuito all'intestazione all'ente, cui non viene riconosciuta una funzione di authorship rispetto ad atti, decisioni o testi emanati. Inoltre si restringe l'ambito di contenuto della norma: non solo essa non riguarda "Academies, Universities and Learned Societies", ma anche ordini religiosi.

### Regole XIX- XXXIII/73: Descrizione bibliografica

Le norme XIX-XX/73 rappresentano la filosofia generale della descrizione bibliografica della pubblicazione. La regola XIX/73 riproduce sostanzialmente la regola VII Baber, nelle sue polarità: trascrizione sintetica del titolo, tale da non pregiudicare la segnalazione dei vari apporti intellettuali e dei possibili ambiti tematici, ma trascrizione filologica. Alla norma Baber originaria, ad accentuare il rispetto rigoroso delle fonti ai fini dell'individuazione delle peculiarità della pubblicazione, sono aggiunte due nuove prescrizioni:

- "the original ortography to be preserved";
- "the number of the edition to be stated when appearing in the title".

### La regola XX/73 recita:

If it be found unnecessary to transcribe the title at full length, and the omission extend to more than a few insignificant words, dots to be substituted to indicate the omission, more particularly in old and rare volumes, in which, however, abridgements should be seldom resorted to.



La formulazione è tale che, in linea di principio, presupporrebbe una regola generale precedente orientata alla *fullness* della registrazione. Il fatto è che la regola XIX/73 e la successiva XX/73 appartengono a due orizzonti descrittivi non omogenei: la regola XIX/73 è ancora interna alla tradizione, sia pure rivisitata. La norma XX/73 – con un ribaltamento assai trasparente – rende eccezione quello che invece nella regola VII Baber era l'indicazione prioritaria, una sintesi descrittiva rispettosa del linguaggio dell'edizione. La regola XX/73 venne eliminata dal testo definitivo (*Report* 1850, Q. 3882, p. 211). Sul piano *formale* tale scelta rese il testo più riconducibile alle linee della tradizione del British Museum, per quanto aggiornata. Nella sostanza del processo di realizzazione del catalogo Panizzi si sforzò di procedere *come se* la XX/73 fosse operante.

#### Regola XXXIV – XLI /73: Anonymous works – Pseudonymous publications

Negli anni Trenta inglesi non si è ancora strutturata la supremazia dei dati di una fonte (il frontespizio) rispetto alle altre interne alla pubblicazione. Un autore, almeno nella linea di tradizione allora prevalente, è identificato come tale in connessione a un titolo se questa correlazione è esplicitata nella pubblicazione nel suo insieme. È la pubblicazione nei suoi paratesti di riferimento (frontespizio, dedica, prefazione...) che costruisce memoria culturale nel lettore, e ne condiziona le strategie di ricerca. A fini catalografici quindi un'opera di cui si conosca attraverso fonti esterne l'autore può essere ritenuta anonima se la correlazione titolo/autore è assente nella pubblicazione. La scelta di connettere, tramite rinvii, l'autore accertato o conosciuto alla pubblicazione anonima non modifica sostanzialmente l'attribuzione di anonimato. Su questo non vi è dissenso da parte di Panizzi. Come evidenziato a suo tempo da Lubetsky, 15 in alcune regole difese da Panizzi davanti alla Royal Commission il potenziale conflitto fra il rispetto dell'obiettività segnica e l'individuazione letteraria delle responsabilità intellettuali, in particolare nel caso di pubblicazioni anonime o nei casi in cui l'autore sia contrassegnato esclusivamente da iniziali, rischia di disperdere edizioni diverse di una medesima opera di un autore. Non così in Hyde e Fysher, tanto meno in Audiffredi. L'anonimato è condizione dell'opera, non dell'edizione, e - come tale - da accertare anche su fonti esterne alla pubblicazione stessa. Proprio in nome di una coerente riconciliazione fra catalografia e bibliografia, il catalogo audiffrediano segnala sotto una comune intestazione di autore (se identificato) tutte le edizioni delle sue opere, anche se – riconosce Audiffredi nella *Praefatio* al primo volume (1761, xii) del Catalogo della Casanatense – "frequenter accidit, ut opera anonyma, vel quasi anonyma, sub propris titulis retulerimus, aliorum non paucorum celeberrimorum bibliographorum exemplo". La letteratura bibliografica francese mantiene su questo punto significativa incertezza. Barbier (1806, 1: i-ii) aveva affermato nella prima edizione del *Dictionnaire*: "On appelle ouvrage *anonyme*, celui sur le frontespice duquel l'auteur n'est nommé [...]. Un ouvrage pseudonyme est celui dont le frontespice contient un nom qui n'est celui de son auteur". Barbier ritornò sulla questione nella seconda edizione (1822), segnalando la differenza fra la propria definizione, di fatto connessa strettamente alla dimensione della storia bibliografica, e quella del Dictionnaire della Académie francaise, sub voce *Anonyme* (5th ed., 1798, 1: 61), connessa invece alla dimensione letteraria: "Qui est sans nome. Il ne se dit que des auteurs dont on ne sait point le nom, et des écrits dont on ne connoit point l'auteur".

