

Wikidata-Enhanced Authority Records: A Project for Personal Names in SBN

Elena Ravelli^(a)

a) Central Institute for the Union Catalog of Italian Libraries and for Bibliographic Information (ICCU), <https://orcid.org/0000-0001-6402-2039>

Contact: Elena Ravelli, elena.ravelli@cultura.gov.it

Received: 16 July 2025; **Accepted:** 18 October 2025; **First Published:** 15 January 2026

ABSTRACT

The catalogue of the National Library Service (SBN), developed and expanded through cooperative cataloguing, has experienced exponential growth over time. Currently, the titles archive contains over 21 million records, while the authors archive includes around 4,200,000 entries. However, not all authority records in SBN meet the level of completeness and accuracy required of a national catalogue. As part of the collaboration between Wikimedia Italy and ICCU, a project was launched to enrich a portion of the authority records related to personal names in SBN that lacked information. This enrichment was made possible thanks to data available on Wikidata, where many records include the SBN identifier. The number of entries in Wikidata associated with SBN identifiers is expected to grow further, also thanks to ICCU's recent decision to make all records related to personal and corporate names in the Index visible in the SBN OPAC, regardless of their authority level.

KEYWORDS

Wikidata; Opac SBN; Authority control; Interoperability.

Il progetto di arricchimento dei nomi di persona di SBN con i dati di Wikidata

ABSTRACT

Il catalogo del Servizio bibliotecario nazionale (SBN), sviluppato e ampliato attraverso la catalogazione partecipata, ha registrato nel tempo una crescita esponenziale. Attualmente, l'archivio titoli conta oltre 21 milioni di record, mentre l'archivio autori raccoglie circa 4.200.000 voci. Tuttavia, non tutte le registrazioni di autorità presenti in SBN possiedono il livello di completezza e accuratezza richiesto da un catalogo nazionale. Nell'ambito della collaborazione tra Wikimedia Italia e l'ICCU è stato avviato un progetto volto ad arricchire una parte delle voci di autorità relative ai nomi di persona che in SBN risultavano prive di informazioni. L'arricchimento è stato possibile grazie ai dati disponibili su Wikidata, dove molti record contengono l'identificativo SBN. Il numero di voci in Wikidata associate a identificativi SBN è destinato a crescere, anche grazie alla recente decisione dell'ICCU di rendere visibili nell'OPAC SBN tutte le registrazioni relative ai nomi di persona e di ente presenti nell'Indice, a prescindere dal loro livello di autorità.

PAROLE CHIAVE

Wikidata; Opac SBN; Authority control; Interoperabilità.

Introduzione

Il progetto di arricchimento dei record relativi ai nomi di persona presenti nel Catalogo del Servizio Bibliotecario Nazionale (SBN) con dati di Wikidata si inserisce nel quadro della convenzione fra l'Istituto centrale per il catalogo unico delle biblioteche italiane (ICCU) e Wikimedia Italia (WMI).¹ L'accordo, oltre a sancire un impegno condiviso nella diffusione dei principi dell'open access e dell'open culture, promuove l'integrazione dei dati e dei materiali dei progetti dell'ICCU con quelli di WMI. Tra le prime azioni realizzate, vi è stata l'integrazione dei dati dell'Anagrafe delle Biblioteche Italiane (ABI)² con Wikidata e OpenStreetMap.³ Un'altra significativa attività, sviluppata tra il 2015 e il 2019, durante la collaborazione di un wikipediano in residenza presso l'ICCU, ha riguardato il collegamento tra le entità di Wikidata e i record relativi ai nomi di persona presenti nell'OPAC di SBN, relazione che è basilare per il progetto qui descritto. A quegli anni risale anche la decisione dell'ICCU di utilizzare una piattaforma Mediawiki per la pubblicazione delle normative per la catalogazione in SBN,⁴ uno strumento che consente una più agevole modalità di aggiornamento dei contenuti e, aspetto più importante, l'accesso libero e gratuito alle normative da parte di tutti i catalogatori e bibliotecari italiani. Uno degli impieghi più recenti dei dati di Wikidata riguarda Alfabetica,⁵ l'ultimo portale sviluppato dall'ICCU. Al suo interno, il percorso "Protagonisti" offre brevi profili degli autori, tratti dalla Nota informativa del record SBN. Quando questa nota non è presente, le informazioni vengono recuperate da Wikidata. Da Wikidata provengono sempre – tramite query SPARQL sull'endpoint LOD e senza alcun sistema di caching – anche il luogo e le date di nascita e di morte, oltre all'immagine del protagonista.⁶ Negli ultimi anni si è stabilito un dialogo attivo e proficuo tra l'ICCU e alcuni membri della comunità Wiki, con un focus particolare sull'Authority file di SBN. Da questa collaborazione sono nate segnalazioni sistematiche da parte dei collaboratori di Wikidata⁷ all'Ufficio Authority di SBN in merito a errori e duplicazioni presenti nel Catalogo nazionale. Questo confronto si è concretizzato in un percorso di formazione da parte dell'ICCU sulle principali funzionalità dell'applicativo di Indice, in modo da consentire a un membro del gruppo Wikidata di intervenire direttamente sulle registrazioni, correggendo le voci di autorità presenti nel Catalogo.

