

Wikidata and SBN: An Assessment of Two Years of Work (2023–2025)*

Carlo Bianchini^(a), Stefano Bargioni^(b),
Camillo Carlo Pellizzari di San Girolamo^(c)

a) University of Pavia, <https://orcid.org/0000-0002-6635-6371>
b) Pontifical University Santa Croce, <https://orcid.org/0000-0001-7679-2874>
c) Scuola Normale Superiore, <https://orcid.org/0000-0003-2699-1693>

Contact: Carlo Bianchini, carlo.bianchini@unipv.it; Stefano Bargioni, bargioni@pusc.it;
Camillo Carlo Pellizzari di San Girolamo, camillo.pellizzaridisangirolamo@sns.it
Received: 17 July 2025; **Accepted:** 24 October 2025; **First Published:** 15 January 2026

ABSTRACT

The article presents the theoretical foundations and the reconstruction of the intensive bibliographic control, correction, and reconciliation activities that have progressively strengthened the link between SBN authority records and Wikidata entities, particularly over the past two years. It provides a detailed account and analysis of the most recent interventions carried out through the collaboration between ICCU and the Wikidata community. Building on this experience, the paper proposes a set of key strategies to further enhance the representation of Italian cultural entities within the Semantic Web and the network of Linked Data. Two main directions are suggested to pursue this goal: developing reliable pathways of access to and from SBN in order to foster interoperability among information systems, and promoting awareness of best practices in metadata creation within ordinary cataloguing workflows.

KEYWORDS

Wikidata; SBN; Authority control; Authority data; Reconciliation.

Wikidata e SBN: un bilancio di due anni di lavoro (2023-2025)

ABSTRACT

L'articolo presenta i presupposti teorici e la ricostruzione dell'intensa attività di controllo bibliografico, correzione e riconciliazione, che hanno consentito che il legame tra le Voci di autorità di SBN e gli elementi di Wikidata si consolidasse progressivamente, in particolare nel corso degli ultimi due anni. L'articolo illustra e analizza in dettaglio gli interventi più recenti, effettuati grazie alla collaborazione tra l'ICCU e la comunità di Wikidata. Sulla base dell'esperienza maturata, vengono proposte alcune strategie fondamentali per rafforzare ulteriormente la presenza delle entità culturali italiane nel web semantico e nella rete dei dati collegati ("Linked Data"). Si suggerisce di perseguire questo obiettivo lungo due direttrici principali: offrire percorsi di accesso affidabili da e verso SBN, favorendo l'interoperabilità tra sistemi informativi, e promuovere la conoscenza delle migliori pratiche di metadattazione all'interno delle pratiche di catalogazione ordinaria.

PAROLE CHIAVE

Wikidata; SBN; Controllo di Autorità; Voci di autorità; Riconciliazione.

* Il testo è stato scritto in totale collaborazione e condivisione: tuttavia vanno ascritti a Stefano Bargioni il paragrafo 4, a Carlo Bianchini i paragrafi 1 e 2 e a Camillo Pellizzari i paragrafi 3 e 5.

1. Introduzione

La profonda trasformazione della catalogazione – tra le attività tecniche più centrali del flusso delle informazioni bibliografiche all'interno delle biblioteche – verso la *metadattazione*, ovvero la 'catalogazione in era digitale' (Guerrini 2020; 2022; Bargioni et al. 2022) si basa, tra l'altro, sullo spostamento del focus dalle risorse bibliografiche, gli oggetti portatori di informazione (es. libri, periodici, articoli, opere di consultazione ecc.) risultanti da un processo di produzione editoriale, alle relazioni bibliografiche che le caratterizzano e che testimoniano quel processo, come quella di autorialità, di curatela, di traduzione ecc. Questo spostamento, che potrebbe sembrare solo di facciata, si fonda su una visione teorica espressa in un modello concettuale evoluto: IFLA LRM (Library Reference Model) (IFLA 2016; 2017; Bianchini 2017; Guerrini e Sardo 2018). In esso, come nei suoi precursori (IFLA Study Group on the Functional Requirements for Bibliographic Records 2009; IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR) 2009; IFLA Working Group on Functional Requirements for Subject Authority Records (FRSAR) 2010), si individuano le entità di interesse e se ne esplicano tanto gli attributi (o proprietà) quanto le relazioni. La riduzione del numero di attributi e l'aumento delle relazioni tra entità sono un aspetto centrale di questo modello e anche degli standard descrittivi internazionali che vi si rifanno, come RDA (JSC 2010; Bianchini e Guerrini 2014). Poiché per creare una relazione in modo efficace ed efficiente è indispensabile definirne con esattezza assoluta gli estremi, ovvero il soggetto e l'oggetto, diventano centrali la questione dell'identificazione delle entità coinvolte e il tema – vasto e complesso – dell' 'identità' (ovvero di stabilire principi per i quali un'entità non è più la stessa ma diventa 'altra'). Su questo secondo problema, la riflessione è ancora tutta agli inizi, come osserva Tiziana Possemato: "i concetti di 'entità' e di 'identità' in ambito catalografico [...] non sono mai stati analizzati in relazione all'attività di identificazione e modellazione degli oggetti del dominio GLAM" (Possemato 2024, 29). Tuttavia, la metadattazione e l'approccio *entity modelling* richiedono un cambiamento di prospettiva nella registrazione dei dati caratterizzato dall'identificazione delle entità stesse, attraverso gli identificatori, strumenti tipici del web semantico (Manzoni 2022). Proprio rispetto all'importanza e al ruolo degli identificatori del web semantico, Wikidata (d'ora in poi WD) svolge una funzione fondamentale anche nell'ambito dei cataloghi.

2. Wikidata e il controllo di autorità

Identificare le entità è un compito fondamentale per creare le condizioni per raggiungere l'obiettivo del Controllo Bibliografico Universale (CBU), il progetto più ambizioso mai ideato in ambito biblioteconomico (Anderson 1974; 2000; Solimine 1995). Il CBU è stato la ragione dello sviluppo di tutti gli standard internazionali nell'ambito della catalogazione, dall'ISBD agli ICP, da FRBR a IFLA LRM. Tra i primi strumenti creati dalla comunità internazionale per condividere gli sforzi delle biblioteche mondiali nell'ambito del lavoro di autorità (Guerrini e Sardo 2003), c'è il Virtual International Authority File,¹ o VIAF (Tillett, Reyad, e Cristán 2006). Si tratta di un progetto basato sullo sforzo congiunto di molte biblioteche e agenzie bibliografi-

¹ <https://viaf.org/>.

che nazionali – come la Library of Congress, la Deutsche Nationalbibliothek, la Bibliothèque nationale de France e, per l'Italia, l'ICCU con i dati di SBN – e realizzato da OCLC. Il VIAF è costituito da insiemi di forme di nomi – detti cluster – che raggruppano e collegano tra loro i punti di accesso autorizzati creati dalle agenzie nazionali in base alle regole in vigore nei diversi paesi. La creazione dei cluster si basa su un software che, tramite algoritmi, individua l'eventuale corrispondenza tra una voce prodotta da un'agenzia e un cluster già presente in VIAF, collegandola o creando per la voce un nuovo cluster. Questo processo complesso può generare molti errori (tipicamente, l'attribuzione di una forma del nome di una persona a un cluster che non la rappresenta o, al contrario, la sua mancata attribuzione al cluster corretto e la conseguente creazione di un cluster distinto). Il lavoro di controllo della qualità dei dati ottenuti con questo processo di clusterizzazione è particolarmente lungo e costoso: la tecnica più efficace ed economica si basa sulla comparazione tra i dati registrati nei cluster VIAF e altri dati di autorità che offrono altrettante garanzie di qualità. Il processo di clusterizzazione infatti prevede sempre un margine di errore, ma l'errore non si ripete sempre identico nei diversi dataset ed è questo che rende efficace il processo di confronto.

Dove trovare un altro insieme di dati di autorità che contiene nomi di persona di provenienza internazionale e intercontinentale? Poiché il VIAF è stato il primo progetto di controllo di autorità di dimensioni internazionali – che ha preceduto nel tempo anche l'International Standard Name Identifier² – la domanda non è puramente retorica. È indispensabile tentare un approccio che coinvolga il web semantico e in particolare un suo 'hub' fondamentale: Wikidata (WD). Anche se il VIAF e WD sono piuttosto diversi per obiettivi, ambito, approccio organizzativo e teorico, raccolta e gestione dei dati, entrambi svolgono un importante ruolo di reciproco controllo della qualità dei dati (Bianchini, Bargioni, e Pellizzari di San Girolamo 2021).