<sup>15</sup> Lubetzky 1953, 37-38; Report 1850, Q. 9838; Petrucciani 1984, 18-19.



Sul tema l'originaria regola XIII Baber recitava:

In case of anonymous works, some prominent or leading word in the title to be selected as the heading to be prefixed to the title. If the author's name should be known or conjectured by the Librarian, the same to be inserted at the end of the title, and included in brackets.

Panizzi riteneva sin dal 1834 tale indicazione di ardua applicabilità, soggetta a discrezionalità e incertezze tali da rendere inutilmente più complesso, lento e costoso il lavoro di catalogazione. Il 10 gennaio 1839, formalizzando una indicazione verbale data ai propri collaboratori, proponeva ai Trustees un paragrafo aggiuntivo alla regola XIII di Baber:

Whenever any doubt arises as to which may be deemed the leading word in the title of anonymous writings, the first word in the title (except it be an article or a preposition) should be preferred, and the work entered alphabetically under it, adding, however, as many cross-references as may be considered necessary from any other word which might be taken as a leading one (*Papers* 1847, 78).

La riformulazione dei Trustees condusse (21 gennaio 1839) a un ben diverso testo:

Whenever no leading word can be fixed upon in the title of anonymous writings, the first word in the title except it be an article or a preposition, may be taken and the work entered alphabetically under it, adding, however, as many cross-references as may be considered necessary from any other word (*Papers* 1847, 80).

La differenza è sostanziale. La premessa di Panizzi significava appena verificata l'impraticabilità; l'altra, verificata definitivamente l'impraticabilità. La scelta dei Trustees inevitabilmente apriva la strada alla individuazione casistica delle possibili tipologie di leading word. Il vincolo posto da tale decisione era molto forte, e rendeva particolarmente complessa la ricerca da parte di Panizzi di una formulazione normativa adeguata nelle regole successive.

La norma XXXIV/73 coincide con la regola XIII Baber riformulata dai Trustees. Essa può essere destrutturata in due blocchi:

- 1. "In case of anonymous works, some prominent or leading word in the title to be selected as the heading to be prefixed to the title. If the author's name should be known or conjectured by the Librarian, the same to be inserted at the end of the title and included in brackets".
- 2. "Whenever no leading word can be fixed upon in the title of anonymous writings, the first word in the title except it be an article or a preposition, may be taken and the work entered alphabetically under it, adding, however, as many cross-references as may be considered necessary from any other word".

Non vi sono esempi a illustrarla: si tratta di una regola quadro, i cui ambiti dovevano essere precisati da norme successive. La regola che segue (XXXV/73) è l'esplicitazione dei criteri specifici d'individuazione della leading word da scegliere come intestazione, a condizione che tale parola sia formalmente nominata nella pubblicazione. Emerge con chiarezza come le funzioni svolte dalla leading word prescelta non siano necessariamente connesse a ruoli di responsabilità intellettuale: se ciò si verifica nel caso di curatore/editor o traduttore, in tutti gli altri casi l'intestazione ha una funzione latamente indicizzatoria con l'intento di non lasciare la pubblicazione orfana nel catalogo.



La regola XXXVI/73 ("Works published under initials, to be entered under the last of them") è solo apparentemente incongrua, anche se spezza la linea argomentativa iniziata con la regola XXXIV/73. La presenza delle sole iniziali, per quanto assimilata a una indicazione di autore, configura una situazione di assai complessa identificabilità: l'intestazione è registrata sotto l'ultima delle iniziali, anche qualora si sia in grado di identificare più compiutamente l'autore.

La regola XXXVII/73 dettaglia, nell'impossibilità di individuare una leading word, l'applicazione del criterio "the first word in the title except it be an article or a preposition". Nella struttura espositiva delle 73 Regole, la relativa vicinanza testuale della soluzione audiffrediana (regola XXXVII/73) rispetto all'esposizione del tema delle anonymous works (regola XXIV-XXXV/73) rende assai più agevolmente riconoscibile tale possibile scelta. Panizzi, con un calibrato sviluppo argomentativo, ha tentato di non renderla residuale e quindi di contenere i danni dell'improvvida riformulazione da parte dei Trustees della regola XIII Baber.

