

# Linked open data e rappresentazione del patrimonio culturale: un caso applicativo per diffondere la conoscenza dei beni culturali ecclesiastici nel web semantico

«DigItalia» 1-2022  
DOI: 10.36181/digitalia-00047

**Claudia Guerrieri**

*Centro servizi per i beni culturali ecclesiastici della CEI (Conferenza Episcopale Italiana)*

*Nel dominio dei beni culturali molte organizzazioni stanno applicando le tecnologie del web semantico e hanno affrontato progetti – alcuni dei quali sperimentali – di pubblicazione dei dati descrittivi del patrimonio culturale in linked open data (LOD): la sfida è far affiorare la conoscenza rappresentata dai dati rendendola immediatamente interpretabile dalle macchine. Attraverso la delicata fase di definizione non ambigua dei concetti e delle relazioni che rappresentano una determinata porzione di realtà (modellazione dei dati) il web semantico si pone come strumento per far emergere significati e produrre nuove relazioni reciproche.*

*Questo contributo ha l'obiettivo di porre le basi per una prospettiva strategica di produzione e pubblicazione di LOD del patrimonio informativo esposto sul portale BeWeB – Beni ecclesiastici in web, scelto come esempio altamente rappresentativo del dominio culturale in virtù non solo dei suoi numeri – circa 12 milioni di schede descrittive di beni culturali di proprietà ecclesiastica – ma soprattutto della modalità in cui è costruito e quotidianamente popolato. Il portale è ideato e coordinato dall'Ufficio nazionale per i Beni Culturali ecclesiastici e l'Edilizia di culto della Conferenza Episcopale Italiana (Ufficio nazionale BCE – CEI), con l'ambizione di offrire una lettura d'insieme sui beni culturali di proprietà ecclesiastica conservati in Italia, dove gli authority data assumono il ruolo di punto di snodo per la tessitura di relazioni tra risorse culturali di natura diversa (risorse storico-artistiche, architettoniche, bibliografiche, archivistiche, fotografiche, istituti di conservazione). La proposta di integrazione nel web semantico del patrimonio informativo esposto sul portale BeWeB mira al raggiungimento di una piena interoperabilità semantica, attraverso l'analisi delle ontologie già presenti nel dominio culturale, per favorirne un loro riuso.*

*BeWeB è dunque un case study presentato e introdotto da un inquadramento generale, utile a spiegare il contesto in cui sono iscritte le tecnologie che abilitano la realizzazione del web semantico e dei progetti di costruzione dei grafi della conoscenza (knowledge graph).*

Questo contributo ha l'obiettivo di porre le basi per una valutazione strategica di produzione e pubblicazione di linked open data (LOD) del patrimonio informativo esposto sul portale *BeWeB – Beni ecclesiastici in web*<sup>1</sup>, scelto come esempio altamente rappresentativo del dominio culturale in virtù non solo dei suoi numeri – circa 12 milioni di schede descrittive di beni culturali di proprietà ecclesiastica –, ma soprattutto della modalità in cui è costruito e quotidianamente popolato<sup>2</sup>.

Il portale è ideato e coordinato dall'Ufficio nazionale per i Beni Culturali ecclesiastici e l'Edilizia di culto della Conferenza Episcopale Italiana (Ufficio nazionale BCE – CEI), con l'ambizione di offrire una lettura d'insieme sui beni culturali di proprietà ecclesiastica conservati in Italia, dove gli authority data<sup>3</sup> assumono il ruolo di punto di snodo per la tessitura di relazioni tra risorse culturali di natura diversa. *BeWeB* si configura come un portale cross-domain presentando risorse culturali di diversa natura: beni architettonici, storico-artistici, librari, archivistici oltre a schede di istituti culturali ecclesiastici (musei, archivi, biblioteche), percorsi e approfondimenti tematici, eventi e news di settore, voci di glossario e pagine descrittive di diocesi e regioni ecclesiastiche.

Il proposito di sperimentare l'utilizzo delle tecnologie del web semantico nell'ambito del portale *BeWeB* va nella direzione di far emergere significati e produrre nuove relazioni reciproche tra i dati che descrivono i beni culturali ecclesiastici e gli altri mondi – più o meno tangenti – collegati non solo al dominio culturale: la sfida è quella di far emergere la conoscenza rappresentata dai dati rendendola immediatamente interpretabile dalle macchine. L'investimento degli ultimi anni sulle attività di authority control da parte dell'Ufficio nazionale BCE, che gestisce il portale *BeWeB*, si configura come presupposto per il collegamento a dataset di autorità nel linked data cloud<sup>4</sup>, come quelli di VIAF e Wikidata<sup>5</sup>.

<sup>1</sup> <<https://beweb.chiesacattolica.it/>>. Si veda: *BeWeB 2020: vent'anni del portale*, a cura dell'Ufficio per i beni culturali ecclesiastici e l'edilizia di culto, Roma: Gangemi, 2020; Stefano Russo, *BeWeB. The cross portal of cultural ecclesiastical heritage*, «JLIS.it», 5 (2014), n. 2, p. 147-157 <<https://www.jlis.it/index.php/jlis/article/view/227/226>>.

<sup>2</sup> Questa ricerca applicata è stata oggetto di un lavoro di tesi di specializzazione: Claudia Guerrieri, *Linked open data del patrimonio culturale ecclesiastico. Una proposta per integrare il patrimonio informativo del portale BeWeB nel web semantico*, tesi di specializzazione, Sapienza Università di Roma. Scuola di specializzazione in Beni archivistici e librari, a.a. 2019-2020, relatore prof. Giovanni Michetti.

<sup>3</sup> CEI-AF è il sistema cross-domain di gestione degli authority file in *BeWeB*. Per approfondire: Paul G. Weston - Francesca D'Agnelli - Silvia Tichetti - Maria Teresa Rizzo - Claudia Guerrieri, *Gli Authority data e l'intersezione cross-domain nei portali ad aggregazione. Il portale BeWeB*, «JLIS.it», 8 (2017), n. 1, p. 138-154.

<sup>4</sup> All'indirizzo <<https://lod-cloud.net>> è possibile apprezzare il costante accrescimento della linked data cloud.

<sup>5</sup> Nel 2019 l'Ufficio BCE della Conferenza episcopale italiana e Wikimedia Italia hanno siglato un accordo per lavorare insieme con l'obiettivo di favorire l'integrazione dei contenuti del portale dei beni culturali ecclesiastici *BeWeB* con i progetti Wikimedia e OpenStreetMap.

