

# Cantautorato e *Digital Humanities*. Per una valorizzazione dell'opera di Fabrizio De André, Lucio Dalla, Gianmaria Testa

Marcello Ranieri<sup>1</sup>

<sup>1</sup> Alma Mater Studiorum - Università di Bologna, Italia – marcello.ranieri2@unibo.it

## ABSTRACT (ITALIANO)

La produzione artistica di famosi cantautori italiani è ampiamente rappresentata nel web. Anche nel semantic web essi sono presenti, ma per dare un'immagine adeguata della loro creatività sarebbe necessario fornire informazioni più complete. Normalmente, i dati disponibili nelle piattaforme della conoscenza per ciascuno di questi artisti costituiscono un semplice segnaposto. Il numero di record recuperabile negli OPAC testimonia di una quantità di pubblicazioni discografiche che supera di gran lunga quello che riportano i più noti aggregatori. Questo contributo presenta un metodo, applicato ai casi di Fabrizio De André, Lucio Dalla e Gianmaria Testa, per convertire in Linked Open Data pronti per la pubblicazione i record bibliografici in formato MARC tratti dall'OPAC del Servizio Bibliotecario Nazionale, ponendo particolare attenzione alla riconciliazione dei dati.

**Parole chiave:** cantautori; discografia; digital library; Linked Open Data; web semantico

## ABSTRACT (ENGLISH)

*The "cantautorato" and the Digital Humanities. For a valorization of the work by Fabrizio De André, Lucio Dalla, Gianmaria Testa*

The artistic production of well-known Italian songwriters is widely represented on the Web. They are on the Semantic Web as well, but to adequately acknowledge their creativity it would be necessary to provide more complete information. Normally, the data available for each of those artists on knowledge platforms serves merely as a placeholder. The number of records available in OPACs bears witness to a quantity of published discography that far exceed what most famous aggregators register. This contribution presents a method, applied to the cases of Fabrizio De André, Lucio Dalla e Gianmaria Testa, for converting to Linked Open Data ready for publication the MARC formatted bibliographic records downloaded from the OPAC of the National Library Service, with particular attention to data reconciliation.

**Keywords:** songwriters; discography; digital library; Linked Open Data; Semantic Web

## 1. INTRODUZIONE

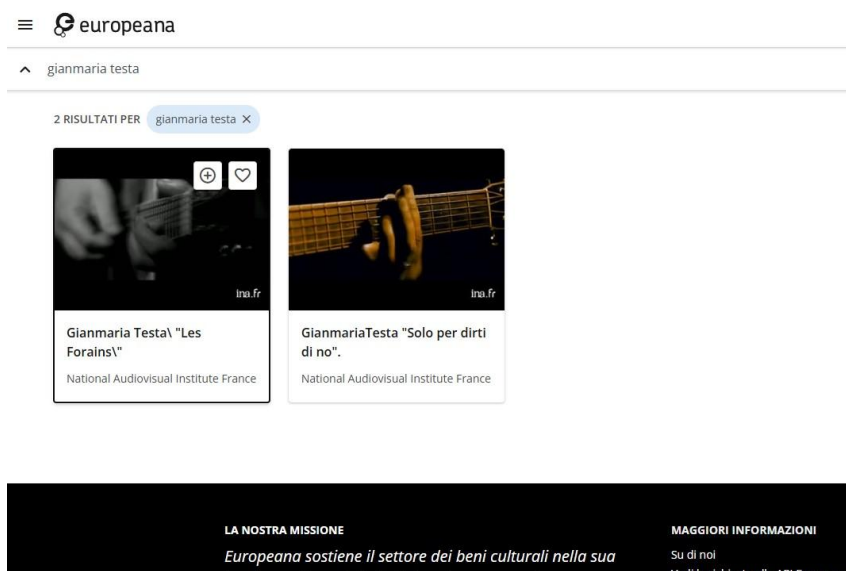
Le opere<sup>1</sup> del cantautorato italiano possono a pieno titolo essere considerate parte del patrimonio culturale per originalità di contenuti, profondità di ricerca musicale, attenzione ai temi sociali, popolarità trasversale.



**Figura 1. Immagine dal sito del Ministero della Cultura: la presentazione della "Casa dei Cantautori"**

<sup>1</sup> Va qui chiarito che la parola *opera* viene usata nel senso di *produzione artistica* che ha comunemente nell'italiano standard e non in quello tecnico assunto nello schema concettuale IFLA FRBR.