L'articolazione testuale delle 91 è profondamente diversa. Non vi è una norma generale che si faccia carico di esplicitare sia il criterio prevalente che una possibile soluzione alternativa nell'impossibilità di applicazione del primo. Prevale da subito nettamente un orientamento casistico. A partire dalla norma XXXIII/91 sino a XXXVII/91 si affrontano situazioni diverse di assenza del nome di un autore nella pubblicazione. Non si usa l'espressione leading word, anche se ciascuna regola di fatto esamina una specifica leading word potenziale, di volta in volta individuata nel nome della persona di cui si tratta, in un ente citato nel titolo, in un luogo ecc. Solo alla fine la norma XXXVIII/91:

In the case of anonymous works, to which none of the foregoing rules can be applied, the first substantive in the title (or if there be no substantive, the first word) to be selected as the heading. A substantive, adjectively used, to be taken in conjunction with its following substantive as forming one word; and the same to be done with respect to adjectives incorporated with their following substantive. The entries which may occur under the same heading to succeed each other in strict alphabetical order.

Regole XLII – XLVIII/73 Raccolte di testi, Traduzioni, Commenti

La regola XLII/73 è un'altra norma di snodo: svolge la funzione prioritaria di delineare tematicamente le politiche catalografiche dell'istituto su un'asse di lungo periodo.

Works of several writers, collectively published, to be entered according to the following rules, and the separate pieces of the various authors included in the collection to be separately entered in the order in which they occur; excepting merely collections of letters, charters, short extracts from larger works, and similar compilations.

Il principio è presente nel Rapporto Baber (*Papers* 1847, 42), per quanto l'applicazione vi sia limitata a raccolte tematizzate, con l'aggiunta: "It would be highly advantageous to the student that the titles of such collected productions should be in some way specified". La segnalazione dovrebbe realizzarsi nella forma di syllabus, di indice dei testi contenuti, senza necessariamente tradursi in registrazioni distinte. La regola XLII/73 generalizza l'indicazione del Rapporto Baber nell'orizzonte di un catalogo che intenda dare ragione di tutti i testi di una collezione, e non solo delle unità fisiche e gestionalmente indipendenti.



Panizzi nel report del 23 febbraio 1836 era andato oltre, sostenendo che le singole opere di una raccolta dovessero essere oggetto di registrazioni separate.

Nella norma corrispondente delle 91 (XLIV/91), testualmente identica, il testo che segue "following rules" sino a "compilations" è collocato entro parentesi quadre con l'aggiunta: "That part of the foregoing rule which is inserted between bracket has not been acted upon, in order to accelerate the printing of the catalogue". Significa che la politica indicizzatoria di lungo periodo non viene attuata per accelerare la realizzazione del catalogo a stampa.

# Regole XLIX – LXIV/73 Rinvii (Cross-references)

La costruzione di una articolata rete di rinvii rappresenta uno dei nuclei di più significativo sviluppo espositivo nel testo delle 73 e nella versione definitiva delle regole. La norma XLIX/73 (con identico testo in LIV/91) costituisce la premessa alla sezione sui rinvii dal punto di vista dei principi di economia e chiarezza del catalogo: "No work ever to be entered twice at full length. Whenever requisite cross references to be introduced".

La successiva L/73, di ambito generale, distingue due tipologie di rinvii: from name to name e from work to work. Nei secondi la connessione fra un nome personale (oggetto di rinvio) e un'opera (in cui figuri, per esempio, nel ruolo di curatore, traduttore ecc.) prevede una registrazione degli elementi essenziali di identificazione della pubblicazione: titolo (e nel titolo la segnalazione dell'oggetto del rinvio), formato e data.

Le norme LI-LXIII/73 dettagliano le due tipologie:

- from name to name: regole LI-LV/73;
- from work to work: regole LVI- LXIII/73.