¹ La Convenzione ICCU Wikimedia Italia è stata sottoscritta per la prima volta nel 2015 (ICCU e Wikimedia Italia 2015) poi è stata rinnovata nel 2018 (ICCU, e Wikimedia Italia 2018). Attualmente è in vigore la Convenzione del 2022 (ICCU e Wikimedia Italia 2015).

² ABI: <https://anagrafe.iccu.sbn.it/it/>

³ Sull'evoluzione del progetto cfr. (Rolleri 2024).

⁴ Norme catalografiche per SBN: https://norme.iccu.sbn.it/index.php?title=Normative_catalografiche. La piattaforma, utilizzata per la prima volta nel 2016, ospita anche il Codice nazionale REICAT, e dal dicembre 2024 la guida a Manus Online.

⁵ Alfabetica: <https://alfabetica.it/web/alfabetica/>.

⁶ L'algoritmo per l'acquisizione delle informazioni si attiva quando l'utente avvia una ricerca sul Protagonista e procede seguendo una gerarchia di controlli. Per prima cosa verifica la presenza della proprietà P396 (identificativo SBN): se disponibile, la ricerca si conclude qui. Se nel record di authority SBN è presente anche il codice ISNI, viene effettuata una query sulla proprietà P213 (ISNI) in Wikidata e, se trovata, il processo termina. In assenza sia del VID sia dell'ISNI in Wikidata, l'algoritmo effettua la ricerca basandosi sulla somiglianza della stringa del Nome, applicando un eventuale filtro per data quando disponibile. In quest'ultimo caso, in presenza di omonimie, può presentarsi la possibilità di errori di associazione in presenza di omonimie.

⁷ Si ringraziano nello specifico Stefano Bargioni e Camillo Carlo Pellizzari di San Girolamo per le liste di segnalazioni inviate nel corso degli ultimi due anni all'Ufficio Authority dell'ICCU.

Oggetto del presente intervento è la descrizione del progetto che consiste nell'utilizzo dei dati di Wikidata per arricchire i nomi di persona che nel Catalogo di SBN risultino privi di informazioni identificative fondamentali, come la Nota informativa, le Datazioni e i codici Lingua e Paese.