Nell’ottica di raggiungere la piena interoperabilità semantica del patrimonio informativo di BeWeB, questo lavoro preliminare è stato affrontato metodologicamente attraverso l’analisi delle ontologie già presenti nel dominio dei beni culturali nell’ottica di un loro riuso: a valle di questa fase di analisi andrà verificata l’eventuale presenza (come residualità) di aree semantiche non ancora indagate e definite dalle ontologie esistenti per questo dominio di conoscenza, considerata la peculiarità del patrimonio ecclesiastico.

Per avviare una valutazione per progetto di produzione di linked open data delle risorse informative presentate sul portale BeWeB, si è fatto riferimento alla sintesi dei passi operativi da seguire, elaborata dal World Wide Web Consortium nello schema seguente<sup>6</sup>.

Produzione LOD: le FASI	
1	Preparare gli stakeholder
2	<i>Selezionare un set di dati</i>
3	<i>Modellare i dati</i>
4	Specificare una licenza appropriata
5	Creare buoni URI per i linked data
6	<i>Utilizzare vocabolari standard</i>
7	<i>Convertire i dati</i>
8	Fornire alla macchina l’accesso ai dati
9	Annunciare nuovi set di dati
10	Riconoscere il contratto sociale

Le fasi operative prese in esame in questo contributo sono quelle evidenziate in corsivo nella tabella precedente, cioè: la fase 2 “Selezionare un set di dati”, la fase 3 “Modellare i dati”, la fase 6 “Utilizzare vocabolari standard”, la fase 7 “Convertire i dati”<sup>7</sup>.

<sup>6</sup> W3C Working Group Note. (2014, January 9). *Best Practices for Publishing Linked Data*. W3C. <<https://www.w3.org/TR/ld-bp/>>.

<sup>7</sup> Sebbene sia di fondamentale importanza, si tralascerà in questo contributo la fase 1 sulla preparazione degli stakeholders. Applicando la filosofia open data, l’ipotesi di un progetto di produzione di LOD del patrimonio informativo pubblicato sul portale BeWeB deve preliminarmente accogliere il principio che i dati pubblicati possano essere usati e collegati da terzi, secondo i vincoli previsti dal tipo di licenza aperta con cui si deciderà di renderli disponibili, senza quindi predefinire un obiettivo conoscitivo o uno specifico uso futuro da parte di chi vi attingerà. In quanto patrimonio conoscitivo universale, questa porzione di conoscenza potrebbe in futuro essere usata da altri per rispondere a domande che ancora non ci si è posti. Pertanto, la fase 4 – che ha a che fare con l’individuazione della licenza d’uso più appropriata – al momento non verrà indagata approfonditamente, dando per scontato che si tratterà di una licenza aperta che dovrà tenere conto delle classi di oggetti che necessitano di maggiori cautele, a causa della sensibilità di alcuni dati. Sebbene di grande rilievo, non sono oggetto di questa riflessione neanche gli aspetti ricompresi nelle fasi 9 e 10 (“Annunciare nuovi set

## La selezione del dataset

Partendo dunque dal primo punto evidenziato nello schema precedente, cioè la fase 2 relativa alla selezione di un dataset rappresentativo, in questa proposta applicativa di produzione di LOD l'Ufficio BCE ha ipotizzato di partire, come campione rappresentativo, dal dataset degli istituti culturali (archivi, biblioteche, musei di proprietà ecclesiastica)<sup>8</sup>, intesi nella loro doppia accezione di istituti con funzioni di servizio verso l'utenza e di contenitori/conservatori di beni culturali: questa scelta è volta anche a indagare un possibile futuro terreno di lavoro comune con i progetti LOD curati dal Ministero della Cultura italiano, anche alla luce dei rapporti e del dialogo tra il portale BeWeB e altri sistemi informativi. L'ontologia sui luoghi e istituti della cultura usata dal Ministero della Cultura in Italia è Cultural-ON<sup>9</sup>, riusata direttamente anche nella rete di ontologie ArCo elaborata dall'Istituto centrale per il catalogo e la documentazione (ICCD)<sup>10</sup>, e più in generale inclusa nella rete di ontologie per la pubblica amministrazione italiana OntoPIA<sup>11</sup>, di cui il settore dei beni culturali fa parte.

di dati" e "Riconoscere il contratto sociale") che riguardano la sostenibilità degli aggiornamenti, alla luce della grande mole di dati in continuo accrescimento sul portale BeWeB. Nel momento in cui si producono LOD è in vigore un contratto sociale implicito: sono annunciati i nuovi set di dati che si stanno esponendo su un dominio autorevole e occorre riconoscere la propria responsabilità nel mantenere i dati una volta pubblicati. Bisogna assicurarsi che i set di dati rimangano disponibili là dove si dichiara che lo sono, e che vengano mantenuti nel tempo. Questo aspetto ha a che fare anche con il punto 5 ("Creare buoni URI per i linked data"), relativo alla strategia di costruzione degli URI, da ponderare, definire e denominare con un piano di implementazione. L'attenzione all'individuazione e denominazione degli oggetti, al supporto multilingue, alla modifica e aggiornamento dei dati nel tempo e alle strategie di persistenza degli URI sono elementi costitutivi dell'utilizzo dei LOD: anche il punto 8 ("Fornire alla macchina l'accesso ai dati") si occupa di questi aspetti, poiché si concentra sulla necessità di fornire ai motori di ricerca e ad altri processi automatizzati diversi modi per poter accedere ai dati, utilizzando meccanismi standard.

<sup>8</sup> L'Anagrafe degli Istituti culturali ecclesiastici (AICE), gestita e coordinata dall'Ufficio BCE della CEI, è fonte dei dati esposti sul portale BeWeB nella sezione Istituti culturali: <<https://beweb.chiesacattolica.it/istituculturali/>>. Si segnala inoltre che grazie a due convenzioni con l'Istituto centrale per il catalogo unico (ICCU) e l'Istituto centrale per gli archivi (ICAR), i dati descrittivi presenti in AICE sono fonte ufficiale delle schede descrittive degli istituti culturali censiti come istituti conservatori di patrimoni culturali di proprietà ecclesiastica nelle rispettive anagrafi nazionali del Ministero della cultura in Italia: ABI (Anagrafe delle biblioteche italiane) e SAN (Sistema archivistico nazionale). I testi dei due accordi sono disponibili agli indirizzi: <<https://bce.chiesacattolica.it/2011/06/13/lettera-circolare-unbce-iccu-per-scambio-dati-anagrafe-delle-biblioteche/>> e <<https://bce.chiesacattolica.it/2015/03/10/convenzione-tra-unbce-e-istituto-centrale-per-gli-archivi-del-mibact-per-la-definizione-di-procedure-condivise-per-il-popolamento-e-laggiornamento-dei-rispettivi-sistemi-beweb-e-san/>>.