Un riconoscimento ufficiale dell'importanza di questa manifestazione creativa nella musica italiana si è avuto con la presentazione a Genova nel 2021 della "Casa dei Cantautori" (vedi Fig. 1), un luogo della cultura dedicato al cantautorato e collocato in una delle città che ne hanno visto le origini.<sup>2</sup> La canzone d'autore è fiorita nel Secondo Dopoguerra, ha raggiunto un altissimo livello negli anni Sessanta del Novecento e lo ha mantenuto costantemente evolvendosi parallelamente e in un certo senso in opposizione (Porciani, 2020: 379) al gusto corrente della musica leggera, rispecchiando allo stesso tempo i mutamenti del costume (Ciofalo, 2022: 335). Molti nomi rientrerebbero nel lungo elenco degli autori di tali testi e musiche da allora ai giorni nostri, permettendo la classificazione delle più o meno celebri canzoni prodotte in svariati sottogeneri. I testi dei cantautori sono stati e sono un mezzo inclusivo di circolazione delle idee e hanno sempre prestato una particolare attenzione alle categorie socialmente marginali e svantaggiate. Al di là di ogni considerazione storico-critica, valutando se l'opera dei cantautori abbia rappresentanza anche nella conoscenza trasposta in modelli semantici, non si può non notare come a una pervasiva presenza della produzione artistica cantautorale nel web non corrisponda una altrettanto completa rappresentazione di essa sotto forma di Linked Open Data. Si prenda a titolo di mero esempio il caso di tre dei più noti e amati tra i cantautori italiani degli ultimi decenni: Fabrizio De André (1940-1999), Lucio Dalla (1943-2012), Gianmaria Testa (1958-2016). Se per ciascuno di essi si verifica la presenza delle rispettive opere nei due noti e importanti luoghi di diffusione di contenuti culturali Europeana e Wikidata,<sup>3</sup> si troveranno i risultati seguenti. Su Europeana, per De André i documenti disponibili risultano 235, per Dalla 139, per Testa 2, come è evidente nella Fig. 2. Su Wikidata, a parte dati di carattere generale per ciascuno dei tre autori considerati e i loro identificatori nei principali social network, la sezione discografica è piuttosto esigua.



**Figura 2. Risultati di una ricerca su Europeana**

Alla voce "Fabrizio De André discography" (contrassegnata come item Q3709717) l'informazione viene suddivisa nei quattro tronconi "Fabrizio De André's albums in chronological order" (item Q76901267), "Fabrizio De André live albums discography" (item Q76901439), "Fabrizio De André singles discography" (item Q80600819), "Fabrizio De André tribute albums" (item Q106409719). Considerando anche questa ultima voce, il totale dei documenti segnalati attraverso la proprietà has part(s) (property P527) è di 27. Per Lucio Dalla i tronconi sono soltanto tre: "Lucio Dalla's albums in chronological order" (item Q64567874), "Lucio Dalla singles discography" (item Q64567991), "Lucio Dalla compilation albums discography" (item Q64620677). Il totale dei documenti in questo caso è di 40. Per Gianmaria Testa, infine, non compaiono dati sulla discografia. Se si verifica invece la presenza di documenti relativi a questi

<sup>2</sup> <https://cultura.gov.it/casadeicantautori> (cons. 25/01/2025)

<sup>3</sup> In particolare le pagine

<https://www.europeana.eu/it/search?page=1&view=grid&query=fabrizio%20de%20andr%C3%A9>,  
<https://www.europeana.eu/it/search?page=1&view=grid&query=lucio%20dalla> e  
<https://www.europeana.eu/it/search?page=1&view=grid&query=gianmaria%20testa> per Europeana e  
<https://www.wikidata.org/wiki/Q25106> , <https://www.wikidata.org/wiki/Q167546> e  
<https://www.wikidata.org/wiki/Q1522516> per Wikidata (cons. 25/01/2025)

tre cantautori con una ricerca sull'OPAC del Servizio Bibliotecario Nazionale (SBN),<sup>4</sup> avendo cura di scegliere la ricerca avanzata e di inserire nome e cognome del cantautore nel campo "Autore" della maschera di ricerca, i documenti risultanti qualificati come "monografia" (in opposizione a "spoglio") e contemporaneamente come "registrazione sonora musicale" risultano essere 948 per Fabrizio De André, 1097 per Lucio Dalla, 53 per Gianmaria Testa. Per quanto differenziati secondo la rispettiva produzione artistica dei tre cantautori scelti, e dunque più o meno grandi a seconda dei casi, questi numeri evidenziano una netta sproporzione rispetto a quanto confluisce nelle due piattaforme della conoscenza sopra considerate. Anche accettando la possibilità di ridondanze dovute a duplicati presenti tra i record SBN, la differenza di visibilità raggiunta nei due diversi canali dal lavoro dei tre cantautori è evidente.