Authority File di SBN

Fra le attività qualificanti del Catalogo di SBN, una delle più rilevanti riguarda la gestione e il monitoraggio dell'Authority file. Attualmente, in SBN vengono gestite voci d'autorità relative a nomi di persona, nomi di ente, titoli dell'opera, titoli dell'opera musicale, soggetti e classi. Inoltre, per le registrazioni bibliografiche anteriori al 1831, sono previsti ulteriori punti di accesso controllati: luogo di pubblicazione, editori e tipografi, e marche tipografiche. Questa attività è coordinata dall'Ufficio Authority dell'ICCU, il cui obiettivo principale è quello di garantire la qualità e il controllo dei punti di accesso ai record bibliografici presenti nel Catalogo nazionale. A tal fine, l'ICCU ha definito specifiche normative per la registrazione delle voci di autorità,⁸ in accordo con gli standard internazionali, e ha promosso percorsi formativi per i catalogatori della rete nazionale,⁹ anche in collaborazione con la Fondazione Scuola nazionale del patrimonio e delle attività culturali, nell'ambito del progetto Dicolab.¹⁰

A supporto del lavoro di authority, è stato costituito un *Gruppo di lavoro per la gestione e la manutenzione dell'Authority file di SBN. Nomi*¹¹ composto da esperti bibliotecari dei Poli di SBN che, per competenza, dislocazione territoriale, specificità di funzioni ed utenza, assicurassero al contesto cooperativo la professionalità e l'impegno indispensabili al compito delicato ed oneroso loro assegnato. Più recentemente è stato costituito un *Comitato di coordinamento per la gestione e la manutenzione dell'Authority file di SBN*¹² con il compito di fornire indirizzo e raccordo ai diversi Gruppi di lavoro operanti in attività di authority in SBN.

Lo sforzo impiegato dall'ICCU e dall'intera comunità di SBN per garantire la correttezza e l'uniformità formale dei punti di accesso ha dovuto misurarsi, nel tempo, con la crescita esponenziale dei dati del catalogo: i soli nomi di persona superano ormai i 3,7 milioni di record. Parallelamente, la progressiva riduzione delle risorse umane ed economiche destinate alla catalogazione, rende sempre più difficile gestire l'Authority file di SBN affidandosi esclusivamente al lavoro umano. Per questo motivo, nei prossimi anni, sarà necessario potenziare gli strumenti di monitoraggio del

⁸ Per le normative per la registrazione di authority in SBN cfr. https://norme.iccu.sbn.it/index.php?title=Norme_comuni/Authority_file. Per i soggetti e le classi, cfr. https://norme.iccu.sbn.it/index.php?title=Norme_comuni/Linee_guida_sul-%27indicizzazione.

⁹ L'ultimo percorso formativo realizzato dall'ICCU si è svolto nel maggio 2025. Per il programma cfr. <https://www.iccu.sbn.it/it/eventi-novita/novita/Corsi-di-formazione-e-aggiornamento-professionale-2025/>.

¹⁰ Il progetto Dicolab si colloca all'interno del Piano Nazionale per la Digitalizzazione del patrimonio culturale e ha come obiettivo quello di aggiornare e migliorare le competenze digitali dei professionisti della cultura. «Catalogare in SBN» è uno dei percorsi realizzati in collaborazione con ICCU, cfr. Catalogare in SBN: un nuovo percorso Dicolab dedicato ai processi di catalogazione digitale - Dicolab.

¹¹ Il gruppo di lavoro è stato più volte aggiornato nel corso degli anni, l'ultimo aggiornamento risale al 4 dicembre 2024. <https://www.iccu.sbn.it/it/attivita-servizi/gruppi-di-lavoro-e-commissioni/gruppo-di-lavoro-per-la-gestione-e-la-manutenzione-dellauthority-file-di-sbn/index.html>.

¹² <https://www.iccu.sbn.it/it/attivita-servizi/gruppi-di-lavoro-e-commissioni/comitato-di-coordinamento-per-la-gestione-e-manutenzione-dellauthority-file-di-sbn/>.