<sup>9</sup> Maggiori informazioni sul progetto di realizzazione dell'ontologia Cultural-ON all'indirizzo: <<http://wit.istc.cnr.it/stlab-lod/#cultural-heritage>>. La versione attuale dell'ontologia Cultural-ON è la 3.2, rifattorizzata nel febbraio 2019 per adeguarla e integrarla con le ontologie di OntoPIA, rilasciate dall'Agenzia per l'Italia digitale (AGID) per la modellazione dei dati prodotti dalla pubblica amministrazione italiana. Per approfondire: <[https://dati.beniculturali.it/cultural\\_on/](https://dati.beniculturali.it/cultural_on/)>.

<sup>10</sup> Chiara Veninata, *Dal Catalogo generale dei beni culturali al knowledge graph del patrimonio culturale italiano: il progetto ArCo*, «DigItalia. Rivista del digitale nei beni culturali», 15 (2020), n.2, p. 43-56.

<sup>11</sup> Maggiori informazioni sulla rete di ontologie e vocabolari controllati della pubblica amministrazione

Sono state inoltre prese in esame diverse ontologie del dominio culturale<sup>12</sup>, oltre ad altre ontologie e schemi di vocabolario non strettamente legati al dominio culturale, ma adottati anche in questo settore nella prassi, perché modellano dei concetti riusabili anche nell’ambito del cultural heritage. Di seguito un prospetto delle ontologie/schemi di vocabolario analizzati:

Principali ontologie usate nel dominio dei beni culturali	
Ontologie Top-level / di dominio (culturale)	CIDOC CRM; EUROPEANA DATA MODEL (EDM ontology); CULTURAL-ON; ARCO (Architettura della conoscenza)
dominio archivistico	RIC-O; EAC-CPF; SAN LOD; OAD Ontology
dominio bibliografico	BIBFRAME; MODS E MADS Ontology; BIBO; FABIO; RDA; ISBD; IFLA-LRM; FRBR-oo
Ontologie della geoinformazione	GEONAMES Ontology; WSG84 GEO POSITIONING; LOCATION CORE VOCABULARY
Ontologie cross-domain	Friends of a Friend (FOAF); ORGANIZATION Ontology (ORG); PROV-O (Provenance ontology)
Ontologie leggere/Schemi di vocabolario	DUBLIN CORE; SCHEMA.ORG; DBPEDIA Ontology

### La modellazione dei dati e le ipotesi di riuso dei modelli ontologici

A partire dal primo dataset selezionato per l’avvio di un progetto sperimentale di produzione di LOD del portale BeWeB – costituito dai dati descrittivi degli istituti culturali come istituti conservatori dei patrimoni culturali – è stato poi avviato l’approfondimento per la fase 3 di “modellazione dei dati”, mediante la definizione non ambigua dei concetti e delle relazioni che rappresentano la porzione di realtà descritta.

È stata dunque estesa l’analisi anche ad altri dataset presenti in BeWeB, cioè a quelli che descrivono le diverse tipologie di risorse culturali rappresentate: beni architettonici, storico-artistici (e a breve fotografici), archivistici e librari.

italiana (OntoPIA), sviluppata dall’Agenzia per l’Italia digitale (AGID), sono disponibili all’indirizzo: <<https://github.com/italia/daf-ontologie-vocabolari-controllati>>.

<sup>12</sup> Per approfondire: Chiara Veninata, *Inside the Meanings. The usefulness of a register of ontologies in the cultural heritage sector*, «JLIS.it», 11 (2020), n. 2, p. 45-58, <<https://www.jlis.it/index.php/jlis/article/view/44>>.

I primi due settori disciplinari sono presenti anche nella gerarchia definita dal modulo ArCo (*arco:*)<sup>13</sup> dell'Istituto centrale per il catalogo e la documentazione (ICCD): è stato quindi avviato un primo confronto per verificare la riusabilità di questo modello per i dati descrittivi di questi patrimoni presenti in BeWeB.

Nel modello ArCo, la superclasse principale *arco:CulturalProperty* rappresenta un bene culturale, che può essere definito materiale o immateriale usando le classi *arco:TangibleCulturalProperty* e *arco:IntangibleCulturalProperty*<sup>14</sup>. Relativamente ai beni materiali, ArCo distingue ulteriormente tra beni mobili (*arco:MovableCulturalProperty*) e immobili (*arco:ImmovableCulturalProperty*).

I settori disciplinari già presenti in BeWeB e modellabili ontologicamente con il modulo ArCo (*arco:*) sono quelli, come anticipato, dei beni architettonici (beni immobili) e dei beni storico-artistici (beni mobili)<sup>15</sup>: all'interno della classe corrispondente al settore disciplinare, saranno presenti sottoclassi più specifiche relative al tipo di oggetto rappresentato dalla scheda di catalogo esposta su BeWeB<sup>16</sup>. Altre tipologie di beni culturali come i beni fotografici, i beni musicali, e i beni immateriali sono già modellati in ArCo, ma saranno presenti in BeWeB in futuro<sup>17</sup>.

<sup>13</sup> Il modulo *arco* importa tutti gli altri moduli della rete di ontologie ArCo e modella le informazioni considerate centrali del dominio dei beni culturali. Per maggiori informazioni sulla rete di ontologie ArCo:

<<http://www.iccd.beniculturali.it/it/progetti/4597/arco-architettura-della-conoscenza-ontologie-per-la-descrizione-del-patrimonio-culturale>>.

<sup>14</sup> È definita come sottoclasse di *CulturalProperty* anche la classe *arco:DemoEthnoAnthropologicalHeritage*, che rappresenta il patrimonio culturale demoetnoantropologico, di tipo immateriale e materiale.