WD ha un background molto diverso da quello di altre istituzioni che si sono occupate del CBU dagli anni 70 del XX secolo. WD non si dedica infatti solo ai dati bibliografici o di autorità, malgrado il suo valore da questo punto di vista sia indiscutibile (Bianchini e Sardo 2022). Proprio per il suo approccio diverso, l'esistenza, l'uso e il riuso dei dati di WD hanno un impatto pratico e teorico molto forte sul modello che la comunità bibliotecaria ha del CBU. WD esemplifica inoltre un approccio differente all'identificazione e alla descrizione delle entità dell'universo bibliografico; infatti, in WD, la differenza teorica tra il punto di accesso autorizzato (perché preferito) e i punti di accesso varianti – che era già stata molto ridimensionata nel cluster del VIAF – sembra diventare sempre meno importante dal punto di vista pratico. Infine, con WD è stata stimolata anche la riflessione sulla descrizione bibliografica, perché l'approccio granulare nella registrazione dei dati (per triple o dichiarazioni) è piuttosto diverso e offre potenzialità che la descrizione tradizionale non consente.

Lo stretto legame tra WD e il mondo dei dati bibliografici si è manifestato in molti modi, ed esiste un'ampia letteratura in merito,³ ma la sua filosofia aperta e l'adesione piena ai principi FAIR⁴

² <https://isni.org/page/what-is-isni/>: l'ISNI è uno standard ISO (27729:2025), utilizzato da numerose biblioteche, editori, database, e organizzazioni per la gestione dei diritti d'autore in tutto il mondo.

³ Si rinvia alla bibliografia di WD su WD (<https://w.wiki/EfQf>) e alla bibliografia su Zotero del Gruppo DARIAH su Wikibase, WD e Digital Humanities (https://www.zotero.org/groups/5639268/dariah_dhwiki/library).

⁴ <https://www.go-fair.org/fair-principles/>.

favorisce molto la sperimentazione di soluzioni integrate con altri prodotti aperti. È il caso dell'estensione AuthorityBox per l'OPAC della Biblioteca della PUSC (Bargioni 2020), che si basa sul software open source Koha, ma anche dei sempre più frequenti casi di riuso dei dati, a partire da *Alphabetica* fino al progetto Parsifal di URBE (Pellizzari di San Girolamo 2024a).

Infine, la caratteristica di WD di identificare non solo le persone, ma qualsiasi oggetto o concetto del mondo reale (esseri viventi, luoghi, proprietà come colore, dimensioni e peso, concetti come amicizia, inflazione ed elettroforesi) la rende uno strumento utile per un confronto anche di strumenti di indicizzazione specializzati e generalisti, come una classificazione o il Nuovo Soggettario (Cencetti, Pellizzari di San Girolamo, e Viti 2025).

3. Wikidata e SBN. Una collaborazione consolidata

Come molti altri cataloghi nazionali, anche l'OPAC SBN si è avvalso delle potenzialità offerte dal confronto con WD per l'analisi della qualità dei suoi dati e l'individuazione di possibili strategie di miglioramento.

Un primo studio si era focalizzato sull'analisi della qualità e della quantità dei dati di autorità di SBN considerando due diversi set di dati e comparandoli a quelli disponibili su WD (Bianchini, Bargioni, e Pellizzari di San Girolamo 2022) al fine di stabilire se la qualità delle voci di autorità (d'ora in poi VID) dei vari livelli previsti da SBN – in particolare i livelli da 90 compreso in su – corrispondeva effettivamente ai requisiti di completezza e di qualità indicati da SBN stesso per quei livelli. Il caso di studio si basava su un principio fondamentale per la qualità dei dati, ovvero il confronto tra dataset diversi che descrivono, in tutto o in parte, lo stesso universo. Il confronto – o riconciliazione – tra dataset può avvenire tra SBN e WD, come nel caso di studio, ma anche con e tra qualsiasi altro dataset. Il risultato è garantito dal processo di riconciliazione, non solo dalla qualità dei dataset di partenza.

Dall'analisi della qualità dei VID e del potenziale offerto da WD come strumento di controllo per il miglioramento delle voci si è virtuosamente sviluppato un progetto di reciproca interconnessione tra WD e SBN.⁵ Anche se già dal 2013 WD aveva cominciato a creare legami verso i VID tramite la proprietà P396 che li identifica in WD (Pellizzari di San Girolamo 2024b, 34), è stato nel 2023 che si è avuta una significativa svolta nel processo di riconciliazione, grazie ad alcune innovazioni cruciali da parte di ICCU nella gestione dei VID: nel luglio 2023 i VID di livello 51 e 71 sono stati resi disponibili nell'OPAC, in aggiunta a quelli di livello 90 o superiore già disponibili, con reindirizzamenti per i VID fusi e implementazione di LOD e API (Atturo e Ravelli 2024).

Di conseguenza negli ultimi due anni (metà 2023-metà 2025), grazie al notevole incremento sia del numero di VID disponibili nell'OPAC e quindi riconciliabili con WD, sia dell'interesse da parte di ICCU e della comunità di WD per il processo di riconciliazione, le collaborazioni e i progetti riguardo ai VID sono molto aumentati, con tangibili miglioramenti in entrambi i dataset.

⁵ <https://w.wiki/EfQs> è la pagina WD che raccoglie tutti i materiali relativi al progetto, costantemente aggiornata.

3.1 L'impegno di ICCU con Wikidata per SBN

Le innovazioni da parte di ICCU nella gestione dei VID non hanno riguardato solo l'apertura dei dati:

- nel dicembre 2024 è stata rinnovata la composizione del *Gruppo di lavoro per la gestione e la manutenzione dell'Authority file di SBN. Nomi*,⁶ scopo di questo gruppo di lavoro è la revisione qualitativa dei VID e la realizzazione di progetti di bonifica mirati, nonché la discussione di *best practices* relative a tali bonifiche;
- nello stesso mese è stato anche nominato un Comitato di Coordinamento per la Gestione e manutenzione dell'Authority File di SBN;⁷
- nel marzo 2025 è stata pubblicata la normativa per il trattamento dei VID relativi agli enti.⁸

Il progetto più ampio portato avanti in questi due anni da ICCU in collaborazione con la comunità di WD, nell'ambito della consolidata collaborazione fra Wikimedia Italia e ICCU (istituzionalizzata tramite una serie di convenzioni firmate nel 2015,⁹ 2018¹⁰ e 2022,¹¹ quest'ultima in corso di validità), ha riguardato l'arricchimento di gruppi di VID con dati ricavati da WD: a partire dall'insieme dei VID registrati in WD, si considera un sottoinsieme di VID carenti di informazioni e in tali VID si inseriscono dati tratti da WD per rendere le persone rappresentate da tali VID univocamente identificabili. Alla fine della prima fase del progetto circa 16200 VID sono stati arricchiti nel luglio 2024.¹² Elena Ravelli, coordinatrice dell'Area di attività per l'elaborazione e diffusione degli standard e delle norme catalografiche, e per la didattica dell'ICCU, ha presentato il progetto durante i Wikidata Days Bologna 2024, un evento dedicato all'interazione tra WD e le biblioteche accademiche (Ravelli 2024). La seconda fase del progetto, che arricchirà un nuovo gruppo di VID, è in avanzata fase di svolgimento.

3.2 La comunità Wikidata per SBN

Anche sul versante della comunità WD, i progressi si sono dispiegati in varie forme.

Un aspetto importante ha riguardato il miglioramento degli strumenti a disposizione degli utenti per lavorare alla riconciliazione, sia manuale sia semiautomatica, dei VID con WD: lo sviluppo di strumenti come Wikilinker per SBN (settembre 2023), WikiPlayground (febbraio 2025) e VID-Q reconciliator (marzo-aprile 2025), di cui si tratta dettagliatamente nel par. 4, ha consentito di progredire molto nella riconciliazione.

Si sono inoltre svolti numerosi eventi, in presenza e online, per spiegare le modalità di riconciliazione tra SBN e WD e coinvolgere più volontari, soprattutto bibliotecari: si segnalano in particolare due workshop relativi a SBN e WD presso il Convegno delle Stelline (21 marzo 2024¹³ e 13

⁶ <https://tinyurl.com/2r5t9k3r>.

⁷ <https://tinyurl.com/2ah6a6he>.

⁸ <https://tinyurl.com/t89a4xtt>.

⁹ <https://tinyurl.com/c3pswe5f>.

¹⁰ <https://tinyurl.com/432ufewz>.

¹¹ <https://tinyurl.com/3nh2rctz>.

¹² <https://tinyurl.com/mua4hd6p>.