Catalogo e realizzare interventi correttivi massivi e centralizzati. In questa direzione, sarà fondamentale anche l'adozione di procedure automatiche per la validazione e l'arricchimento dei dati, basate sul confronto con altre banche dati, sia nazionali che internazionali. L'obiettivo è rafforzare l'interoperabilità del Catalogo nazionale, promuovendo lo scambio e il riutilizzo dei dati, così da migliorare la qualità dell'Authority file, sia in termini di completezza che di affidabilità dei record. Del resto, l'idea stessa di Controllo Bibliografico Universale si basa sulla possibilità di condividere i dati bibliografici e le voci di autorità prodotti dalle agenzie bibliografiche nazionali (Bianchini, Bargioni, e Pellizzari di San Girolamo 2022, 248).¹³ In quest'ottica, tra gli sviluppi previsti con il progetto Indice 3 (Atturo e Ravelli 2024, 56), si prevede un uso più esteso del campo 017 di UNIMARC, utile per collegare i record SBN con identificatori esterni, diversi da ISNI, come ad esempio VIAF e Wikidata. Questi collegamenti potranno contribuire in modo significativo al controllo e al miglioramento della qualità dei dati di authority presenti nel nostro Catalogo.

La prima fase del progetto

L'arricchimento massivo dell'Authority file di SBN si basa sul riutilizzo dei dati di un progetto, come quello di Wikidata, basato sull'archiviazione di una grande quantità di dati strutturati. Alla base di questo progetto c'è un'attività, ormai consolidata nel tempo che consiste nel collegamento delle entità di Wikidata con le voci dell'Authority file presenti nell'OPAC di SBN, mediante la registrazione nella scheda Wiki della proprietà P396, ovvero dell'identificativo del corrispondente record di SBN.¹⁴ A tale proposito, va ricordato che, fino a giugno del 2023, nell'OPAC di SBN venivano pubblicate solo le voci con un livello di autorità compreso tra 90 e 97, cioè meno del 10% del totale delle voci registrate nell'Indice SBN.¹⁵ Nel giugno del 2023, l'ICCU decideva di includere anche le voci con livello di autorità 51 e 71, che fino a quel momento non erano state esportate nell'OPAC, nella consapevolezza che in quella porzione di record fossero concentrate le maggiori criticità dell'Authority file di SBN: voci duplicate, voci non controllate e prive di forme di rinvio, voci per le quali spesso non è stata verificata la coerenza dei titoli collegati. Inoltre, la maggior parte dei record con livello 51 e 71 contiene informazioni identificative scarse o addirittura assenti. Tuttavia, in un'ottica di maggiore trasparenza e apertura dei dati delle proprie piattaforme, l'ICCU ha scelto di rendere pubblici anche questi dati, che costituiscono la parte numericamente più consistente. Di conseguenza, nell'OPAC di SBN sono ora presenti tutti i record relativi a nomi di persona e di ente, con esclusione di quelli con livello di autorità 05.¹⁶ L'auspicio è che questa maggiore apertura dei dati favorisca una migliore interoperabilità con altre piattaforme, anche non necessariamente bibliografiche, che gestiscono le medesime entità.

¹³ Sulla condivisione e il riuso dei dati bibliografici cfr. anche (Guerrini 2020, 38-42); sull'importanza della combinazione di processi manuali e automatici di validazione e arricchimento dei dati, cfr. anche (Casalini 2022).

¹⁴ Per una storia della proprietà 396 di Wikidata, cfr. (Pellizzari di San Girolamo 2024).

¹⁵ Si trattava delle stesse voci che l'ICCU invia anche a VIAF. A partire dal 2009 l'ICCU ha inviato annualmente al VIAF le voci di autorità con un alto grado di completezza ed esaustività, corrispondenti al livello 97 (Authority) e 95 (Super). Dal 2023, nell'ottica di ampliare la visibilità dei record di SBN, ha iniziato a inviare anche le voci con livello 90 (Massimo).

¹⁶ "Il livello 05 è predisposto per le retroconversioni da cataloghi o da repertori, i dati previsti sono quelli del trattamento del nome a livello minimo (51), se presenti e desumibili dalla scheda o dal repertorio di riferimento, in particolare i dati da inserire, se possibile, sono quelli relativi ai codici di qualificazione e i dati necessari alla disambiguazione del nome", cfr. (Atturo e Ravelli 2024, 48-9).