<sup>15</sup> Si ricorda che i beni architettonici e i beni storico-artistici sono descritti utilizzando la normativa nazionale a standard ICCD, nei relativi progetti CEI di catalogazione di settore (rispettivamente i progetti CEI-A e CEI-OA). In particolare viene adottato lo standard ICCD-A per i beni architettonici, arricchito con informazioni relative all'adeguamento liturgico subito dalla chiesa dopo la riforma liturgica del Concilio Vaticano Secondo, oltre che di alcune informazioni base sulla accessibilità delle chiese ai diversamente abili e il rispetto delle normative per gli impianti, e lo standard di livello inventariale ICCD-OA-D-N per i beni storico-artistici, integrato con paragrafi di precatalogo e catalogo e con altri campi di natura strettamente ecclesiastica. Ad esempio, è stato integrato nel tracciato ICCD un paragrafo chiamato DAE (Dati di amministrazione ecclesiastica) che rileva informazioni precise sulla geografia territoriale/amministrativa ecclesiastica, un campo per rilevare dati sull'uso liturgico del bene culturale (USOL – Uso liturgico) e sul livello di rischio del bene oggetto di catalogazione (LIVR – Livello di rischio: alto/medio/basso).

<sup>16</sup> Per i beni architettonici e per i beni storici e artistici presenti in BeWeB si ritiene possibile il riuso anche del modulo di ArCo Catalogue ("a-cat:"), che si occupa di conferire maggiore potere espressivo alla modellazione della singola scheda di catalogo. In potenza, questo riuso potrebbe essere esteso anche per la modellazione ontologica dei beni fotografici (a breve pubblicati in BeWeB) e per i beni immateriali (in futuro). Per quanto attiene invece ai beni musicali e in particolare agli strumenti musicali, ricordiamo che la schedatura degli organi a canne storici è svolta direttamente nel sistema coordinato dall'ICCD (SIGEC-web) e il risultato di questa attività non è esposto su BeWeB. Potenzialmente, attraverso l'individuazione di una relazione di tipo contenuto-contenitore in grado di collegare l'organo a canne storico al bene architettonico in cui è conservato (consultabile su BeWeB) si potrebbe realizzare un proficuo dialogo tra i sistemi CEI e quelli del Ministero della Cultura italiano per poter visualizzare in futuro questo collegamento anche in BeWeB.

<sup>17</sup> In particolare, i beni fotografici, catalogati nell'ambito del progetto CEI-F seguendo lo standard

Dal punto di vista delle relazioni di tipo contenuto-contenitore il modello dati in BeWeB è il seguente:

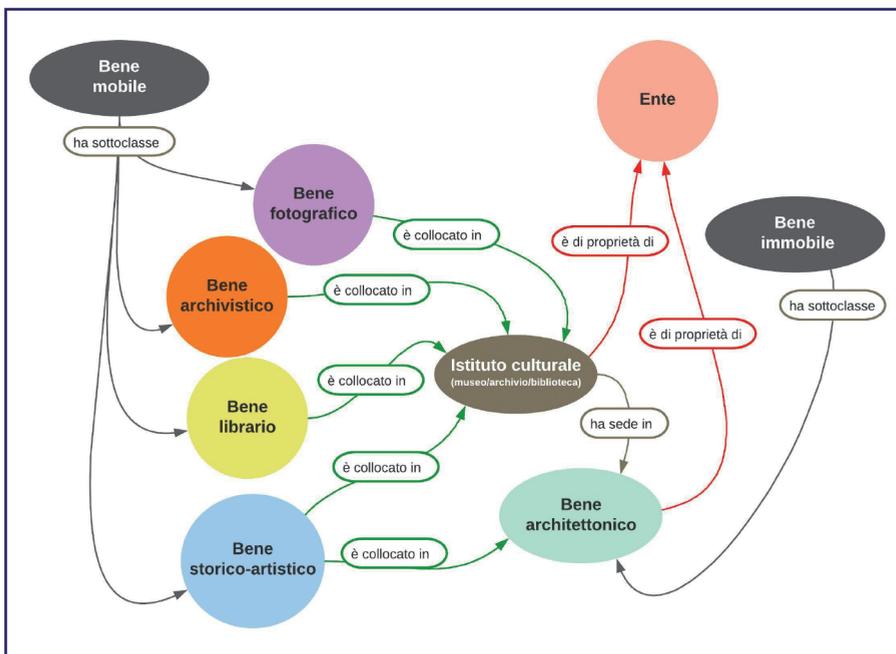


Figura 1. Relazioni contenuto-contenitore tra bene librario/archivistico/fotografico/storico-artistico (contenuti) e istituto culturale/bene architettonico (contenitori)

BeWeB presenta anche i risultati delle campagne di descrizione e catalogazione del patrimonio archivistico e librario conservato presso gli istituti culturali di proprietà ecclesiastica in Italia. Pertanto, per questi due settori – non modellati nella rete di ontologie ArCo – si potrebbe sperimentare l'applicazione delle ontologie di dominio più usate e più supportate a livello internazionale:

- per i beni librari, si potrebbe valutare l'adozione del modello BIBFRAME<sup>18</sup> (Bibliographic framework) della Library of Congress;
- per i beni archivistici, potrebbe essere interessante avviare un'applicazione sperimentale del modello RiC-O (*Records in Context – Ontology*)<sup>19</sup> proposto dall'International Council on Archives (ICA) e oggetto di dibattito nella comunità archivistica.

ICCD-F (con alcune estensioni), saranno a breve consultabili sul portale BeWeB, mentre i beni musicali, cioè gli strumenti musicali come gli organi a canne storici, sono schedati in Sigec-web mediante standard ICCD-SMO. I beni immateriali non sono ancora stati pubblicati come risorse culturali sul portale BeWeB, ma lo saranno in futuro.

<sup>18</sup> Il modello ontologico BIBFRAME è disponibile all'indirizzo: <<http://id.loc.gov/ontologies/bibframe/>>.

<sup>19</sup> La versione 0.2 del modello ontologico RIC-O è stata rilasciata a febbraio 2021, disponibile all'indirizzo: <[https://www.ica.org/standards/RiC/RiC-O\\_v0-2.html](https://www.ica.org/standards/RiC/RiC-O_v0-2.html)>: questa versione è conforme all'ultima

Relativamente ai beni librari, il modello BIBFRAME è organizzato in tre classi o livelli principali di astrazione: *work* (cioè l'opera, il livello di astrazione più alto, l'essenza concettuale della risorsa collegata), *instance* (cioè la pubblicazione, una particolare forma pubblicata dell'opera) e *item* (copia effettiva di una pubblicazione, oggetto fisico)<sup>20</sup>.