¹³ <https://w.wiki/EfQd>.

marzo 2025¹⁴) e un corso relativo a WD e il controllo di autorità presso il Servizio Bibliotecario di Ateneo dell'Università di Pisa (9-12 giugno 2025¹⁵); la riconciliazione tra SBN e WD è stata anche l'esercizio svolto da 14 studenti tra il novembre 2024 e l'aprile 2025 nell'ambito del corso di *Archivi digitali, digital curation and preservation* tenuto da Chiara Storti presso l'Università Telematica Internazionale UNINETTUNO, con un totale di 444 VID aggiunti a WD.¹⁶

Si è anche cercato di facilitare ai volontari di WD il processo di comunicazione di problemi nei VID: per questo nel marzo 2024 è stata creata in WD una pagina per segnalazioni di errore,¹⁷ che si affianca all'indirizzo mail reso disponibile da ICCU per le segnalazioni di errore (ic-cu.AFno-mi@cultura.gov.it) in modo da meglio distribuire il carico di segnalazioni.

Al tempo stesso, per facilitare ai bibliotecari che possono modificare i VID l'uso di WD per trovare in essi errori da correggere, sono state allestite delle query per verificare potenziali errori¹⁸ e sono state redatte delle istruzioni per individuare tramite WD e risolvere i casi di VID duplicati da fondere.¹⁹

Tra fine 2024 e inizio 2025, nell'ambito della sua tesi magistrale di informatica umanistica (De Monaco 2025), è stato svolto da Sara De Monaco un progetto di creazione e revisione in WD delle quasi 500 persone descritte in (Frattarolo 1975) e di sistemazione dei loro principali omonimi, aggiungendo in tutti gli elementi ove possibile un VID; inoltre, tutti i 253 errori incontrati in SBN (soprattutto record bibliografici male attribuiti, VID erroneamente privi di qualificazione e VID duplicati da unire) sono stati raccolti in una lista di segnalazioni (De Monaco e Pellizzari di San Girolamo 2025), poi verificate e risolte.

¹⁴ <https://w.wiki/EfQb>.

¹⁵ <https://w.wiki/EfQa>.

¹⁶ <https://w.wiki/EfQZ>.

¹⁷ <https://w.wiki/EfQX>.

¹⁸ <https://w.wiki/EfQU>.

¹⁹ <https://w.wiki/EfQV>.

3.3 Dati statistici

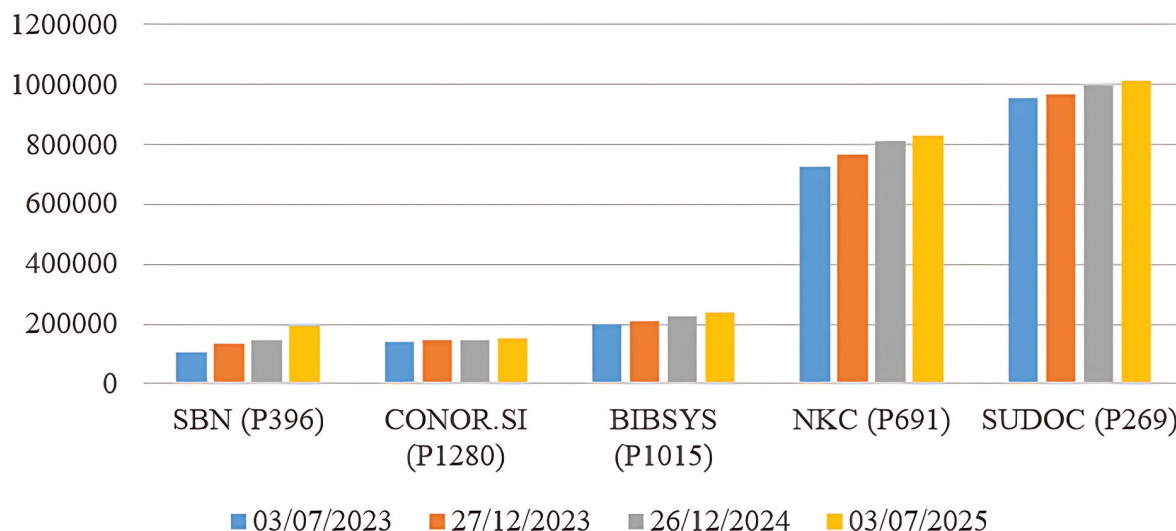


Tabella 1: numero di valori presenti in WD (come dichiarazioni) per cinque proprietà di *authority file* europei membri del VIAF²⁰

Dalla Tabella 1 si nota come in due anni i valori di P396 siano cresciuti molto di più (+80%) rispetto a quelli di P691 (+14,4%) e di P269 (+5,7%); una parte significativa di questa crescita, soprattutto nei primi mesi del 2025, è dovuta all'efficacia del metodo VID-Q che sarà descritto nel par. 4. La crescita nel numero di utenti coinvolti in WD nell'aggiunta di P396 è stata meno significativa: il numero di utenti che hanno aggiunto almeno 100 valori di P396 è cresciuto solo lievemente, da 51 a 60 (di cui, in entrambi i casi, 4 bot, cioè programmi automatici), tra il 1° aprile 2023²¹ e il 1° luglio 2025;²² al 1° luglio 2025 sono 1161 gli utenti totali che hanno aggiunto almeno un valore di P396 da quando la proprietà esiste (per confronto, sono 1420 per P691²³ e 3302 per P269²⁴). Se, quindi, il progresso negli strumenti a disposizione per l'abbinamento di P396 ha consentito di accelerare molto il ritmo di crescita dei valori di P396, sia grazie a metodi semiautomatici sia rendendo più efficiente il lavoro manuale degli utenti, è ancora necessario aumentare il numero di utenti coinvolti in questa attività.

4. Strumenti di riconciliazione e visualizzazione dati specifici per SBN

Il controllo di autorità svolto in SBN è un'attività complessa e articolata e saranno necessari tempi ancora lunghi per giungere alla pubblicazione completa dei dati di autorità prodotti in Italia, ma

²⁰ I dati sono tratti dalla cronologia di <https://w.wiki/EfQj>.

²¹ <https://tinyurl.com/4y4u7rw6>.

²² <https://tinyurl.com/2a35ysu8>.

²³ <https://tinyurl.com/4r5xntd8>.

²⁴ <https://tinyurl.com/4jzxm8d3>.

è indispensabile che il bene comune rappresentato dai VID sia reso disponibile al meglio, integrandosi con quanto viene realizzato all'estero e incrementando il loro riuso nel web semantico in forma di Linked Open Data (Bianchini, Bargioni, e Pellizzari di San Girolamo 2021; 2022).

La soluzione più praticabile è quella di pubblicare e diffondere i VID attraverso gli hub del web semantico specialistici come VIAF o generalisti come WD, secondo il quarto principio dei Linked Data (Berners-Lee 2006). Tuttavia, il processo è molto differente dal punto di vista della fattibilità. Nel caso del VIAF, solo ICCU è in grado di svolgere questo compito, in quanto tra le fonti primarie di VIAF,²⁵ attraverso l'invio dei dati con cadenza all'incirca annuale (nell'aprile 2024 VIAF conteneva 200041 VID).²⁶

Nel caso di WD, invece, la sua apertura alle modifiche e soprattutto l'uso della proprietà P396 per registrare i VID permettono a chiunque di contribuire all'identificazione e al collegamento dei VID con qualsiasi altra realtà produttrice di dati culturali che abbia identificatori della stessa entità.

In WD la proprietà P396 si può aggiungere con una modifica manuale,²⁷ ma, al momento esistono anche tre strumenti che ne facilitano l'aggiunta: WikiPlayground 2.0 (WPG), WikiLinker per SBN + AuthorityBox SBN (WikiLinker) e VID-Q reconciliator (VID-Q), le cui caratteristiche sono riassunte nella Tabella 2 (Appendice 1).

4.1 WikiPlayground

WikiPlayground (WPG) è un'applicazione web²⁸ e mobile friendly realizzata da Lorenzo Gobbo, bibliotecario addetto ai Servizi Digitali per le biblioteche dell'Università della Svizzera Italiana. Dopo aver aperto una sessione in WD nel browser, si esegue una ricerca su WD tramite una query SPARQL allo scopo di individuare un gruppo di autori presenti in WD che abbiano cittadinanza italiana,²⁹ identificativo VIAF e siano privi di P396. La query è personalizzabile, ovvero si possono modificare i criteri di ricerca degli elementi privi di P396. A partire dal gruppo di elementi estratti da WD mediante la query, viene effettuata una ricerca per nome su SBN e vengono presentati i possibili abbinamenti (Figura 1). Tra questi l'operatore, confrontando altri dati o le opere associate in VIAF e in SBN, può decidere di confermare la corrispondenza tra l'elemento WD e il VID o di scartarla. In caso di conferma, WPG aggiunge il VID come valore di P396 in WD, con il qualificatore "soggetto indicato come" (P1810) che registra la forma preferita in SBN.

²⁵ <https://viaf.org/en/contributors?id=ICCU>.