## 2. METODOLOGIA ADOTTATA

La differenza di risultati nei casi osservati al paragrafo precedente non avrebbe motivo di esistere. La sproporzione con il portale Alfabetica curato dall'Istituto Centrale per il Catalogo Unico delle Biblioteche Italiane e per le informazioni bibliografiche (ICCU), ad esempio, è netta.<sup>5</sup> Non si esaminano qui le possibili motivazioni della dotazione ancora scarsa di molte piattaforme del web semantico, né si insiste sull'importanza della presenza di un dataset nello spazio globale dei dati, considerando l'assodata versatilità dei Linked Open Data nel creare relazioni tra i dati e nel permettere di "sfruttare tali dati in modo coerente, consapevole e arricchito dal punto di vista di uno o più mediatori culturali" (Daquino & Tomasi, 2016: 33). Semplicemente, si ipotizza un auspicabile riequilibrio della presenza della produzione cantautorale in un'organizzazione della conoscenza che rispecchi il più possibile il mondo reale. Una condivisione proficua di dati in quest'ottica è possibile, come tra l'altro testimoniano diverse infrastrutture basate sulla semantica, come ad esempio il portale Linked Open Data del Coordinamento Biblioteche Speciali e Specialistiche di Torino (CoBiS).<sup>6</sup> Oppure, in un contesto internazionale e per un diverso genere musicale, il portale Linked Jazz (Thorsen & Pattuelli, 2016).<sup>7</sup> In assenza di una sistematica operazione di *harvesting* che consenta alle basi di conoscenza di aggregare dati pubblicati dai fornitori, si individua qui una via praticabile per coadiuvare un simile riequilibrio. La metodologia è quella di individuare le risorse presenti nell'OPAC SBN e confrontarle con i risultati delle medesime ricerche in Europeana e in Wikidata per inserirvi quelle mancanti. Le strategie possibili per compiere tale lavoro sarebbero naturalmente molte e diverse. Si presenta qui un modello di procedura semiautomatizzata, applicata al caso di Europeana, nella quale ci si avvale anche di software in linguaggio Python per la conversione e il trattamento dei dati. La strategia di lavoro ideata consiste di una prima fase in cui si recuperano dati dall'OPAC SBN. In un secondo momento i record vengono convertiti dal formato UNIMARC adottato da SBN a quello dei Linked Open Data dell'Europeana Data Model. A questo punto viene effettuato il confronto semiautomatizzato con i risultati prodotti da una query all'API fornita da Europeana. Tutti i casi di record già presenti in Europeana vengono registrati in un apposito file disponibile per un controllo ulteriore da parte dell'operatore umano che individua eventuali falsi positivi per aggiungerli nuovamente alla lista di quelli da inserire. Infine, è previsto l'inserimento vero e proprio in Europeana, anch'esso da effettuarsi massivamente per mezzo di apposito software.

## 3. RISULTATI

Per la fase di ricognizione iniziale, dato l'ambito musicale delle risorse analizzate, si è individuato un polo particolare tra quelli presenti in SBN: quello dell'Istituto Centrale Beni Sonori e Audiovisivi (ICBSA).<sup>8</sup> Si è ritenuto che per lo scopo dimostrativo della ricerca effettuata i risultati forniti dall'OPAC del polo SBN dell'ICBSA fossero assolutamente qualificati a rappresentare il patrimonio documentale posseduto dalle biblioteche italiane. Per ciascuna delle tre ricerche si seleziona "Esporta Unimarc" tra le funzioni di servizio e si ottiene il relativo file, in formato testuale. Questa funzione è particolarmente apprezzabile per osservare i dati completi che compongono il record bibliografico. Si riporta in Fig. 3 come esempio il record

<sup>4</sup> <https://opac.sbn.it/web/opacsbn/risultati-ricerca-avanzata#1744293646260>, <https://opac.sbn.it/web/opacsbn/risultati-ricerca-avanzata#1744293719667> e <https://opac.sbn.it/web/opacsbn/risultati-ricerca-avanzata#1744293767985> (cons. 25/01/2025)