A livello di *work* si instaurano le relazioni con il *creator* (l'autore – *agent* – dell'opera), con il *subject* (cioè il soggetto, l'argomento trattato, che può essere un luogo, un agente come persona/famiglia/ente, un'espressione temporale, un evento, un'opera, un oggetto o altro). A livello di *instance* si instaurano relazioni con l'editore della pubblicazione, con la data di pubblicazione, con il formato ecc. A livello di *item* si instaurano relazioni con la collocazione fisica (o virtuale) della copia effettiva, con il barcode, o con il possessore della copia. Le proprietà di BIBFRAME descrivono le caratteristiche della risorsa descritta, nonché le relazioni tra le risorse.

Questo modello è in grado di accompagnare la transizione dei dati dei cataloghi in formato MARC<sup>21</sup> verso le tecnologie dei linked open data, con l'obiettivo di integrare i cataloghi bibliotecari nel web semantico.

Una delle criticità dal punto di vista operativo nell'ipotesi di applicazione dell'ontologia BIBFRAME ai beni librari esposti su BeWeB riguarda la classe *work*. Nel Servizio bibliotecario nazionale (SBN), di cui BeWeB espone il sottoinsieme di record bibliografici catalogati dalle biblioteche di proprietà ecclesiastica<sup>22</sup>, ci sono pochi record corredati da un punto di accesso autorizzato al titolo dell'opera (*work*): i record bibliografici sono carenti di registrazioni di autorità per i titoli dell'opera e questo comporta una riduzione delle potenzialità di questo modello. Si tratta di una riduzione delle potenzialità, ma non di impossibilità: andrebbero messe in campo delle azioni di bonifica massiva e di cura del catalogo per far emergere questi concetti in modo da rendere appieno il potenziale del modello che, peraltro, non fa altro che attingere, nella traduzione in BIBFRAME, al modello FRBR preesistente, sebbene con qualche semplificazione (ad esempio il livello *espressione* di FRBR è sfumato in BIBFRAME). Oltre al grappolo dei nomi di persona/famiglia/ente (e delle loro forme varianti) rappresentati come *Agenti* a diverso titolo nei vari settori disciplinari in BeWeB (sistema di clustering già pienamente operativo come dispositivo di raccordo nell'AF cross-domain di BeWeB<sup>23</sup>),

versione *Records in Contexts-Conceptual Model* (RiC-CM), RiC-CM 0.2, pubblicata successivamente, a luglio 2021, e disponibile all'indirizzo: <<https://www.ica.org/en/records-in-contexts-conceptual-model>>.

<sup>20</sup> In Bibframe esistono poi altre classi aggiuntive, molte delle quali sono sottoclassi delle classi principali.

<sup>21</sup> Le specifiche di conversione da MARC a BIBFRAME sono gestite dalla Library of Congress e disponibili qui: <<https://www.loc.gov/bibframe/mtbf/>>.

<sup>22</sup> Cioè da tutte le biblioteche di proprietà ecclesiastica aderenti al SBN, indipendentemente dal polo bibliotecario di appartenenza (sia aderenti al polo di biblioteche ecclesiastiche (PBE) che aderenti ad altri poli).

<sup>23</sup> Per maggiori informazioni sul sistema di gestione authority data cross-domain in BeWeB: Francesca M. D'Agnelli – Claudia Guerrieri – Maria Teresa Rizzo – Silvia Tichetti, *L'autorità work nel sistema*

l'adozione di questo modello per i beni librari consentirebbe la generazione di un altro punto di accesso autorizzato alle risorse, coincidente con il livello di *work*: si formerebbe dunque allo stesso modo un grappolo di titoli (e forme varianti) che rimanderebbero alla medesima classe di *work* (opera). Ciascuna istanza (o pubblicazione) sarebbe quindi collegata al relativo *work*, e la disponibilità della copia sarebbe presente a livello di *item* con la possibilità di fruire dei servizi dell'OPAC di BeWeB (ad esempio, i servizi di prenotazione per la consultazione e il prestito riservati agli utenti registrati al portale). La potenzialità di questo modello (Fig. 2) naturalmente si estende in un portale trasversale come BeWeB, che aggrega beni culturali di diversa natura: alla stessa opera, infatti, sarebbe possibile collegare un bene archivistico in cui se ne parla, un bene storico e artistico da cui trae ispirazione, un bene architettonico che ne rappresenta il soggetto cioè l'argomento dell'opera, e così via.

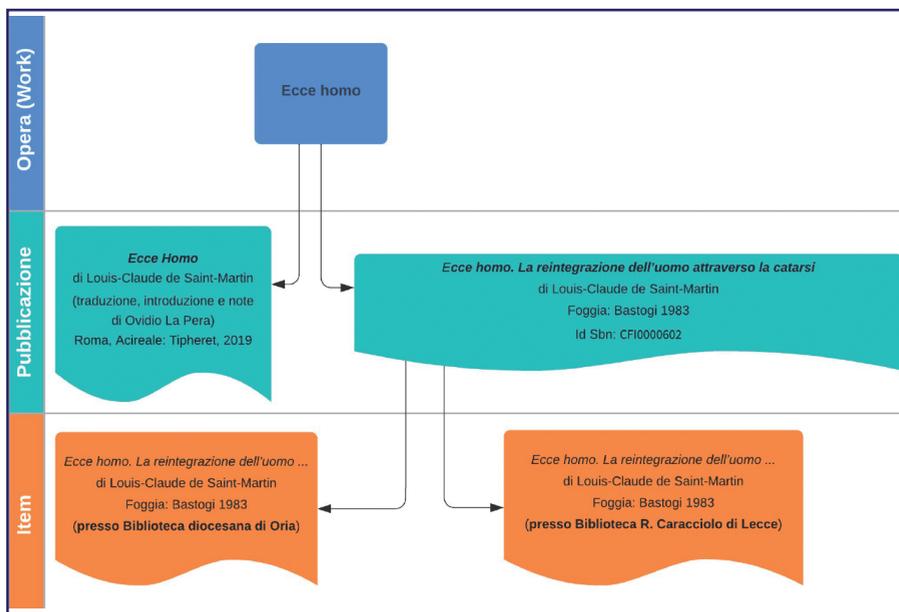


Figura 2. Livelli Work-Instance-Item nel modello BibFrame

Per quanto riguarda i beni archivistici<sup>24</sup>, l'ipotesi è quella di valutare un'applicazione sperimentale dell'ontologia RiC-O per la modellazione dei dati delle schede de-

*dei beni culturali ecclesiastici*, «DigItalia. Rivista del digitale nei beni culturali», 16 (2021), n. 2, p. 91-107; *Gli authority data per l'integrazione cross-domain dei beni culturali: riflessioni su un approccio alla lettura trasversale dei beni culturali della Chiesa cattolica italiana*, a cura del Gruppo di lavoro sugli authority file dell'Ufficio per i beni culturali ecclesiastici e l'edilizia di culto della CEI, «AIB studi», 57 (2017), n.1, p. 106-108.