Una struttura così articolata riflette una riflessione e un confronto di lungo periodo. Il Sub-Committee, tramite il Segretario Forshall, aveva chiesto a Panizzi il 2 febbraio 1839 (*Papers* 1847, 82) di illustrare "in the form of rules supplementary to the numbered III the principles upon which he considers the names of authors presenting variations in the titles of their books or in ordinary usage should be entered in the catalogue". Con i due report successivi dell'8 febbraio 1839 e del 21 febbraio 1839 Panizzi presenta proposte articolate in forma di future norme di dettaglio:

- la lingua di servizio del catalogo, che Panizzi propone sia l'inglese (*Papers* 1847, 82-83);
- la formalizzazione della regola utilizzata per i nomi preceduti da preposizione, articolo o da entrambi. Diventerà, con pressoché nulle modifiche, la regola XIII/73 (*Papers* 1847, 83-85);
- una articolazione complessa delle due tipologie di cross-references: from name to name e from work to work. Diventeranno, con scarse modifiche, proprio le regole L-LXIII/73 secondo la struttura indicata nel Report del 21 febbraio.

#### Quel che mancava nelle 73

Dal testo delle 73 regole mancavano "Academies, Almanacks, Calendars or Ephemerides; Biblia, including the Old and New Testament, Liturgies, and probably Councils and periodicals in general". Esaminiamone il trattamento nelle 91, per vedere se la logica normativa sia o meno coerente con alcune premesse e indicazioni del draft.



# Form headings e group headings

Form e Group headings hanno una lunga tradizione. La refrattarietà per lungo tempo a un trattamento catalografico per le opere anonime diverso dalla leading word e la preferenza per talune tipologie di form headings e di group (class) headings, dura a lungo, di fatto sino ai primi anni Ottanta del secolo XIX.

Hyde nella *Praefatio* illustra gli ambiti di adozione dei *generalia capitula*: "Sub generalibus capitulis qualia sunt Lexica, Concordantiae, Jus, Concilia, etc. exhibui librorum eo spectantium pleniorem copiam" (p. v). L'adozione di tali tipologie intestative cambia nel tempo e da catalogo e catalogo: se Hyde usa Lexica, Concordantiae, Jus, Concilia, Statuta, abbandona però Acta, Leges, Synoda utilizzati prima da James. E per alcuni (Concilia, Statuta) con un uso più parco che nella precedente prassi (Verona 1962, 297). La grande tradizione della Bodleiana tende a contenerne l'uso. Panizzi non nutriva una particolare simpatia per le intestazioni con funzioni di raggruppamento. Non condivideva l'adozione, in un catalogo alfabetico che avesse il suo focus sugli autori, i titoli e le opere, di intestazioni di natura semantica. E la ricerca della prominent word del titolo spesso si traduceva nella ricerca della parola più significativa in un orizzonte semantico. Nel testo delle 73, il combinato disposto della regola IX/73, della regola XXXIV/73 e di un'acuta organizzazione strutturale che individuava contiguità, differenze e livelli conduceva prima ancora che a un alleggerimento quantitativo delle regole, a un alleggerimento casistico e, con esso, alla riduzione mirata di soluzioni intestative per specifiche categorie di pubblicazioni o di group headings. Rispetto alle convinzioni espresse a suo tempo da Panizzi nel report del 23 febbraio 1836, vi era già stata nelle 73 una concessione significativa alla tradizione, o comunque il riconoscimento dei vincoli di contesto. Il catalogo prospettato allora (Papers 1847, 54) riduceva fortemente le intestazioni non riconducibili all'autore o al titolo: "collections of every description should be entered either under the collector's name or as anonymous works, and their contents entered as distinct works; and the same should be done with respect to transactions and acts of societies". E sotto Societies deve leggersi il trittico Academies, Universities, Learned Societies.

La struttura normativa nelle 91 per Academies, Universities, Learned Societies, riproposta nella sua articolazione interna anche per i Periodicals, fu oggetto durante le audizioni della Royal Commission di critiche significative proprio in ragione della sua farraginosità e non trasparenza per il lettore. Nel testo finale delle 91 continueranno ad essere assenti norme specifiche per i Councils, ma in aggiunta vi saranno norme per Dictionaries (LXXXVIII/91) e Encyclopaedias (LXXXIX/91), laddove anonimi, e Catalogues (LXXXV – LXXXVII/91).

Che necessità vi era di una norma specifica per i dizionari e le enciclopedie anonime? Se anonime, si sarebbero applicati – laddove possibile – i criteri di cui a XXXV-XXXVI/73: per un dizionario o un'enciclopedia, qualora nel libro fosse assente una segnalazione di autore, si sarebbe potuto utilizzare come intestazione il curatore (editor). Altrimenti si sarebbe ricorsi alla regola XXXVII/73, individuando come intestazione il primo sostantivo del titolo. Analogamente per i cataloghi.