Il progetto di arricchimento ha avuto avvio nel maggio 2023, quando gli identificativi di SBN collegati a entità di Wikidata erano 103.148. Da questo elenco sono stati selezionati solamente i record relativi ai nomi di persona¹⁷ che, nell'Indice SBN, risultavano privi di Nota informativa, Datazioni, codice Lingua e codice Paese. In questa prima fase l'attenzione è stata rivolta solo ai record che non contenevano alcuna informazione, se non i campi ISNI o Fonte. I record con queste caratteristiche erano 17.145. Per queste voci sono state selezionate alcune proprietà di Wikidata destinate a completare, nel record di Indice, i campi Datazioni, Nota informativa e, quando assente, il campo ISNI. In particolare, le proprietà selezionate sono state le seguenti: P19 (Luogo di nascita), P20 (Luogo di Morte), P569 (Data di nascita), P570 (Data di morte), P106 (Occupazione), P611 (Ordine religioso), P123 (ISNI). Un requisito fondamentale è stato che ciascuna proprietà fosse supportata da almeno una fonte di riferimento.

Con l'esperienza e la consapevolezza che le operazioni massive sulle basi dati comportano necessariamente un margine di errore, è stata condotta una verifica estremamente approfondita sulle proprietà estratte, al fine di individuare tutte le possibili tipologie di errore e valutarne l'impatto sul Catalogo. La situazione più critica si è verificata in caso di errata associazione fra entità, errore riscontrato in un numero residuale di casi. Più frequente è stata invece l'associazione in Wikidata con identificativi che in SBN potevano fare riferimento a entità duplicate. Tuttavia, anche questo errore non era tale da inficiare il buon risultato del progetto. In ogni caso, tutti gli errori rilevati durante il test sono stati immediatamente corretti: in particolare, si è intervenuti sulle entità duplicate in SBN, attività che ha consentito di estendere le correzioni anche a Wikidata. Inoltre, sono stati necessari alcuni interventi per adattare le informazioni provenienti da Wikidata ai formalismi richiesti per la registrazione nei campi dei record di SBN.

Le principali problematiche che hanno richiesto un intervento di ICCU sono illustrate nei punti seguenti.

- **Luogo di nascita e il luogo di morte**

Poiché i luoghi di nascita e morte registrati in Wikidata possono indicare siti specifici, come un ospedale o un castello, e non necessariamente un'unità amministrativa, è stata necessaria una revisione degli elenchi dei luoghi. Questo ha permesso di inserire nella Nota informativa di SBN l'entità amministrativa corrispondente alla nascita e/o alla morte, generalmente il comune o, in mancanza di questo, la regione o il Paese.

- **Occupazione**

Poiché in Wikidata possono essere registrate più occupazioni, non tutte sempre accompagnate da una fonte, capita talvolta che l'unica occupazione referenziata sia quella meno qualificante o comunque insufficiente a identificare con precisione la persona. Ad esempio, molti nomi risultavano qualificati solo con l'occupazione "politico", un'indicazione poco significativa per persone che hanno ricoperto ruoli pubblici di rilievo, come presidente del consiglio, presidente della repubblica, senatore, e simili. Tale informazione è registrata nella proprietà P39 (Carica ricoperta), che in una prima fase non era stata considerata.

¹⁷ Sono stati quindi esclusi i nomi di ente, compresi quelli di tipografi, editori, librai, identificati tramite la forma inversa del nome (cognome, nome), poiché in SBN il record relativo all'autore è distinto da quello riferito al tipografo, anche quando si tratta della stessa entità.

- **Ordine religioso**

Analogamente al caso dell'occupazione, anche l'indicazione dell'ordine religioso risultava spesso insufficiente a identificare correttamente una persona che, nel corso della sua vita, aveva ricoperto incarichi di rilievo, come quello cardinalizio o papale. Tali informazioni sono registrate nella proprietà P39 (Carica ricoperta). Inoltre, nel corso di questa prima fase di verifica, è emerso che per i religiosi è particolarmente importante estrarre anche la proprietà P411 (Stato di canonizzazione), inizialmente non presa in considerazione.