<sup>24</sup> Le risorse archivistiche presenti in BeWeB sono descritte seguendo gli standard internazionali ISAD-G (*General International Standard Archival Description*) e ISAAR-CPF (*International Standard*

scrittive dei fondi archivistici pubblicate su BeWeB<sup>25</sup>. In questo contributo non verranno approfonditi i dettagli dell'ontologia RiC-O, che traduce nelle sue relazioni principali il modello concettuale RiC-CM<sup>26</sup>.

Sin qui abbiamo dunque ipotizzato quali potrebbero essere le ontologie riusabili per modellare gli istituti culturali (Cultural-ON), i beni storico-artistici e architettonici (ArCo, che consentirebbe di modellare anche i dati descrittivi di beni fotografici e immateriali, una volta pubblicati su BeWeB), e i beni archivistici e librari (facendo sommariamente riferimento a Bibframe e RiC-O).

Successivamente alla fase di individuazione delle ontologie per la modellazione dei dati, le buone pratiche per la produzione di LOD proposte dal W3C indicano altri due passi fondamentali nello sviluppo di un progetto di questo tipo: l'utilizzo di vocabolari standard (fase 6) e il processo di conversione dei dati in RDF<sup>27</sup> (fase 7).

### L'utilizzo di vocabolari standard

Nell'ipotesi di progetto di produzione LOD del patrimonio esposto su BeWeB si propone, relativamente alla buona pratica dell'utilizzo di vocabolari standard, di utilizzare SKOS (Simple Knowledge Organization Systems)<sup>28</sup> per favorire l'intero-

*Archival Authority Records—Corporate Bodies, Persons, and Families*), che di fatto sono alla base dell'elaborazione del modello concettuale RiC-CM insieme agli altri due standard internazionali per la descrizione delle funzioni (ISDF) e dei soggetti conservatori (ISDIAH). Per approfondire: <<https://www.ica.org/en/archival-arrangement-and-description>>.

<sup>25</sup> Ad oggi BeWeB considera come risorse archivistiche principali, tra i risultati di ricerca, esclusivamente le schede relative al livello di descrizione "fondo" e non espone direttamente le altre tipologie di livello di descrizione (sezioni, serie, unità archivistiche ecc.) ma si configura come principale punto di accesso ai livelli descrittivi gerarchici inferiori: questi ultimi, infatti, sono navigabili, per mezzo della funzione "esplora il fondo", qualora ne sia stata rilasciata liberatoria dagli enti proprietari dei beni. Questa funzione consente la classica visualizzazione ad albero rovesciato del patrimonio archivistico descritto: è in corso di sviluppo la progettazione di una restituzione dei dati dei livelli descrittivi inferiori al fondo, in modo da renderli singoli oggetti (tipizzati a seconda del livello di descrizione di riferimento), per renderli direttamente ricercabili come schede descrittive da parte dell'utente web, che potrà essere condotto direttamente sul livello descrittivo (fondo, sezione, serie, sottoserie, unità ecc.) in cui compare il risultato della sua ricerca. Attualmente invece l'utente viene a conoscenza della presenza di risultati compatibili con la sua ricerca a livello di fondo archivistico, e a partire da questo livello di contesto può esplorare i livelli inferiori di appartenenza.

<sup>26</sup> Per approfondire: Florence Clavaud – Tobias Wildi, *ICA Records in Contexts-Ontology (RiC-O): a semantic framework for describing archival resources*, «Ceur Workshop Proceedings», 3019, settembre 2021, <<http://ceur-ws.org/Vol-3019/>>; *Descrivere gli archivi al tempo di RiC-CM*, a cura di G. Di Marcantonio – F. Valacchi, EUM Edizioni Università di Macerata, 2018.

<sup>27</sup> RDF (*Resource Description Framework*) è una raccomandazione del W3C dal 2004: il modello è stato ideato per essere facilmente generalizzabile e poter rappresentare informazioni riguardanti entità opportunamente identificabili sul web (risorse e relazioni che tra di esse intercorrono), rendendo in tal modo effettivamente praticabile la descrizione di qualunque cosa.

<sup>28</sup> Maggiori informazioni all'indirizzo: <<https://www.w3.org/TR/skos-primer/>>. Questa raccomandazione prevede sia la conversione dei sistemi esistenti in SKOS (ne è un esempio il soggetto della Library of Congress che ora è disponibile anche nella sua versione SKOS), sia la definizione di nuovi

perabilità semantica. La raccomandazione SKOS del W3C è finalizzata proprio alla costruzione di sistemi di organizzazione della conoscenza. Dall'analisi dei dati esposti su BeWeB emerge che alcuni concetti, come ad esempio quello di ente proprietario, richiedono l'adozione di una più puntuale classificazione. Se si applica questa proposta all'esempio concreto di una scheda descrittiva di un istituto culturale, la proprietà/relazione "è proprietario di" è quella che collega ad esempio una classe "istituto culturale" a una classe "ente". Se nei sistemi informativi del Ministero della Cultura italiano gli istituti culturali di proprietà ecclesiastica (presenti in BeWeB) confluiscono tutti nella macrocategoria "ente/istituzione di culto cattolico", per BeWeB è fondamentale distinguere la tipologia di ente di culto cattolico. Il vocabolario controllato al momento disponibile distingue già tra le seguenti tipologie di enti: Diocesi, Parrocchia, Ordine religioso, Confraternita, Capitolo canonico, Università/Istituti superiori di scienze religiose, Associazione di fedeli, Confraternita, Fabbriceria. Se per gli enti Diocesi e Parrocchia il sistema di gestione degli authority data attribuisce per ogni voce un identificativo univoco ed è scientificamente valido<sup>29</sup>, sarebbe opportuno codificare anche il vocabolario presente per le altre tipologie di enti. Sarebbe utile l'avvio di un lavoro di produzione di un vocabolario controllato in SKOS, con l'obiettivo di individuare la gerarchia degli enti ecclesiastici che appartengono alla categoria Istituti di vita consacrata e Società di vita apostolica, al fine di ottenere una classificazione più analitica e supportata da una rappresentazione utilizzabile nel contesto del web semantico. Il vantaggio in termini di LOD è quello di poter identificare univocamente – ad esempio – un ordine religioso e tutte le sue emanazioni amministrative sul territorio (come la curia generale o la provincia): avere la possibilità di identificare un ordine religioso e, come struttura organizzativa gerarchicamente inferiore, una determinata provincia sul territorio italiano (identificata univocamente e inserita all'interno di un sistema di classificazione coerente) consentirebbe di poter collegare tutti gli istituti culturali disseminati sul territorio che condividono lo stesso ente proprietario<sup>30</sup>. Concretamente, la costruzione del sistema di classificazione dovrebbe partire – per gli istituti di vita consacrata e società di vita apostolica – dall'analisi di quanto già presente online, verificando se esistono già agenzie autorevoli che hanno prodotto vocabolari controllati in SKOS per queste tipologie di ente (ad esempio, verificando quanto già realizzato dalla Library of Congress o quanto