# Conclusioni

Jamais [...] des règles semblabes n'ont été communiquèes au public, avec autant d'etendue, par les rédacteurs de catalogues de bibliothéque ayant cette importance pour on travail de cette importance et de cette durée on ne doit pas s'en rapporter à des instructions verbales

(Olivier Barbier, rec. a Catalogue of printed books in the British Museum. Vol. 1 (A), 1841. *Bulletin du Bibliophile, Septième sèrie*, 1845, n. 6 (juin), 242-247: cit., p. 243)

La recensione di Barbier rappresenta un unicum all'epoca nel confronto dialettico fra due tradizioni bibliografiche e catalografiche. Barbier dedica uno spazio molto ampio alla pubblicazione delle 91 Regole: prima la recensione; poi nei numeri successivi di luglio e agosto la loro traduzione integrale. Alla fine della traduzione un impegno: "Nous donnerons dans le prochain numéro les examples auxquels renvoie l'astérisque placé à la fin de quelques unes des regles". 16 Il filo conduttore della recensione è quello di uno stupore critico ammirato nei confronti dell'impresa compiuta da Panizzi, "le savant et activ bibliothecaire du British Museum". Competenza, energia, profonda cultura: qualificazioni che troveremo anche in Querard nella breve narrazione del suo incontro con Panizzi, "honorable et savant directeur". 17 Barbier riconosce la novità assoluta della pubblicazione di un insieme di regole così completo e articolato, a maggior ragione per una biblioteca di tale importanza. Si è dinanzi ad una transizione fondamentale: il passaggio da norme e procedure interne di catalogazione, spesso verbali, a un documento di rara complessità reso noto, con il catalogo, alla comunità dei professionisti (bibliotecari e bibliografi) e degli utilizzatori dei cataloghi. Il precedente è così significativo, aggiungiamo, proprio perché inatteso nella forma comunicativa: la forma codice, sia pure ancora in una fase iniziale e non ancora nominata in quanto tale, segna l'ingresso nella modernità ottocentesca della riflessione catalografica. E lo segna con una maturità, per l'epoca, straordinaria. Barbier conclude la recensione con una citazione da Constantin Hesse (1839, 96) sulla complessità di costruzione di un catalogo: "Il faut se livrer à l'ouvrage comme si la perfection étoit chose possible, et cette illusion peut seule faire faire quelque chose de bon dans ce genre". Miglior complimento non poteva essere fatto a Panizzi. Erano anni di spinte culturali fortemente contraddittorie: scambi tecnici, culturali e professionali fra bibliotecari e bibliografi convivevano con rivendicazioni, in Europa, di tradizioni e primati nazionalistici. L'orizzonte metodologico proposto da Barbier non ebbe seguito negli anni immediatamente successivi. Nel 1850 apparve nel Bulletin un intervento di Gustave Brunet (1850), ma le riflessioni risultavano fuori tempo massimo non tanto per la relativa lontananza dalla pubblicazione del primo volume del catalogo del British Museum, quanto per l'obsolescenza dell'approccio.<sup>18</sup>

<sup>&</sup>lt;sup>16</sup> La pubblicazione degli esempi non fu mai edita nel *Bulletin* né, ci risulta, altrove. Non si sarebbe trattato di pochi esempi: avrebbero riguardate le regole 2-12, 14-17, 20-21, 32-33, 35-44, 47-49, 56-68, con focus su scelta e forma dell'intestazione, pubblicazioni anonime e rete di rinvii.

<sup>&</sup>lt;sup>17</sup> Querard 1827-1863, 11 (1854), s.v., autobiografica "Querard Joseph-Marie", 588-598; cit. p. 591. La voce è siglata con il *nom de plume* prevalente di Querard, "Mar. Jozon d'Erquar".

<sup>&</sup>lt;sup>18</sup> Per le reazioni e il dibattito coevo in Europa sulle 91 Regole cfr. Schacht 1979-1980, in particolare Pt. 2 (1980).