²⁶ Secondo il dump VIAF del 01/04/2024, file `viaf-20240401-links.txt.gz`; l'ultimo dump VIAF disponibile è di agosto 2024 (<https://viaf.org/en/viaf/data>).

²⁷ Il processo è descritto in questo video del GWMAB: <https://www.youtube.com/watch?v=wndO45DWcpw>.

²⁸ WPG è presentato in <https://github.com/labaib/WikiPlayground2.0> ed è utilizzabile tramite <https://labaib.github.io/WikiPlayground2.0/>.

²⁹ Cittadinanza italiana include Regno d'Italia (pre 1946) e il caso di P396 "nessun valore" equivale alla presenza di P396.

The screenshot shows the WikiPlayground 2.0 interface. At the top, it says "WikiPlayground 2.0" and "Connected to Wikidata". Below that, it says "Query SPARQL". The main content is a list of seven potential matches, each with a profile picture, name, description, Wikidata ID, and a count of matches.

Name	Description	Wikidata ID	Match Count
Barbara Longhi	pittrice italiana	Q31146	0
Gigi Villorosi	pilota automobilistico italiano (1909-1997)	Q171514	0
Stophan El Shaarawy	calciatore italiano (1992-)	Q12984	0
Ricardo Martinoli	politico e imprenditore panamense (1952-)	Q57491	0
Renzo Zorzi	pilota automobilistico italiano	Q172231	1
Corrado Pardini	politico svizzero	Q81968	0
Ernesto Bertarelli	imprenditore svizzero	Q123702	2

Figura 1. Abbinamenti potenziali proposti da WPG 2.0.

4.2 WikiLinker

WikiLinker è costituito da una coppia di programmi³⁰ richiamati dal proprio browser tramite un'estensione, come p.es. Code Injector,³¹ che li esegue quando si sta visualizzando il risultato di una ricerca su Voci di autorità/Nomi nell'OPAC SBN.³² La videata dei risultati include una prima colonna con il VID, copiabile con un click, una seconda con un link per mostrare i cluster VIAF

³⁰ WikiLinker per SBN (introduzione <https://w.wiki/EfQN>, codice <https://w.wiki/EfQQ>) e AuthorityBox SBN (codice <https://w.wiki/EfQS>).

³¹ <https://github.com/Lor-Saba/Code-Injector>. Questa estensione non è più mantenuta; inoltre il browser Google Chrome, se aggiornato alle versioni più recenti, ne rifiuta l'installazione a causa di nuovi requisiti molto stringenti, motivati da ragioni di sicurezza; si può ancora installare in Firefox. Esistono altre estensioni simili a Code Injector, ma meno semplici da utilizzare.

³² <https://opac.sbn.it/voci-controllate-nomi>.

collegati alla forma preferita di SBN, e una terza riporta invece il link a WD nel caso in cui il VID sia già esistente in WD (Figura 2).³³

The screenshot shows the OPAC SBN interface. At the top, there is a navigation bar with the logo and text 'OPAC SBN INFORMAZIONI - RICERCA AVANZATA VOCI DI AUTORITY - ALTRI CATALOGHI - SERVIZI - BIBLIOTECHE'. Below this, a breadcrumb trail reads 'OPAC SBN > Voci di autorità > Nomi > Risultati Autori'. The main content area displays a list of 94 authors. Each row includes a checkbox, a small icon, the author's name with a count in parentheses, the type 'Persona', and a 'WikiLinker per SBN' button with a dropdown arrow. The dropdown menu is open for the first few authors, showing the SBN VID and a 'WD' link. For example, 'Achterbusch, Herbert (70)' has VID 'CFIV030785' and a 'WD' link. Other authors listed include Adlon, Percy (38), Alon, Fatih <1973-> (54), Becker, Wolfgang <1954-2024> (6), Benedix, Hugo (1), Berrini, Anne (1), Blank, Richard <1939-> (6), Breuer, Hans <1870-1929> (5), and Bruni Tedeschi, Valeria (138).

	Nome	Tipo	WikiLinker per SBN
<input type="checkbox"/>	Achterbusch, Herbert (70)	Persona	CFIV030785 VIAF WD
<input type="checkbox"/>	Adlon, Percy (38)	Persona	RAVV090321 VIAF WD
<input type="checkbox"/>	Alon, Fatih <1973-> (54)	Persona	VEAV487437 VIAF WD
<input type="checkbox"/>	Becker, Wolfgang <1954-2024> (6)	Persona	L01V311477 VIAF WD
<input type="checkbox"/>	Benedix, Hugo (1)	Persona	TSAV633207 VIAF WD
<input type="checkbox"/>	Berrini, Anne (1)	Persona	L01V553026 VIAF WD
<input type="checkbox"/>	Blank, Richard <1939-> (6)	Persona	T00V034002 VIAF WD
<input type="checkbox"/>	Breuer, Hans <1870-1929> (5)	Persona	UB0V280344 VIAF WD
<input type="checkbox"/>	Bruni Tedeschi, Valeria (138)	Persona	HILV193395 VIAF WD

Figura 2. WikiLinker per SBN

L'utente può lavorare sulle voci non ancora associate a WD: il passaggio del mouse sulla voce visualizza il numero di pubblicazioni associate,³⁴ mentre il passaggio del mouse ma con il tasto "maiuscole" premuto mostra tutti i campi del VID in una finestra che include anche i titoli di 20 pubblicazioni.

In caso di corrispondenza tra VID e Item di WD, l'utente potrà aggiungere manualmente il valore del VID in WD (P396).

Se non c'è già una corrispondenza, l'utente potrà creare in modo guidato l'elemento WD mancante. Si procede con un click sulla forma preferita di SBN, che aprirà il VID corrispondente arricchito dall'AuthorityBox, cioè il secondo programma di WikiLinker.³⁵ Il riquadro dell'AuthorityBox ha diverse funzioni ma, in questo caso, sono disponibili due bottoni (Figura 3): uno per la ricerca

³³ Un tutorial di WikiLinker si trova in <https://www.youtube.com/watch?v=YGMj8j6t8BE>.

³⁴ Il numero viene visualizzato accanto alla voce e si può cliccare per mostrare tutte le opere associate.

³⁵ AuthorityBox è stato introdotto nel catalogo della Pontificia Università della Santa Croce nel 2019 e viene presentato in (Bargioni 2020) e discusso in (Bianchini 2024).

in WD tramite stringa del nome di un VID SBN,³⁶ e uno per la creazione dell'elemento WD una volta verificato che non esista, tramite l'utilizzo di QuickStatements (QS).³⁷

DATAZIONE	n. 1880
NOTA INFORMATIVA	Poetessa, docente di lingua italiana, latino, storia, filosofia, collaborò attivamente al 'Mattino' di Napoli. Nata a Napoli.
FONTI	World biographical index. Internet-edition. K. C. Saur Electronic Publishing Munchen: www.saur-wbld.de Casalena, Maria Pio. Scritti storici di donne italiane. Bibliografia 1800-1945, Firenze, Olschki, 2003 Catalogo dei libri italiani dell'Ottocento, 1801-1900. Milano, Edarico Bibliografica, 1991
RECOLE DI CATALOGAZIONE	REICAT
ISNI	0000 0000 4355 8072
DATA DI AGGIORNAMENTO	26/01/2024
IDENTIFICATIVO SBN	ITVCCU\SBLV223806
<i>click sull'identificativo per copiarlo</i>	

AuthorityBox
Scalera, Anna
Non esiste un elemento corrispondente in Wikidata, o non ha un collegamento con SBN tramite la proprietà P396.
[Cerca in Wikidata](#) per aggiornarlo, o [Crea](#) se non lo hai trovato.
[Info e Impostazioni](#)

Figura 3. Il VID SBLV223806 (Anna Scalera) con le aggiunte dell'AuthorityBox evidenziate nei riquadri blu

Premendo sul bottone “Crea”, viene composto il gruppo di istruzioni (o comandi) per la creazione dell'elemento (Figura 4). Si noti che i dati di luogo di nascita (Napoli) e occupazione/i (poetessa, docente) non possono fare parte delle istruzioni perché non facilmente isolabili e soprattutto perché non riconciliati con i corrispondenti elementi WD. Questa operazione sarà possibile solo manualmente in WD, una volta creato l'elemento. La composizione di P569 per la data di nascita segue il formato indicato da WD. Il riferimento indica che il dato proviene dall'OPAC SBN (P248 = Q105086090), in particolare dal VID SBLV223806, consultato (P813) in data 7 maggio 2025.³⁸

³⁶ <http://id.sbn.it/bid/SBLV223806>.

³⁷ <https://quickstatements.toolforge.org/>.

³⁸ Inoltre P31 indica istanza di umano (Q5) e P21 genere femminile (Q6581072), tramite uno dei bottoni in basso a sinistra. La descrizione in italiano (Dit) è ottenuta modificando manualmente la nota informativa.