- **Data**

In Wikidata, le date sono espresse secondo lo standard ISO 8601 che definisce i formati per la rappresentazione di date e orari. Oltre alla stringa che indica anno-mese-giorno, viene associato un valore che ne specifica il grado di precisione: ad esempio il valore 11 indica che sono noti giorno, mese e anno; il valore 10 indica che sono noti solo il mese e l'anno; il valore 9 che è noto solo l'anno; il valore 8 che è noto il decennio e così via.¹⁸ Un ulteriore indicatore segnala se la data è incerta ("ca."= circa), o solo ipotizzata ("?"). L'informazione relativa alla data andava riportata sia nel campo Datazioni, sia nel campo Nota informativa di SBN: nel campo Datazioni la data è stata trascritta utilizzando solamente caratteri numerici, secondo la normativa, recentemente aggiornata, per la registrazione delle Datazioni nei record di authority di SBN.¹⁹ Nella Nota informativa, invece, i dati sono stati rielaborati e trascritti in forma discorsiva, adottando locuzioni diverse a seconda delle differenti modalità in cui la data si poteva presentare, ad esempio: "nato il 10 novembre 1824, nato nell'ottobre del 1994, vissuto nel XVI secolo", etc.

- **ISNI**

A eccezione dei casi di associazione errata di codici ISNI in Wikidata, difficili da individuare, la principale criticità riscontrata riguardava la presenza di entità duplicate in ISNI. Tuttavia, grazie al controllo presente in Interfaccia Diretta, che impedisce l'inserimento di uno stesso ISNI in più record, durante la fase di importazione degli ISNI da Wikidata sono stati rilevati diversi casi di duplicazione, che sono stati tempestivamente corretti.

Al termine della prima fase di verifica delle informazioni estratte da Wikidata, è emerso che, oltre alla Carica ricoperta e allo Stato di beatificazione, utili a completare le informazioni sulle entità presenti nel Catalogo, erano importanti anche altre tre proprietà inizialmente non prese in considerazione: P21 (Sesso o genere), P27 (Paese di cittadinanza), P6886 (Lingua di scrittura).

¹⁸ Sulla modalità di registrazione delle date in Wikidata vedi <https://www.wikidata.org/wiki/Help:Dates/it>.

¹⁹ https://norme.iccu.sbn.it/index.php?title=Norme_comuni/Authority_file/Nomi_di_persona/Registrazione_di_autho-riety/Datazioni. La normativa per la registrazione delle datazioni delle voci di authority in SBN prevede l'utilizzo di soli caratteri numerici nella forma: anno.mese.giorno, quando tutti e tre conosciuti, altrimenti solo l'anno o il decennio o il secolo, secondo diverse casistiche. Una modalità che va incontro a un progressivo adeguamento agli standard internazionali, cfr. (Atturo e Ravelli 2024, 51-2).

- **Sesso o genere**

Pur non essendo previsto un codice di genere in SBN, questa informazione risulta fondamentale per poter declinare correttamente al maschile o al femminile le occupazioni e le espressioni “nato/nata” e “morto/morta” all’interno della Nota informativa.

- **Paese di cittadinanza**

Poiché il record SBN consente l’associazione di un solo codice Paese, si è deciso di utilizzare la proprietà di Wikidata solo nei casi in cui fosse presente una sola occorrenza, in assenza di un principio di gerarchia che definisse un paese principale di cittadinanza.

- **Lingua di scrittura**

Secondo le norme SBN per i Nomi di persona, la lingua da indicare è quella utilizzata dall’autore nelle sue pubblicazioni. Si tratta tuttavia di una proprietà non molto rappresentata in Wikidata. Come per il codice Paese, anche questa informazione è stata acquisita solo quando risultava una singola occorrenza.