L'esame dei testi ha tentato di cogliere il disegno di Panizzi e gli intrecci e i conflitti di culture e modelli bibliografici in un periodo particolarmente denso. Al fondo delle 73 Regole c'è una scommessa intellettuale: nei vincoli dati, andare oltre i confini della tradizione inglese dell'epoca con l'apertura ad altre culture bibliografiche. Queste culture erano rappresentate dal recupero della grande tradizione di Hyde e Fysher e innanzitutto dal Catalogo della Casanatense di Audiffredi. Il corto circuito più stimolante proviene dalla lezione di Audiffredi mediata da un principio di exactitude che, nei suoi riferimenti più immediati, ha radici nelle punte più alte della cultura bibliografica francese ottocentesca (Barbier e Querard), ma che trae la sua linfa dalla tradizione erudita del Settecento italiano (Tiraboschi). È il progetto di catalogo per una grande biblioteca nazionale di ricerca: "We do not make catalogues for Mr. Bolton Corney: this is a great national undertaking", aveva affermato con forza e orgoglio Panizzi dinanzi alla Royal Commission (Report 1850, Q. 9751). A Bolton, che aveva sostenuto che il nuovo catalogo era realizzato su un piano "too elaborate" e che un tale dettaglio non era necessario neppure per i cataloghi delle collezioni speciali del British Museum (Report 1850, Q. 6127), Panizzi contrapponeva l'idea di un catalogo costruito su una prospettiva di generazioni. Il suo catalogo worthy of the nation significava un'impresa bibliografica di lungo periodo, che tentasse – come nei grandi modelli a lui cari – di riconciliare bibliografia, ricerca e catalografia.

La mediazione che venne proposta con le 73, in una articolazione completamente nuova per l'epoca, era orientata a una interpretazione non casistica e, per il modello di catalogo, a un catalogo alfabetico strutturato per autore e titolo. Non dobbiamo ritenere che le 73 corrispondessero compiutamente all'idea di catalogo di Panizzi: essa era più ampia e coerente della mediazione proposta nel draft e, a maggior ragione, nelle 91. I suoi nuclei fondanti sono esplicitati nella Lettera a Lord Ellesmere (Panizzi 1848), testo esemplare di ricostruzione del processo di analisi bibliografica e letteraria, nelle audizioni del 1848-1849, ma proposti già nelle linee essenziali nei report del 23 febbraio 1836 e del 12 gennaio 1838. Un testo come le 91, costitutivo di un nuovo genere non può mai essere oggetto di una lettura esclusivamente tecnica, come l'esame del passaggio dal draft al testo finale non poteva essere esclusivamente tecnico. Nel momento in cui le norme divengono pubbliche, incontrano comunità diversificate di professionisti, non solo nazionali. Le 91 sono un *testo di fase*. Come disse Lubetzky: "Panizzi raised the catalog to a new level in its evolution". 19

Una lettera di Jewett del 29 aprile 1847, di calorosi apprezzamenti al funzionamento della Biblioteca del British Museum venne prodotta da Panizzi durante l'audizione del 6 febbraio 1849 (*Report* 1850, Q. 4293: p. 265-266). Jewett si era espresso in merito alla polemica contro Panizzi, sostenendo con forza le ragioni del Keeper e affermando a proposito del nuovo catalogo alfabetico del British Museum: "It should be a work of bibliographical authority". *Bibliographical authority*, *undertaking*, *monument of bibliographical labour and learning*: osono tutte qualificazioni, in momenti diversi, dell'impresa catalografica perseguita da Panizzi, la costruzione di un catalogo che durasse nel tempo proprio perché strumento di ricerca fondato su un sapere incorporato (*learning*). Fagan ricorda l'invito di Jewett a Panizzi a partecipare alla First General Convention of Librarians che

<sup>&</sup>lt;sup>19</sup> Lubetzky 1979 (citato in Lubetsky 2001, 347-348); anche Lubetzky 2001, 421: "The man who started it all".

<sup>&</sup>lt;sup>20</sup> Quest'ultima qualificazione in Nichols 1865, 35.



si sarebbe tenuta a New York dal 15 al 17 settembre 1853.<sup>21</sup> Panizzi (Fagan 1880a, vol. 1, 170-171) ne scrisse all'amico Francis Haywood (1796-1858):

I have to tell you of a dream, which I should like to become a reality. There is going to be a Congress of Librarians in the United States, which is to open on the 15th of September next, and where all the great questions connected with the management of a great Library are to be discussed and uniform principles adopted. The Americans have always been my friends, and the principles which will prevail are mine.

A New York si sarebbe discusso del grande progetto di catalogo cooperativo incentrato sullo Smithsonian Institution come soggetto coordinatore e sulle regole di On the construction of catalogues of libraries, ben noto alla comunità bibliotecaria americana. Si sarebbe discusso di quei principi. E se il testo di Jewett fosse stato, agli occhi di Panizzi, una possibile incarnazione, pure mediata in tutt'altro contesto e con soluzioni talora difformi, dei suoi principi? E l'affermazione di Panizzi ("the principles which will prevail are mine") fosse una sorta di indiretta rivendicazione di una vicinanza e di una contiguità e, forse, di una continuazione ideale?