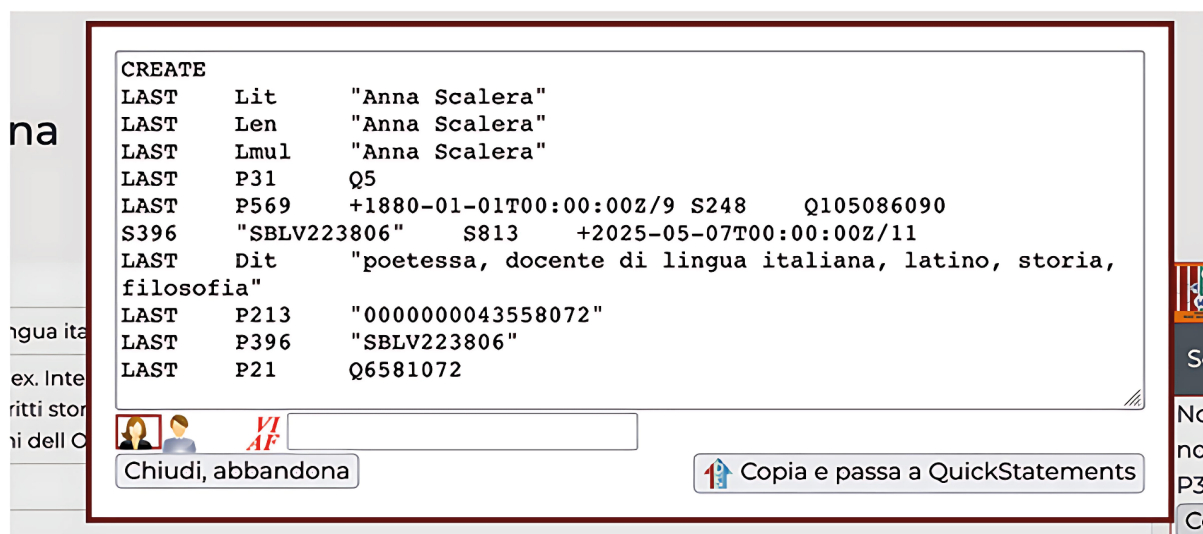


Figura 4. Le istruzioni di QuickStatements per la creazione di un nuovo elemento WD

4.3 VID-Q reconciliator

WPG e WikiLinker facilitano molto il lavoro di riconciliazione, ma la velocità operativa resta bassa perché trattano un caso alla volta. Considerata la quantità di lavoro necessaria, nemmeno una squadra di numerosi esperti può portare grandi risultati, se non in termini di qualità. D'altra parte, non è prudente delegare l'aggiunta di P396 ad operazioni del tutto automatizzate: per la riconciliazione dei nomi di autori devono essere valide due condizioni di corrispondenza: uguaglianza della forma preferita o di una forma variante con l'etichetta in italiano in WD e date di nascita e morte (anno o giorno). Se c'è corrispondenza solo per l'anno di nascita o di morte, occorre un ulteriore criterio di coincidenza.³⁹ Perciò, è stato creato un insieme di script in linguaggio Perl⁴⁰ – basati sulle API di ICCU-SBN⁴¹ – ognuno dei quali associa un VID a un elemento Q di WD, da cui il nome VID-Q reconciliator o VID-Q sotto forma di comandi QS del tipo:

```
Qnnn P396 "VID"
```

Le tre API di Nomi sono *search*, *vid* e *browse* e hanno l'endpoint comune <https://api.iccu.sbn.it/nomi/1.0.0/>. In VID-Q sono utilizzate le prime due, soprattutto *search*, che ricalca la ricerca per campi disponibile in Voci di autorità/Nomi. Per ottenere abbinamenti sicuri, in base ai criteri già citati, sono stati preparati 5 script, uno per ogni tipo di ricerca. Nelle Tabella 3 e 4 (Appendice

³⁹ Questo tipo di riconciliazioni per associazione di attributi comuni tra entità presenti in database diversi è utilizzato anche in (Tillman e Oaxaca 2025).

⁴⁰ Su Zenodo è stato pubblicato lo script principale, che permette la ricerca di VID via API in modo analogo alla ricerca in Voci di autorità/Nomi: (Bargioni, Bianchini, e Pellizzari di San Girolamo 2025).

⁴¹ <https://api.iccu.sbn.it/devportal/apis>. Le API richiedono un token di autenticazione (di tipo OAuth2; <https://www.rfc-editor.org/rfc/rfc6749>) basato su un account personale ottenibile gratuitamente, richiesto con una chiamata specifica, rinnovato ogni 3600 secondi e associato ad ogni chiamata a una API. Non vi è un limite di utilizzo delle API, né quantitativo né basato sulla frequenza delle chiamate. È buona pratica lasciare almeno un secondo tra una chiamata e l'altra, a tutela del lavoro ordinario dei server di SBN.

1) sono riportati rispettivamente i metodi di ricerca utilizzati in VID-Q e, per ogni metodo, un esempio di abbinamento ottenuto, ma sono possibili altri tipi di abbinamento.

Va comunque notato che le parole chiave e soprattutto le date utilizzate nelle ricerche delle voci SBN si possono trovare ovunque nel VID e perciò non sempre sono pertinenti. Lo script Perl infatti verifica se il VID trovato rispetta il criterio di abbinamento ma controlla anche in quali campi del record SBN si trova il dato; perciò le voci cercate su WD sono solo una parte, e di questa gli abbinamenti sono una parte ridotta, se non a volte nulla.

4.4 Dati quantitativi sugli abbinamenti

Gli abbinamenti del tipo 1 si basano su ricerche per coppie di decenni, di nascita e di morte, dal 1700 in poi, per un totale di 1587 abbinamenti (alcuni esempi in Tabella 5, Appendice 1).

Gli abbinamenti del tipo 2 si basano sulle decadi di nascita da 1700 fino a 1799 abbinate a decadi di morte da 1720 a 1899, sulle decadi di nascita da 1800 fino a 1899 abbinate alle decadi di morte da 1820 a 1999, e sulle decadi di nascita da 1900 fino a 1999 abbinate alle decadi di morte da 1920 a 2025, per un totale di 1587 abbinamenti (rispettivamente 186, 1016 e 385).

Gli abbinamenti del tipo 3, basati su nome e data con precisione superiore ad anno, sono stati applicati a 106 mesi, da gennaio 1900 a dicembre 2005, e sono raggruppati per decenni nella Tabella 6 (Appendice 1).

Per il tipo 4, basato su nome, anno di nascita e parola chiave dell'occupazione, ma solo per alcuni casi (colonna 1 della Tabella 7, Appendice 1), viene presentata una statistica più articolata, che permette di comprendere meglio il peso delle operazioni svolte da ogni passaggio.

La ricerca su SBN, a volte eseguita per troncamento (storic* per storico o storica), restituisce il numero di voci riportate in colonna 4. Il filtro permette di effettuare la ricerca su WD su un numero un po' inferiore di casi (colonna 5). Gli abbinamenti ottenuti sono invece quasi sempre molto inferiori (colonna 6), per varie cause: elemento non presente in WD, nonostante la ricerca per uguaglianza della forma scelta in SBN con l'etichetta in italiano "it", o inglese "en" o multilingua "mul",⁴² o con i possibili alias;⁴³ assenza in WD o mancata coincidenza delle date o dell'occupazione; abbinamento già presente in WD. Il totale degli abbinamenti ottenuti in questa fase è 9513; sono state trovate 173855 voci di SBN ed esaminati 151084 elementi WD.

Sempre per gli abbinamenti di tipo 4, ma con il confronto con la proprietà P611, sono state eseguite anche alcune ricerche per verificarne l'efficacia, riportate nella Tabella 8 (Appendice 1).

Benché la casistica delle prime tre colonne non sia certo esaustiva, la Tabella 8 mostra come si ottengano meno abbinamenti con questo ulteriore passaggio. Ma va comunque tenuto presente che ogni ricerca era seguita dalla registrazione in WD degli abbinamenti ottenuti, e quindi diminuiva via via la possibilità che una ricerca apparisse come efficace se l'abbinamento era già stato registrato.

Il criterio di abbinamento 5, basato sulla presenza di un identificativo di ambito italiano (IDita), ha dato risultati più significativi rispetto ai precedenti (Tabella 9, Appendice 1). Anche in questo caso, ogni estrazione poteva vedere ridotti gli elementi abbinati in funzione dei risultati delle precedenti.

⁴² <https://w.wiki/CjFz>.

⁴³ Non sono state esaminate le eventuali varianti riportate in SBN, per non moltiplicare il lavoro. Potrebbe essere un ulteriore metodo di ricerca, se si riesce a individuare facilmente le voci di SBN con varianti.