Una volta ricomposte in un unico file tutte le proprietà estratte separatamente da Wikidata, è stato necessario concatenare le informazioni per elaborare le stringhe di carattere discorsivo da inserire nella Nota informativa del record di SBN. Le espressioni sono state adattate e declinate in modo diverso a seconda delle tipologie di dati da ricomporre.

Di seguito alcuni esempi di formulazioni da inserire nel campo Nota informativa:

- Agronomo, giornalista. Nato a Cerro Maggiore, 31.12.1878, morto a Roma, 18.04.1960,
- Omeopata. Nato nel 1840, morto nel 1901,
- Archivista, filosofo, scrittore. Nato in Cina, presumibilmente intorno al 571 a.C., morto presumibilmente intorno al 490 a.C.

Al termine del lavoro, è emerso che per circa 2.000 autori mancavano informazioni relative all’attività svolta, alla carica ricoperta o all’eventuale stato di beatificazione. Nella maggior parte dei casi, tali dati potevano essere ricavati dalla descrizione presente nella voce di Wikidata, di seguito all’etichetta con il nome. Questa descrizione viene generata automaticamente sulla base di informazioni non sempre contenute nei campi formalizzati o, se presenti, non sempre accompagnate da fonti. Un’indagine a campione ha confermato che queste descrizioni risultavano, nel complesso, attendibili. Si è quindi proceduto a estrarle ed eventualmente a tradurle in italiano nei casi in cui comparissero in lingua inglese, tedesca o francese.

Nel luglio 2024, completati tutti i test e apportate le necessarie correzioni e integrazioni, sono stati inseriti nell’Indice SBN 16.102 record arricchiti con nuove informazioni identificanti. Tali record sono riconoscibili dalla Nota del catalogatore, che riporta un’espressione analoga alla seguente: “Informazioni da Wikidata 01.07.2023–31.03.2024 [wikidata.org/wiki/Q94409375](https://www.wikidata.org/wiki/Q94409375)”.

La prosecuzione del progetto

Nell'autunno del 2024 è stata effettuata una nuova estrazione di dati da Wikidata. Rispetto alla precedente selezione, le entità con proprietà P396 (Identificativo SBN dell'autore) risultavano pari a 144.347,²⁰ inclusi i 16.102 che erano stati già arricchiti con informazioni provenienti da Wikidata. Anche in questa fase, l'obiettivo era individuare i record di SBN privi di informazioni identificanti. L'attenzione si è concentrata non solo sui record con Nota informativa vuota, ma anche su quelli in cui la Nota risultava compilata con contenuti non pertinenti, quali la data, le fonti, o altre informazioni di servizio.²¹ Per questi casi si è proceduto a una bonifica dei record, riportando le informazioni nei campi corretti oppure eliminando le informazioni non appropriate. Di conseguenza, il numero di record presenti in Indice privi di Nota informativa è cresciuto in modo significativo, raggiungendo circa 36.000 unità.

A differenza della prima fase, si è ritenuto opportuno intervenire anche sui record di SBN che presentavano alcuni campi già compilati, come Datazione, Lingua o Paese. Si è infatti ritenuto che l'aggiunta di elementi come il luogo di nascita o di morte, l'occupazione, le cariche ricoperte o l'eventuale stato di canonizzazione, potesse contribuire a migliorare sensibilmente la qualità complessiva del record.

In questa seconda fase, è stata sviluppata una procedura specifica per l'estrazione dei dati di Wikidata, progettata per recepire tutte le esigenze emerse durante la fase iniziale del progetto. Le specifiche della procedura sono illustrate nelle 'Tabelle per estrazione dati per ICCU' in Wikidata.²² Il progetto è tuttora in corso e procede attraverso successivi affinamenti ma ha già superato la fase sperimentale iniziale, restituendo risultati decisamente incoraggianti. La procedura messa a punto ha permesso non solo di arricchire oltre 51.000 record di SBN precedentemente privi di informazioni identificanti, ma anche di individuare e correggere molteplici criticità: voci duplicate, altre che necessitavano di essere disambiguate, note informative improprie, datazioni non aggiornate e altre incongruenze presenti nell'Indice di SBN. Il miglioramento della qualità dei record non ha solo un impatto sull'affidabilità complessiva del catalogo, ma apporta benefici su più livelli. Per gli utenti, significa disporre di informazioni più chiare, dettagliate e coerenti; per i catalogatori, rappresenta uno strumento prezioso per facilitare le attività di *authority control*, agevolando l'univocità delle intestazioni e la coerenza nell'organizzazione dei dati bibliografici.