<sup>&</sup>lt;sup>21</sup> Cfr. Librarians' Convention 1853.



# Riferimenti bibliografici

Barbier, Antoine-Alexandre. 1806-1809. *Dictionnaire des ouvrages anonymes et pseudonymes*. 4 voll. Paris: Imprimerie bibliographique.

Barbier, Antoine-Alexandre. 1822-1827. Dictionnaire des ouvrages anonymes et pseudonymes. Second édition, revue, corrigée et considérablement augmentée. 4 voll. Paris: Imprimerie bibliographique.

Biagetti, Maria Teresa. 2001. Teoria e prassi della catalogazione nominale: i contributi di Panizzi, Jewett e Cutter. Roma: Bulzoni.

Brunet, Pierre-Gustave. 1850. "Quelques mots au sujet des difficultés que présente la catalographie." *Bulletin du Bibliophile, Neuvième sèrie* 16: 563-67.

Carpenter, Michael, and Elaine Svenonius, eds. 1985. Foundations of cataloging: a sourcebook. Littleton, Colorado: Libraries unlimited.

Carpenter, Michael. 2002. "The original 73 Rules of the British Museum: a preliminary analysis." *Cataloging & Classification Quarterly* 35 (1-2): 23-36. https://doi.org/10.1300/J104v35n01\_03.

Fagan, Louis. 1880a. The life of Sir Anthony Panizzi K.C.B., late principal librarian of the British Museum, senator of Italy. 2 vols. London: Remington & Co.

Fagan, Louis, a c. di. 1880b. Lettere ad Antonio Panizzi di uomini illustri e di amici italiani: (1823-1870). Pubblicate da Luigi Fagan. Firenze: G. Barbera.

Foscolo, Ugo. 1826. "On the antiquarians and critics of Italian history." *Retrospective Review* 14, pt.1: 136-53.

Foscolo, Ugo. (1826) 2012. *Antiquarj e critici (On the antiquarians and critics)*. Edizione critica bilingue a cura di Paolo Borsa. Milano: Ledizioni.

Goldgar, Anne. 2000. "The British Museum and the virtual representation of culture in the Eighteenth century." *Albion: a quarterly journal concerned with British studies* 32 (2): 195-231.

Guerrini, Mauro, e Franco Neri. 2020. "La tormentata formulazione delle Regole del British Museum del 1839." In *Scaffali come segmenti di storia: studi in onore di Vincenzo Trombetta*. A cura di Rosa Parlavecchia e Paola Zito, 153-65. Roma: Edizioni Quasar.

Hallam, Henry. 1837-1839. *Introduction to the literature of Europe in the Fifteenth, Sixteenth and Seventeenth centuries*. 4 voll. London: J. Murray.

Hesse, Leopold Auguste Constantin. 1839. Bibliothéconomie. Paris: Techener.

Hobsbawm, Eric, e Terence Ranger, a c. di. 1983. *The invention of tradition*. Cambridge: Cambridge University Press.

Librarians' Convention. 1853. "Proceedings of the Librarians' Convention, held in New York City, September 15, 16 and 17. 1853." *Norton's literary gazette and publishers' circular* 3, no. 10 (15 October 1853): 170-176. Anche in *Norton's library and education register for 1854*: 49-94.

Lubetzky, Seymour. 1953. Cataloging rules and principles: a critique of the A.L.A. rules for entry and a proposed design for their revision. Washington, D.C.: Processing Department, Library of Congress.



Lubetzky, Seymour. 1956. "Panizzi vs. the Finding catalog." *Journal of cataloging and classification* 12 (July 1956): 152-56. Anche in Carpenter, Michael, and Elaine Svenonius, a c. di. 1985: 173-79.

Lubetzky, Seymour. 1979. "Ideology of bibliographic cataloging: progress and retrogression." In *The nature and future of the catalog.* A cura di M.J. Friedman e Michael Malinconico, 5-19. Phoenix, AZ: Onyx Press. Anche in Lubetzky, Seymour. 2001: 345-66.

Lubetzky, Seymour. 1999. "On the use of form headings in an alphabetic catalog." *Library Quarterly* 69: 222-32.

Lubetzky, Seymour. 2001. Writings on the classical art of cataloging. Compiled and edited by Elaine Svenonius and Dorothy McGarry. Englewood: Libraries Unlimited.