5. Conclusioni e prospettive future

Le diverse procedure di abbinamento tra VID di SBN ed elementi WD illustrate suggeriscono alcuni accorgimenti per migliorare ulteriormente il procedimento di riconciliazione, tanto relativi all'accessibilità e alla riusabilità dei dati di SBN che le procedure di creazione e modifica dei VID all'interno di SBN.

La prima necessità è che anche gli URI dei VID di livello 05 siano resi visibili nella sezione Voci di autorità; attualmente, infatti, per i VID di livello 05 WD è costretta a linkare a <https://opac.sbn.it/risultati-autori/-/opac-autori/detail/VID?core=autoriall> anziché a <https://opac.sbn.it/nome/VID> proprio per poter creare link.

Un secondo problema è che in Interfaccia Diretta (di seguito InD) mancano alcune funzioni che sono disponibili nell'interfaccia di ricerca Voci di autorità/Nomi e che sono molto utili per trovare VID da bonificare (es. cercare stringhe in tutti i campi del VID consente di trovare VID aventi, nella nota informativa, informazioni da spostare in altri campi, come l'ISNI⁴⁴).

Un altro miglioramento del processo riguarda l'aspetto legale del riuso dei dati contenuti nei VID; è necessario che venga indicata la licenza con cui sono rilasciati. Se si adottasse la licenza CC0, come per l'Anagrafe delle Biblioteche Italiane,⁴⁵ sarebbero possibili importazioni massive di dati in WD. Infine, il riuso dei dati e la realizzazione di analisi statistiche sarebbero molto avvantaggiate dal rilascio dei dump periodici dei VID, che, se contenessero anche il livello di autorità di ciascun VID (al momento non ottenibile da OPAC SBN), potrebbero fornire informazioni utili anche sulla distribuzione dei VID per livello e su come rendere la gestione dei livelli più efficiente. Se i dump contenessero anche i VID reindirizzati, si potrebbero usare periodicamente per aggiornare i VID in altri database.

Faciliterebbe il riuso dei dati anche la creazione di un endpoint SPARQL che permettesse di interrogare l'OPAC SBN, o almeno i VID, sul modello di BNF⁴⁶ e IDREF.⁴⁷

Riguardo alle procedure di creazione e modifica dei VID (tramite InD e i gestionali di polo), sarebbe necessario implementare alcuni campi UNIMARC. Per esempio, per migliorare l'identificabilità delle persone descritte nei VID, sarebbe utile introdurre un campo analogo al 670 ("source data found") del MARC21 usato dalla Library of Congress per indicare nelle sue voci di autorità quale è stato il record bibliografico per il quale la voce è stata originariamente creata; stante il manuale UNIMARC Authorities (IFLA 2024, s.v. 810), il campo "source data found" in UNIMARC è l'810 (utilizzato ad es. nelle voci di autorità della BNF); tuttavia in SBN il campo 810 è usabile solo per indicare la presenza della persona nei repertori; l'indicazione del record bibliografico originariamente usato si potrebbe quindi inserire o nella nota del catalogatore (campo UNIMARC 830) o in un nuovo campo appositamente introdotto.

Per rendere possibile strutturare identificativi diversi dall'ISNI (che è inserito nel campo UNIMARC 010), sarebbe da introdurre il campo UNIMARC 017 ("other identifier") – invece che

⁴⁴ <https://opac.sbn.it/risultati-autori?core=autori&item%3A1016%3AKeywords=ISNI>.

⁴⁵ <https://anagrafe.iccu.sbn.it/it/open-data/>.

⁴⁶ <https://data.bnf.fr/sparql/>.

⁴⁷ <https://data.idref.fr/sparql/>.

utilizzare il campo 810 Fonti o 830 Nota del catalogatore –, ripetibile in modo da ospitare più identificativi (es. nello stesso VID un ID WD, un ID VIAF e un ID ORCID) facilitandone il riuso. Sempre in un’ottica di miglior riuso dei dati, sarebbe utile disporre di campi più granulari per suddividere le informazioni che al momento vengono inserite in forma testuale nel campo 300 Nota informativa (eventualmente separate da “//”), come il campo 104 Principali date dell’entità per la datazione. Sarebbe utile prevedere inoltre: 1) uno o due campi appositi per luogo di nascita e luogo di morte (p. es. il campo 301\$a per luogo di nascita e 301\$b per luogo di morte), registrando i luoghi come entità identificate in database/vocabolari controllati (es. Geonames), così da garantirne l’identificazione e da migliorarne la traducibilità e l’interoperabilità; 2) un campo apposito, ripetibile, per le occupazioni/professioni, anch’esso preferibilmente con valori attestati in vocabolari controllati (es. nel Thesaurus del Nuovo Soggettario).

Vista l’importanza dell’ISNI in quanto standard ISO e anche unico identificativo inseribile nei VID in modo strutturato, sarebbe indispensabile poterlo utilizzare per specifiche funzioni di mantenimento dei dati: cercare un VID a partire da un ISNI; estrarre periodicamente ISNI da WD, come già vengono annualmente estratti dal VIAF; aggiornare periodicamente in modo massivo gli ISNI reindirizzati o cancellati; indagare la possibilità per i catalogatori di SBN, o almeno a una parte di essi, di creare direttamente gli ISNI mancanti, visto che ICCU è tra le ISNI Registration Agencies,⁴⁸ così da dotare soprattutto gli autori italiani di un ISNI in modo più sistematico.

L’uso attuale dei livelli di autorità per i VID dovrebbe essere ripensato affinché essi possano più efficacemente svolgere il loro ruolo positivo di proteggere i VID di buona qualità da possibili modifiche improprie, senza invece svolgere un ruolo negativo di ostacolo al miglioramento di VID di scarsa qualità o semplicemente ‘vuoti’ (ossia contenenti solo il campo 200 compilato). Purtroppo attualmente non è tecnicamente possibile abbassare di livello gruppi di VID (es. quelli di livello 90 e 95 ‘vuoti’); anche per questo, tuttavia, sarebbe importante stabilire norme sull’uso dei livelli in modo che, almeno per i VID lavorati in futuro, il livello scelto dal catalogatore sia proporzionale all’insieme dei campi compilati, e VID che sicuramente avranno bisogno di modifiche future (es. quelli di persone viventi) vengano posti a livelli che consentano a un adeguato numero di persone di apportare tali modifiche, riducendo così i possibili colli di bottiglia.

Si suggerisce infine di creare una procedura ufficiale tramite la quale coloro che già catalogano in SBN possano richiedere accesso in visualizzazione, e soprattutto in modifica, a InD, e al contempo di fornire dei corsi di formazione all’uso di InD con periodicità regolare, in modo da coinvolgere nella bonifica dei VID un maggior numero di persone dopo adeguata formazione.

Le prospettive future per la collaborazione tra WD e l’*authority file* di SBN si collocano nei tre ambiti fondamentali che si sono consolidati in questi anni: l’abbinamento reciproco, l’arricchimento reciproco, la verifica incrociata.

L’abbinamento tra elementi WD e VID ha visto un’accelerazione negli ultimi due anni grazie soprattutto a efficaci metodi di abbinamento automatico (VID-Q) e agli strumenti forniti agli utenti per rendere il lavoro manuale più efficiente (WPG, WikiLinker). Per aumentarne ulteriormente la velocità è necessario un approccio con metodi automatici sull’intero insieme dei VID – che sarebbe facilitato dalla disponibilità di dump periodici – e convogliare più forze verso il lavoro

⁴⁸ <https://isni.org/page/isni-registration-agencies/>.

di abbinamento manuale, pressoché indispensabile per i VID privi di informazioni identificative. WD e SBN possono beneficiare molto dai dati l'una dell'altro: per SBN, i risultati del primo arricchimento di VID 'vuoti' (concluso nel 2024) e del secondo (in corso nel 2025) sono buoni e si sta stabilizzando il *workflow* per farlo diventare una procedura periodica; per WD sarebbe possibile creare massivamente nuovi elementi a partire dai VID di buona qualità ma per farlo sarebbero necessari dump periodici e soprattutto l'attribuzione della licenza CC0 ai VID, sul modello di quanto fatto dalla Biblioteca Nazionale della Repubblica Ceca a partire dal 2019 (Jansová, Maixnerová, e Štastná 2024, 8–10).