<sup>5</sup> <https://alphabetica.it/risultati/-/s/results?input=Fabrizio+De+Andr%C3%A9>, <https://alphabetica.it/risultati/-/s/results?input=Lucio+Dalla> e <https://alphabetica.it/risultati/-/s/results?input=Gianmaria+Testa> (cons. 25/01/2025)

<sup>6</sup> Il portale aggrega per Fabrizio De André 59 risultati, per Lucio Dalla 32, per Gianmaria Testa 7:

<https://dati.cobis.to.it/ricerca?q=Fabrizio%20De%20Andr%C3%A9&start=1>, <https://dati.cobis.to.it/ricerca?q=Lucio%20Dalla&start=1> e <https://dati.cobis.to.it/ricerca?q=Gianmaria%20Testa&start=1> (cons. 25/01/2025)

<sup>7</sup> <https://linkedjazz.org> (cons. 25/01/2025)

<sup>8</sup> <https://cloud.sbn.it/opac2/DDS/ricercaSemplice> (cons. 25/01/2025)

bibliografico del disco *Montgolfieres*, pubblicato da Gianmaria Testa nel 1995, come appare nel file testuale ricavato dalla codifica UNIMARC. Si noter , tra l'altro, anche la presenza dell'URL della versione digitalizzata della risorsa, all'interno del campo 899.<sup>9</sup> Essa, sia pur presente anche su Internet Culturale,<sup>10</sup> l'aggregatore del Ministero della Cultura, non   citata da Europeana (vedi Fig. 2) e non compare in Wikidata (come del resto nessun'altra opera pubblicata dal cantautore). Si ripete dunque la ricerca per ciascuno dei tre cantautori prescelti in modo da procedere. La ricerca avanzata di opere che hanno come autore Fabrizio De Andr , dopo aver eliminato gli spogli e trattenuto le sole "monografie", cio  le opere complessive dalle quali essi siano espunti, e tra le monografie le sole "registrazioni sonore musicali", fornisce 233 risultati. Per Lucio Dalla sono invece 192 e per Gianmaria Testa soltanto 5. Per la fase di conversione dei dati ricavati dal formato UNIMARC e mappatura di questi al formato Linked Open Data

```
LEADER 01174nm0a22002773i 4500
001 DDS0107776
005 20250105064301.0
035 $a(OCOLC)1116290524
071 01$a74321291902 $cMusicultura
100 $a20151117d1995 ||||0itac50 ba
101 | $aund
102 $ait
126 $aag|||||||b|||
181 1$6z01$ad $bbxxa
182 1$6z01$aa
183 1$6z01$asd$2RDACarrier
200 1 $aMontgolfieres$fGianmaria Testa
210 $aRecanati (AN)$cMusicultura$d1995
215 $a1 Compact Disc$c1.4 m./sec (compact disc)
700 1$aTesta$b, Gianmaria$c3L01V189636$4590
712 02$aMusicultura$3DDSV049071$4650
801 3$aIT$bIT-000000$c20151117
850 $aIT-RM0200
899 1$aIstituto Centrale per i Beni Sonori e Audiovisivi$bRM0200 $e$uhttp://opac2.icbsa.it/vufind/Record/IT-DDS0000017350000000
950 0$aIstituto Centrale per i Beni Sonori e Audiovisivi$d IC1C 2944$e IC 0002053455 Cda
copia 1$fT $h19970919$i20151119$lCopia 1$mCopia 1$d IC2C  92$e IC 0002053465 Cda copia 2$fT $h19970919$i20151119$lCopia 2$mCopia 2
977 $a IC
```