Guardando al futuro, l'auspicio è quello di ampliare ulteriormente il dialogo e le possibilità di interoperabilità con altre banche dati, sia nazionali che internazionali. Questo approccio mira a rafforzare l'integrazione tra piattaforme di dati diverse e a promuovere un ecosistema informativo più aperto, dinamico e collaborativo, in grado di valorizzare al meglio l'Authority file di SBN e, più in generale, il patrimonio documentario e culturale a cui esso dà accesso.

²⁰ Al momento dell'elaborazione del presente contributo (luglio 2025), gli elementi in Wikidata con proprietà P396 hanno superato le 190.000 unità.

²¹ In molti casi queste informazioni derivano da lavorazioni o importazioni massive precedenti a Indice 2 quando non esisteva il campo *Nota informativa* ma solo una generica *Nota* che poteva essere compilato con contenuti vari.

²² https://www.wikidata.org/w/index.php?title=Property_talk:P396&oldid=2341991629#Tabelle_per_estrazione_dati_per_ICCU.

Riferimenti bibliografici*

Atturo, Valentina, e Elena Ravelli. 2024. «The evolution of authority work in SBN. From origins to Alphabetica and future prospects.» *JLIS.it* 15 (1): 45-58. <https://doi.org/10.36253/jlis.it-572>.

Bianchini, Carlo, Stefano Bargioni, Camillo Carlo Pellizzari di San Girolamo. 2022. «Le voci di autorità dei nomi di persona in SBN e Alphabetica: problemi e prospettive.» *Bibliothecae.it* 11 (1): 247-314. <https://doi.org/10.6092/ISSN.2283-9364/15078>.

Casalini, Michele. 2022 «The future of bibliographic services in light of new concepts of authority control.» *JLIS.it* 13 (1):107-15. <https://doi.org/10.4403/jlis.it-12766>.

Guerrini, Mauro. 2020 «Dalla catalogazione alla metadattazione. Tracce di un percorso.» Roma: Associazione italiana biblioteche.

ICCU (Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche), e Wikimedia Italia. 2015 *Convenzione tra ICCU e Associazione Wikimedia Italia - As-sociazione per la diffusione della conoscenza libera*. https://www.iccu.sbn.it/export/sites/iccu/documenti/2016/convenzione_Wikimedia.pdf.

ICCU (Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche), e Wikimedia Italia. 2018 *Convenzione tra ICCU e Associazione Wikimedia Italia - As-sociazione per la diffusione della conoscenza libera*. <https://www.iccu.sbn.it/export/sites/iccu/documenti/2020/ICCU-Wikimedia.pdf>.

ICCU (Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche), e Wikimedia Italia. 2022 *Convenzione tra ICCU e Associazione Wikimedia Italia - As-sociazione per la diffusione della conoscenza libera*. https://www.iccu.sbn.it/export/sites/iccu/documenti/2022/ICCU_Wikimedia_2022.pdf.

Pellizzari di San Girolamo, Camillo Carlo. 2024. «The reconciliation of SBN Authority records with Wikidata. Progresses and perspectives after a decade of work (2013-2023).» *JLIS.it* 15 (1):v33-44. <https://doi.org/10.36253/jlis.it-573>.

Rolleri, Davide. 2024. «Importazione dell'Anagrafe delle Biblioteche Italiane in Wikidata.» (slide e testo dell'intervento ai Wikidata Days, Bologna, novembre 8-9, 2024).

* Ultima data di consultazione dei siti web: 11 novembre 2025.