I dati di WD e di SBN hanno ovviamente delle divergenze: analizzarle può permettere di rintracciare errori di abbinamento, nonché errori fattuali da entrambe le parti. Tuttavia, è difficile per ragioni di tempo aggiungere l'analisi di queste divergenze al flusso di lavoro ordinario dei bibliotecari impegnati quotidianamente nella creazione e manutenzione dei VID e sarebbe più fattibile inserire queste attività di controllo della qualità dei dati nell'ambito di progetti tematici che riguardino un insieme coerente di persone da descriversi in WD, come p. es. gli autori del *Dizionario degli scrittori italiani contemporanei pseudonimi* di Renzo Frattarolo (De Monaco 2025); rendere l'identificazione di una persona in SBN parte strutturale di ogni progetto di ricerca che descriva in WD un corpus di persone legate all'Italia consentirebbe di coinvolgere anche i ricercatori nella manutenzione dei VID SBN, rendendoli quindi patrimonio comune non solo al mondo delle biblioteche e iniziando una loro possibile apertura anche ad altre istituzioni culturali e al mondo della ricerca, secondo il modello del GND. In generale, tanto più l'*authority file* di SBN riuscirà a diventare un punto di riferimento nell'ambito dei Linked Open Data anche al di fuori del mondo delle biblioteche, tanto più aumenterà il numero e la diversità dei suoi utenti, e di conseguenza saranno sempre più individuati e corretti errori e imprecisioni, tanto più ne aumenterà la qualità, instaurando così un circolo virtuoso.

Appendice 1

Nome completo	Nome breve	Operatività	Ambito di partenza	Interfaccia	Linguaggio
WikiPlayground	WPG	manuale	WD	web	HTML e JavaScript
WikiLinker per SBN + AuthorityBox SBN	WikiLinker	manuale	SBN	web	JavaScript
VID-Q reconciliator	VID-Q	batch	SBN	linea di comando da terminale	Perl

Tabella 2 - Metodi di riconciliazione SBN-WD disponibili

Criterio di abbinamento	Esempio di ricerca	Note	
1. nome e due anni	191* 199*	nati tra gli anni 1910 e 1999	anni inclusi nel qualificatore <an-am>
2. nome e due anni	191* 199*	nati tra gli anni 1910 e 1999	anni ricavati dal campo Datazione
3. nome e data con precisione superiore ad anno	1910mm*	nati nel mese “mm” dell’anno 1910	anno mese e giorno nel campo Datazione
4. nome, anno di nascita e parola chiave dell’occupazione	19* chimic*	chimico/a del XX secolo	anno e occupazione ovunque
5. presenza di identificativo di ambito italiano (IDita)	query WD che estrae elementi con IDita ma senza SBN	il risultato della query viene analizzato da uno script apposito	le fonti degli identificativi sono indicate in Tabella 9

Tabella 3. Tipi di ricerca utilizzati in VID-Q

Critero di abbinamento	Comando utilizzato	SBN	Wikidata
1. nome e due anni da qualificatore	perl cerca_range_date.pl 193 200	RMSV963070 Arnold, Hans Joachim <1932-2006>	Q17322103 Hans-Joachim Arnold P569 31 mar 1932 P570 20 feb 2006
2. nome e due anni da campo Datazione	perl cerca_data_datazione.pl 193 200	UBOV397234 Advis, Luis Datazione: 1935-2004	Q3038827 Luis Advis P569 10 feb 1935 P570 9 set 2004
3. nome e data di nascita con precisione superiore ad anno, da campo Datazione	perl cerca_data_datazione.pl 19300*	UBOV624476 Fain, Veniamin Moiseevič Varianti Fain, Benjamin Datazione: 1930.02.17-2013.04.15	Q2897209 Benjamin Fain P569 17 feb 1930
4. nome, anno di nascita e parola chiave dell'occupazione	perl cerca_data_occup.pl chimic Q593644 19	SBLV286830 Asinger, Friedrich Datazione: 1907-1999 Nota informativa: Chimico ...	Q84786 Friedrich Asinger P569 26 giu 1907 P106 chimico (Q593644)
5. presenza di identificativo di ambito italiano (IDita)	perl q_IDita.pl file_elementi_estratti_dalla_query	UM1V037628 Carolina Santoni 1808 1878	Q100149180 Carolina Santoni

Tabella 4. Esempi di abbinamenti ottenuti con VID-Q

Decade anno nascita	Decade anno morte	Abbinamenti ottenuti
170*	173*	0
185*	192*	29
191*	199*	25
195*	202*	2

Tabella 5. Alcuni abbinamenti tra nome, date nascita e morte per decenni presenti nel qualificatore

Decade	Abbinamenti ottenuti
1900-1909	23
1910-1919	19
1920-1929	61
1930-1939	110
1940-1949	115
1950-1959	87
1960-1969	45
1970-1979	25
1980-1989	10
1990-1999	5
2000-2005	1
TOTALI	501

Tabella 6. Conteggi degli abbinamenti di tipo 3

Parola occupazione cercata in SBN	Q occupazione (valore di P106)	Data nascita troncata (secolo)	Trovati in SBN	Cercati in WD	Abbinamenti ottenuti	Durata minuti
pittore	Q1028181	19	698	660	69	19
pittrice	Q1028181	19	1006	871	85	20
sacerdote	Q42603	18	2001	1646	15	45
sacerdote	Q42603	19	5113	4144	46	110
filolog*	Q13418253	18	771	668	20	19
filolog*	Q13418253	19	2807	2481	49	77
giornalista	Q1930187	19	10986	9781	971	314
giornalista	Q1930187	18	2115	1806	42	51
traduttore	Q333634	19	3008	2628	203	73
traduttrice	Q333634	19	1537	1323	93	38

Tabella 7. Conteggi degli abbinamenti di tipo 4

Parola relativa all'ordine religioso	Q ordine religioso (valore di P611)	Data nascita troncata (secolo)	Trovati in SBN	Cercati in WD	Abbinamenti ottenuti	Durata minuti
gesuita	Q36380	15	251	203	11	6
gesuita	Q36380	16	543	431	13	13
gesuita	Q36380	17	402	325	4	9
gesuita	Q36380	18	437	378	31	10
gesuita	Q36380	19	898	770	49	21
domenican*	Q131479	14	66	47	1	1
domenican*	Q131479	15	118	74	1	2
domenican*	Q131479	16	104	61	1	3
domenican*	Q131479	17	86	61	1	3
domenican*	Q131479	18	141	104	7	3
domenican*	Q131479	19	323	260	22	10
francescano	Q913972	15	127	90	1	3
francescano	Q913972	16	184	120	2	3
francescano	Q913972	17	192	141	1	4
francescano	Q913972	18	451	401	6	11
francescano	Q913972	19	770	706	13	18
cappuccino	Q913972	15	42	30	0	1
cappuccino	Q913972	16	97	77	0	2
cappuccino	Q913972	17	97	73	0	2
cappuccino	Q913972	18	166	138	0	3
cappuccino	Q913972	19	319	277	0	7
TOTALI			5814	4767	164	135

Tabella 8. Conteggi degli abbinamenti di tipo 4, in base alla proprietà P611

Proprietà	Nome	Estratti con query	Abbinamenti ottenuti	% ottenuti
P8290	Archivio Storico Ricordi	2465	568	23,04%
P8982	Pontificio Istituto Archeologia Cristiana	1985	413	20,81%
P6715	Sistema Informativo Unificato per le Soprintendenze Archivistiche	558	56	10,04%
P7203	Dizionario biografico dei Friulani	433	50	11,55%
P9625	Dizionario Biografico dell'Educazione (1800–2000)	178	26	14,61%
P10844	Pontificia facoltà teologica Teresianum	787	157	19,95%
P8985	Dizionario bio-bibliografico dei bibliotecari italiani del XX secolo	88	0	0,00%
P5619	Fondazione Franco Fossati	70	10	14,29%
P5739	Pont. Univ. S. Croce	17345	3012	17,37%
P8034	Biblioteca apostolica vaticana	71609	8760	12,23%
P12458	Parsifal (URBE)	76894	4862	6,32%
P5731	Pontificia Università S. Tommaso	14089	17	0,12%
P8153	Accademia delle Scienze di Torino	759	119	15,68%
P1986	Dizionario Biografico degli Italiani	11871	635	5,35%
TOTALI		199131	18685	9,38%

Tabella 9. Fonti e conteggi degli abbinamenti di tipo 5

Riferimenti bibliografici

Anderson, Dorothy. 1974. *Universal Bibliographic Control. A long term Policy — A Plan for action*. München: Verlag Dokumentation.

Anderson, Dorothy. 2000. «IFLA's Programme of Universal Bibliographic Control: Origins and Early Years.» *IFLA Journal* 26 (3): 209-14. <https://doi.org/10.1177/034003520002600309>.

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>.

Bargioni, Stefano. 2020. «From Authority Enrichment to AuthorityBox: Applying RDA in a Koha environment.» *JLIS.it* 11 (1): 175–89. <https://doi.org/10.4403/jlis.it-12595>.