**Figura 3. Un record UNIMARC tratto dall'OPAC SBN**

di Europeana, serializzati in un file JSONLD, ci si   serviti di un apposito script in linguaggio Python. Tutto il software realizzato nell'ambito di questa ricerca   disponibile su Github.<sup>11</sup> Simili modalit  operative sono documentate in diversi ambiti del contesto dei Beni Culturali (Lorenzini, 2016: 214-215). Per la fase successiva del confronto con i documenti presenti in Europeana, si interroga l'*endpoint* per ottenere la lista in formato JSONLD.<sup>12</sup> A questo punto, mediante l'esecuzione di un apposito script in linguaggio Python (presente nel *repository* Github) il confronto viene automatizzato: in due cicli annidati un record per volta del file convertito da UNIMARC viene messo a confronto con un record per volta del file dei record tratti da Europeana fino a esaurimento o fino al riscontro di una similarit . Il confronto avviene tra autore e titolo presenti nei due record. Se sono simili (l'algoritmo misura se la similarit  tra le stringhe supera una certa soglia, per gestire anche i casi di grafie leggermente discordanti) il record di partenza viene accantonato in un apposito file che memorizza i record gi  presenti e il confronto riparte dal record SBN successivo. Tutti i record del file di partenza esportato dall'OPAC SBN e convertito in JSONLD che non trovano corrispondenza nei record Europeana vengono trascritti invece in un apposito file pronto per l'inserimento in Europeana. Un esame successivo effettuato direttamente da un operatore umano sui record accantonati verifica l'esattezza del riconoscimento per individuare eventuali falsi positivi che, quando presenti, vengono trasferiti nel file dei record pronti per l'inserimento. Un ultimo script Python, anch'esso caricato in Github,   stato predisposto per effettuare tramite l'API l'inserimento massivo dei record mancanti in Europeana.

#### 4. LAVORI FUTURI

Stabilire un flusso di lavoro che permetta di andare oltre il *proof of concept* e ottenere risultati apprezzabili, considerando la mole di dati da elaborare in modo sistematico, richieder  un progetto collettivo che veda la collaborazione tra pi  soggetti coinvolti con un intento condiviso e sotto un coordinamento efficace. Il presente contributo vuole essere uno stimolo affinch  questo avvenga, con

<sup>9</sup> <http://opac2.icbsa.it/vufind/Record/IT-DDS0000017350000000> (cons. 25/01/2025)

<sup>10</sup> <https://www.internetculturale.it/it/16/search/detail?id=oai%3A192.168.10.31%3A22%3ARM0200%3AIT-DDS0000017350000000&mode=all&teca=ICBSA> (cons. 25/01/2025)

<sup>11</sup> <https://github.com/arts-lod/songwriters> (cons. 25/01/2025)

<sup>12</sup> <https://api.europeana.eu/console/index.html?url=docs/v3/search.json> (cons. 25/01/2025. Per l'interrogazione dell'API   necessario ottenere una chiave di accesso che si pu  richiedere solo registrandosi)

l'auspicio che l'attenzione futura riservata a una fase storica fondamentale della musica italiana porti a trasferire anche nel web semantico tutta la conoscenza archiviata di essa. L'inserimento sulle piattaforme della conoscenza di dati ottenuti attraverso le procedure descritte richiederà in molti casi che vengano stipulati specifici accordi tra istituzioni che dovranno essere sanciti in via canonica con tutte le approvazioni necessarie.<sup>13</sup>

## RINGRAZIAMENTI

Il presente lavoro nasce spontaneamente dall'ammirazione per figure di artisti che con il loro talento e la loro applicazione hanno mutato il panorama sonoro italiano. Non è possibile citare tutti e tutte, ma è a loro che va una immensa gratitudine. Il software depositato su Github è stato ottimizzato utilizzando ChatGPT.

## BIBLIOGRAFIA

- Ciofalo, G. (2022). La televisione e il cantautorato: lo scenario italiano degli anni Ottanta tra mediazione e mediatizzazione. *Comunicazioni Sociali*, 44(3), 1. [https://doi.org/10.26350/001200\\_000166](https://doi.org/10.26350/001200_000166)
- Daquino, M., & Tomasi, F. (2017). Linked Cultural Objects: dagli standard di catalogazione ai modelli per il web of data. Spunti di riflessione dalla Fototeca Zeri. *Umanistica Digitale*, (1). <https://doi.org/10.6092/issn.2532-8816/7195>
- Lorenzini, M. (2016). Le API di Europeana come esempio di integrazione e rappresentazione delle risorse culturali. *Archeologia e calcolatori*, 211-216. ISSN 2385-202X
- Porciani, E. (2020). Bob Dylan e gli strimpellatori. La popular music nella narrativa del Settantasette e dintorni. *Between (Cagliari)*, 10(19). <https://doi.org/10.13125/2039-6597/4042>
- Thorsen, H. K., & Pattuelli, M. C. (2016). Linked Open Data and the Cultural Heritage landscape. In E. Jones & M. Seikel (A c. Di), *Linked Data for Cultural Heritage*. Facet Publishing. ISBN 9781783301621

---

<sup>13</sup> <https://pro.europeana.eu/page/aggregators> (cons. 25/01/2025)