Mari, Michele. 1990. "Il genio freddo: la storiografia letteraria di Girolamo Tiraboschi." N. monografico di *Bergamon: bollettino della civica Biblioteca A. Mai di Bergamo*, 80, n. 4.

Mari, Michele. 1999. Il genio freddo: la storiografia letteraria di Girolamo Tiraboschi. Milano: CUEM.

MacNeil, Heather. 2016. "Catalogues and the collecting and ordering of knowledge (I): ca. 1550–1750." *Archivaria* 82 (Fall): 27-53. <a href="https://archivaria.ca/index.php/archivaria/article/view/13580">https://archivaria.ca/index.php/archivaria/article/view/13580</a>.

MacNeil, Heather. 2017. "Catalogues and the collecting and ordering of knowledge (II): debates about cataloguing practices in the British Museum and the forebears of the Public Record Office of Great Britain, ca. 1750–1850." *Archivaria* 84 (Fall): 1-35. <a href="https://archivaria.ca/index.php/archivaria/article/view/13612">https://archivaria.ca/index.php/archivaria/article/view/13612</a>.

Miller Edward. 1967. Prince of librarians: the life and times of Antonio Panizzi of the British Museum. London: Andre Deutsch.

Nichols, Thomas. 1866. A handbook for readers in the British Museum. London: Longmans, Green and Company.

Panizzi, Antonio. 1837. A Letter to His Royal Highness, the President of the Royal Society, on the new catalogue of the library of that institution now in the press. London: C. Whittingham.

Panizzi, Antonio. 1847a. "Observations on the Report of Sir H. Ellis, dated Nov. 14Th, 1846." In *Papers* 1847: 1-19.

Panizzi, Antonio. 1847b. "Observations on two reports of Sir H. Ellis, one dated April 30th 1834, and the other February 11th, 1809." In *Papers* 1847: 20-39.

Panizzi, Antonio. 1848. "Mr. Panizzi to the Right Hon. The Earl of Ellesmere." In Carpenter, Michael, and Elaine Svenonius, a c. di. 1985: 18-47.

Papers relating to the Alphabetical Catalogue of printed books. 1847. Private and confidential. London: George Woodfall and Son. <a href="http://dbooks.bodleian.ox.ac.uk/books/PDFs/590119436.pdf">http://dbooks.bodleian.ox.ac.uk/books/PDFs/590119436.pdf</a>.

Petrucciani, Alberto. 1984. Funzione e struttura del catalogo per autore. Firenze: Giunta regionale Toscana & La nuova Italia.

Querard, Joseph-Marie. 1827-1864. *La France littéraire, ou Dictionnaire bibliographique des savants, bistoriens et gens de lettres de la France.* 12 voll. Paris: Firmin Didot.



Report from the Select Committee on British Museum: together with the Minutes of evidence, Appendix and Index. 1836. London: [s.n.].

Report of the Commissioners appointed to inquire into the constitution and government of the British Museum: with Minutes of evidence. 1850. London: William Clowes.

Roscoe, William. 1806. *The Life and Pontificate of Leo the Tenth*. Second edition corrected. London: printed by McCreery for T. Caddell and W. Davies.

Schacht, Sigrid. 1979-80. "Antonio Panizzi's librarianship and its representation in the pioneering library journals." Part. 1 Libri 29 (4): 273-310; Part. 2 Libri 30 (3): 177-231.

Thomas, Ralph. 1867. A Martyr of bibliography: a notice of the life and works of Joseph-Marie Querard, bibliographer, by Olphar Hamst. London: John Russell Smith.

Tinti, Paolo. 2014. "Lo spazio della biblioteca nelle lettere di Girolamo Tiraboschi (1731-1794)." In *Cartas-Lettere: discursos, practicas y representationes epistolares (siglos XIV-XX)*, a cura di Antonio Castillo Gomez e Veronica Sierra Blas, 111-24. Alcaklà de Henaves: Universidad de Alcalà.

Ugoni, Camillo. 1820-1822. *Della letteratura italiana nella seconda metà del secolo xviii.* 3 voll. Brescia: per Nicolò Bettoni.

Verona, Eva. 1956. "A historical approach to corporate bodies." Libri 7 (1): 1-40.

Verona, Eva. 1962. "Form headings in catalogues of the past and the present." *Library resources and technical services* 6 (Fall): 295-317.

Willison, Ian R. 1989. "The national library in historical perspective." *Libraries & Culture* 24 (1): 75-95.

Willison, Ian R. 1996. "The development of the British National Library to 1837 in its European context." *Library history* 12: 31-48.