Bargioni, Stefano, Elisa Bellistri, Giovanni Bergamin, Carlo Bianchini, Giulio Blasi, Alessandra Boccone, Marilena Daquino, Roberto Delle Donne, Pierluigi Feliciati, Claudio Forziati, Alberto Gambardella, Stefano Gambari, Maurizio Lana, Valeria Lo Castro, Federico Meschini, Alessandra Moi, Rossana Morriello, Stefano Parise, Valdo Pasqui, Tiziana Possemato, Roberto Raieli, Cristina Silvani, Francesca Tomasi, Federico Valacchi, Chiara Veninata, Maurizio Vivarelli. 2022. «Ventisei bibliotecari e professori di biblioteconomia rispondono a dieci domande poste da Mauro Guerrini su temi legati alla metadattazione». In *Metadattazione*, 107–272. Milano: Editrice Bibliografica.

Bargioni, Stefano, Carlo Bianchini, e Camillo Carlo Pellizzari di San Girolamo. «Dati e script dell'articolo "Wikidata e SBN. Un bilancio di due anni di lavoro (2023-2025)»». Zenodo, 7 luglio 2025. <https://doi.org/10.5281/zenodo.15829966>.

Berners-Lee, Tim. 2006. *Linked Data – Design Issues*. <https://www.w3.org/DesignIssues/Linked-Data.html>.

Bianchini, Carlo. 2017. «Osservazioni sul modello IFLA Library Reference Model.» *JLIS.it* 8 (3): 86–99. <https://doi.org/10.4403/jlis.it-12416>.

Bianchini, Carlo. 2024. «Nessun catalogo è un isola». In *Parsifal. Un modello di collaborazione bibliotecaria per condividere la conoscenza registrata*, a cura di Silvano Danieli, 57–72. Firenze: Firenze University Press. <https://doi.org/10.36253/979-12-215-0356-2.10>.

Bianchini, Carlo, Stefano Bargioni, e Camillo Pellizzari di San Girolamo. 2021. «Beyond VIAF. Wikidata as a complementary tool for authority control in libraries.» *Information Technology and Libraries* 40 (2): 1–31. <https://doi.org/10.6017/ital.v40i2.12959>.

Bianchini, Carlo, Stefano Bargioni, e 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): 1–67. <https://doi.org/10.6092/issn.2283-9364/15078>.

Bianchini, Carlo, e Mauro Guerrini. 2014. *Introduzione a RDA. Linee guida per rappresentare e scoprire le risorse*. Milano: Editrice Bibliografica.

Bianchini, Carlo, e Lucia Sardo. 2022. «Wikidata: a new perspective towards universal bibliographic control.» *JLIS.it* 13 (1): 291–311. <https://doi.org/10.4403/jlis.it-12725>.

Cencetti, Elena, Camillo Carlo Pellizzari di San Girolamo, e Elisabetta Viti. 2025. «Termini, dati e collegamenti: ‘conversazioni’ tra il Thesaurus del Nuovo soggettario e Wikidata.» *Imagines* 12: 85–94.

De Monaco, Sara, «Il Dizionario degli scrittori italiani contemporanei pseudonimi in Wikidata. Metodologia e risultati» (Tesi di laurea magistrale, Università di Pisa, 2025). <https://etd.adm.unipi.it/t/etd-05132025-160428>.

De Monaco, Sara, e Camillo Carlo Pellizzari di San Girolamo. 2025. «Il Dizionario degli scrittori italiani contemporanei pseudonimi in Wikidata. Metodologia e risultati (dati)». Zenodo. <https://doi.org/10.5281/ZENODO.15474929>.

Fratтарolo, Renzo. 1975. *Dizionario degli scrittori italiani contemporanei pseudonimi. 1900-1975. Con un repertorio delle bibliografie nazionali di opere anonime e pseudonime*. Ravenna: Longo.

Guerrini, Mauro. 2020. *Dalla catalogazione alla metadattazione: tracce di un percorso*. Roma: AIB.

Guerrini, Mauro. 2022. *Metadattazione: la catalogazione in era digitale*. Milano: Editrice Bibliografica.

Guerrini, Mauro, e Lucia Sardo. 2003. *Authority control*. Roma: Associazione italiana biblioteche.

Guerrini, Mauro. 2018. *IFLA Library Reference Model (LRM). Un modello concettuale per le biblioteche del XXI secolo*. Milano: Editrice Bibliografica.

IFLA (International Federation of Library Associations and Institutions). 2016. *Overview of differences between IFLA LRM and the FRBR-FRAD-FRSAD models*. Draft. https://www.ifla.org/files/assets/cataloguing/frbr-lrm/transitionmapping_overview_20161207.pdf.

IFLA (International Federation of Library Associations and Institutions). 2017. *IFLA Library Reference Model. A Conceptual Model for Bibliographic Information*. A cura di Pat Riva, Patrick Le Boeuf, e Maja Zumer. Den Haag: IFLA. https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla_lrm_2017-03.pdf.

IFLA (International Federation of Library Associations and Institutions). 2024. *UNIMARC Authorities Format Manual*. Online edition. Version: 1.1.0, 2024. Den Haag: IFLA. <https://www.ifla.org/unimarc-updates/unimarc-authorities-format-manual-online-ed/>.

IFLA (International Federation of Library Associations and Institutions). Study Group on the Functional Requirements for Bibliographic Records. 2009. *Functional Requirements for Bibliographic Records: Final Report*, approved by the Standing Committee of the IFLA Section on cataloguing. September 1997; as amended and corrected through February 2009. München: K.G. Saur. http://www.ifla.org/files/cataloguing/frbr/frbr_2008.pdf.

IFLA (International Federation of Library Associations and Institutions). Working Group on Functional Requirements and Numbering of Authority Records (FRANAR). 2009. *Functional Requirements for Authority Data: a Conceptual Model. Final report*. München: K.G. Saur.

IFLA (International Federation of Library Associations and Institutions). Working Group on Functional Requirements for Subject Authority Records (FRSAR). 2010. *Functional Requirements for Subject Authority Data (FRSAD). A conceptual model*. München: K. G. Saur. <https://www.ifla.org/files/assets/classification-and-indexing/functional-requirements-for-subject-authority-data/frsac-final-report.pdf>.

- Jansová, Linda, Lenka Maixnerová, e Petra Štátná. 2024. «Library Data in Wikimedia Projects: Case Study from the Czech Republic.» *O-Bib. Das Offene Bibliotheksjournal* 11 (4): 1–19. <https://doi.org/10.5282/O-BIB/6081>.
- JSC (Joint Steering Committee for Development of RDA), a c. di. 2010. *Resource Description & Access. RDA*. American library association; Canadian library association; CILIP.
- Linked Data for Production, *Wikidata as a hub for identifiers*. 11 giugno 2020, <https://t.ly/677QP>.
- Manzoni, Laura. 2022. *Identificatori*. Roma: Associazione italiana biblioteche.
- Pellizzari di San Girolamo, Camillo Carlo. 2024a. «Storia della collaborazione tra Wikidata e le biblioteche della Rete URBE nel controllo di autorità». In *Parsifal. Un modello di collaborazione bibliotecaria per condividere la conoscenza registrata*, a cura di Silvano Danieli, 147–62. Firenze: Firenze University Press. <https://doi.org/10.36253/979-12-215-0356-2.17>.
- Pellizzari di San Girolamo, Camillo Carlo. 2024b. «The reconciliation of SBN authority records with Wikidata. Progresses and perspectives after a decade of work (2013-2023).» *JLIS.it* 15 (1): 33–44. <https://doi.org/10.36253/jlis.it-573>.
- Possemato, Tiziana. 2024. *Entity modeling: la terza generazione della catalogazione*. Firenze: Firenze University Press. <https://doi.org/10.36253/979-12-215-0393-7>.
- Ravelli, Elena. 2024. «Authority file di SBN e Wikidata. Un progetto per l'arricchimento dei dati del Catalogo nazionale.» (Intervento presentato ai Wikidata Days, Bologna, novembre 8, 2024). https://commons.wikimedia.org/wiki/File:Authority_file_di_SBN_e_Wikidata_-_Wikidata_Days_Bologna_2024.pdf.
- Solimine, Giovanni. 1995. *Controllo bibliografico universale*. Roma: Associazione italiana biblioteche.
- Tillett, Barbara B., Khaled Mohamed Reyad, e Ana Lupe Cristán, a c. di. 2006. «A Virtual International Authority File (VIAF)». In *IFLA Cataloguing Principles: Steps towards an International Cataloguing Code*, 3, 90–116. Berlin, Walter de Gruyter – K. G. Saur. <https://doi.org/10.1515/9783598440335.1.90>.
- Tillman, Ruth Kitchin, e Gala Campos Oaxaca. 2025. «Crossing Silos: Assessing the Utility of Identity Attributes in Name Reconciliation.» *portal: Libraries and the Academy* 25 (2): 299–319. <https://dx.doi.org/10.1353/pla.2025.a955947>.