sistemi di organizzazione della conoscenza in questo linguaggio, basato su RDF, considerato come una tecnologia intermedia tra l'elevato formalismo logico dei linguaggi OWL per la costruzione di ontologie e l'insufficiente (quasi assente) strutturazione semantica delle informazioni sui siti web.

<sup>29</sup> Attualmente il paragrafo "ente proprietario" ha una gestione di recupero codificata univocamente per enti di tipo Diocesi (226 enti) e Parrocchia (circa 26.000 enti).

<sup>30</sup> Ad esempio la Provincia di Sant'Antonio dei Frati minori è un'articolazione dell'Ordo Fratrum Minorum (Ordine dei Frati minori: <<https://ofm.org/about/provinces-custodies/>>) ed è proprietaria di diversi istituti culturali presenti in BeWeB.

siano utilizzabili i Getty vocabularies<sup>31</sup>, il VIAF o altri repository). Qualora i vocabolari già presenti non siano soddisfacenti andrebbe avviata una fase di studio di fonti autorevoli, come ad esempio l'Annuario Pontificio e il secondo volume di Acolit<sup>32</sup> relativo agli ordini religiosi: i dati raccolti, verificati e arricchiti all'occorrenza, dovrebbero essere caricati in una struttura che rispetti i formalismi di SKOS, e le istanze da collegare corrisponderebbero a un authority file della classe "ente", sottoclassi "istituti di vita consacrata" e "società di vita apostolica"<sup>33</sup>. In questo modo ogni istanza potrebbe beneficiare, una volta identificata univocamente nel sistema CEI-AF, del trattamento tramite OLAF<sup>34</sup> per l'interlinking con altre banche dati esterne (come Wikidata e VIAF). La presenza di queste schede Ente all'interno dell'AF cross-domain, in una forma standardizzata e controllata, garantirebbe il loro utilizzo e recupero anche in tutte le banche dati (in back-end) che sono utilizzate quotidianamente dagli operatori coinvolti nei progetti di descrizione e catalogazione dei beni culturali di proprietà ecclesiastica. La ripercussione positiva di questo lavoro potrebbe essere quella di poter esplorare tutti gli archivi, biblioteche e musei (e i relativi beni culturali in essi conservati) di proprietà di un determinato istituto di vita consacrata/società di vita apostolica, con la potenzialità espressiva di poterli navigare su mappa, strumento utile anche per la didattica oltre che per la ricerca. Tale vocabolario controllato, che potrebbe popolare direttamente l'AF CEI cross-domain, garantirebbe l'allineamento tra i servizi, la circolarità delle informazioni e l'univocità (senza duplicazioni) dei record di autorità di tipo "ente" appartenenti alle sottoclassi istituti di vita consacrata e società di vita apostolica<sup>35</sup>. Il recupero e il trattamen-

<sup>31</sup> I *Getty vocabularies* formalizzano concetti espressi in formato semantico: ne sono esempi l'*Art and architecture thesaurus* (AAT), che contiene una terminologia strutturata per l'archeologia, l'arte, l'architettura, le arti decorative, e il *Thesaurus of geographic names* (TGN), che individua i nomi geografici, il *Cultural Objects Name Authority* (CONA), che contiene titoli, attribuzioni, soggetti raffigurati e altri metadati su opere d'arte, architettura e patrimonio culturale, e *Getty Iconography Authority* (IA), un thesaurus che copre argomenti rilevanti per l'arte, l'architettura e le discipline correlate (ad esempio Adorazione dei Magi, Diluvio universale, Rivoluzione francese). Nel *Getty vocabulary semantic representation* è disponibile una descrizione del set completo di URI disponibili: <<https://www.getty.edu/research/tools/vocabularies/lod/#uris>>.

<sup>32</sup> Sebbene Acolit risulti datato: *Acolit. Autori cattolici e opere liturgiche: Ordini religiosi*, a cura dell'Associazione Bibliotecari Ecclesiastici Italiani, v. 2, Milano: Editrice bibliografica, 2000.

<sup>33</sup> Questa classificazione è prevista dal Codex Iuris Canonici nel libro II, parte III: "Gli istituti di vita consacrata e le società di vita apostolica". Per approfondire: <[https://www.vatican.va/archive/cod-iuris-canonici/ita/documents/cic\\_libroII\\_573-606\\_it.html](https://www.vatican.va/archive/cod-iuris-canonici/ita/documents/cic_libroII_573-606_it.html)>.

<sup>34</sup> OLAF (Open Linked Authority File) è uno strumento di riconciliazione semiautomatica (con validazione da parte del redattore centrale) tra i record CEI-AF e quelli provenienti da banche dati esterne collegate Linked Open Data Cloud. Per approfondire: <<https://synapta.it/blog/olaf-per-i-beni-culturali-un-passo-verso-lautomatizzazione-dei-collegamenti-tra-le-fonti/>>.

<sup>35</sup> L'allineamento con la banca dati di Authority file cross-domain e con l'Anagrafe degli istituti culturali ecclesiastici (AICE) agevolerebbe il controllo delle informazioni relative alla gerarchia istituzionale di istituti di vita consacrata e società di vita apostolica sul territorio e consentirebbe il recupero univoco dell'intestazione dell'ente (senza nuove immissioni di intestazioni foriere di errori).

to a livello centrale, in un sistema di classificazione validato centralmente, garantirebbe una gerarchia degli enti sistematica e coerente, senza duplicazioni, riutilizzabile direttamente sia per scopi di gestione all'interno dei sistemi di back-end sia per favorire la consultazione sul front-end di BeWeB, potendo inoltre allargare ancora di più le potenzialità descrittive grazie agli allineamenti con sistemi esterni di identificazione degli authority file<sup>36</sup>.

Nell'esempio seguente (Fig. 3) è possibile visualizzare, partendo dal concetto di "agente", un tentativo di organizzazione di un sistema di classificazione degli istituti di vita consacrata e società di vita apostolica (sulla base di SKOS, nella prospettiva di un suo utilizzo per un progetto di produzione LOD).

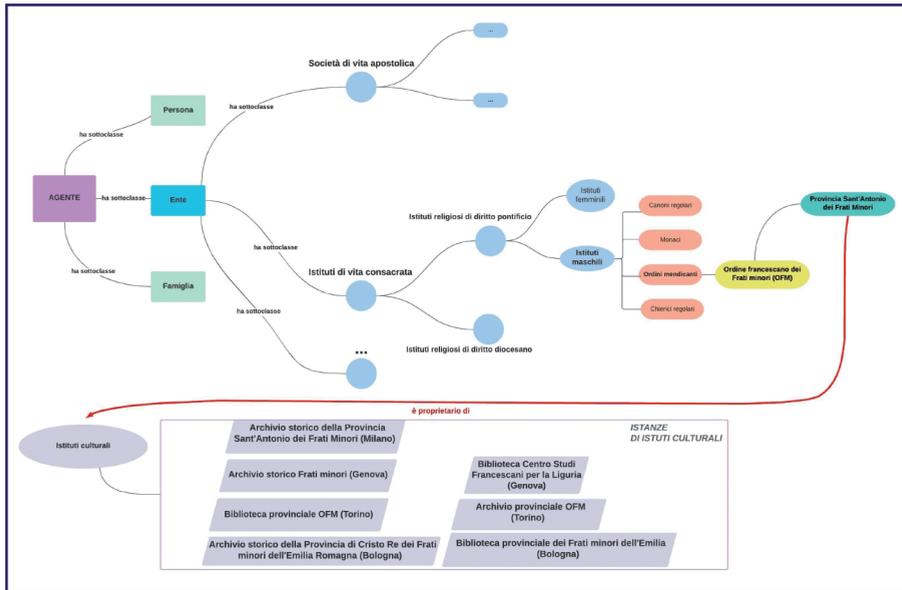


Figura 3. Ipotesi di classificazione istituti di vita consacrata/società di vita apostolica: la Provincia Sant'Antonio dei Frati minori (identificata univocamente) intrattiene la relazione "è proprietario di" con n istituti culturali ecclesiastici sul territorio italiano.

### Aboutness: la questione dell'identificazione dei soggetti

La necessità dell'utilizzo di vocabolari controllati è ancora più cruciale se si pensa all'identificazione dei soggetti rappresentati dalle risorse culturali: non sempre i progetti di catalogazione e descrizione del patrimonio hanno rispettato rigide regole di composizione della stringa di soggetto o adottato sistemi di classificazione

<sup>36</sup> Alcuni esempi di identificazione dell'Ordine dei Frati minori in altri repository: <<https://id.loc.gov/authorities/names/n79015671.html>>; <<https://viaf.org/viaf/147415964/#skos:Concept>>; <<https://www.wikidata.org/wiki/Q913972>>; <<https://www.wikidata.org/wiki/Q165005>>.

codificata come quelle proposte da Iconclass (IC)<sup>37</sup> e dall'Art & Architecture Thesaurus (AAT)<sup>38</sup>, e pertanto non è sempre agevole comprendere quale sia il soggetto di una determinata risorsa, con la conseguenza negativa di non poter rilevare i significati veicolati da un bene culturale e non poter fare ricerche – anche cross-domain – su un determinato soggetto rappresentato.

Confrontando questi due schemi di classificazione del dominio storico-artistico, IC e AAT, ci si rende conto che trovare una corrispondenza tra i soggetti di IC e i concetti del AAT sia un compito molto impegnativo: si tratta infatti di schemi molto diversi, in cui la struttura gerarchica delle voci non trova corrispondenza tra uno schema e l'altro rendendo difficili eventuali relazioni di equivalenza tra medesimi concetti. I soggetti usati in IC possono essere visti come una descrizione più generale del concetto corrispondente in AAT: si potrebbe quindi collegare un soggetto IC al concetto AAT più particolare per aggiungere uno strato di semantica. Inoltre, andrebbero disambiguati i significati delle voci usate poiché esse appaiono più di una volta negli schemi classificatori con significati diversi: per questi motivi questi strumenti ancora sono in corso di evoluzione per rispondere pienamente alle esigenze del semantic web.

Le domande più frequenti poste al dominio culturale riguardano proprio questa dimensione, necessaria per comprendere “cosa è raffigurato” o “di cosa parla” una determinata risorsa (la cosiddetta *aboutness*), attività per la quale è necessaria la mediazione di un professionista dei beni culturali che dovrebbe sempre tenere presente il suo fondamentale contributo alla diffusione della conoscenza durante le attività di descrizione e catalogazione. Se non si rilevano adeguatamente in fase di input del dato, cioè durante la schedatura, quali siano i soggetti rappresentati oppure non si struttura correttamente l'informazione mediante una sintassi standard, non è possibile interpretare correttamente la risorsa consultata. Gli aspetti

<sup>37</sup> Iconclass (IC) è un sistema di classificazione progettato per l'arte e l'iconografia. È stato ideato da Henri van de Waal, professore di storia dell'arte all'Università di Leiden, sin dagli anni Cinquanta; completato negli anni successivi al 1972 da un folto gruppo di studiosi e pubblicato tra il 1973 e il 1985 dalla Royal Netherlands Academy of Arts and Sciences (KNAW). Dal 1990 al 2001 si sono succedute diverse edizioni a cura dell'Università di Utrecht. È lo strumento scientifico più diffuso per la descrizione e il recupero di soggetti rappresentati in immagini (opere d'arte, illustrazioni di libri, riproduzioni, fotografie ecc.) ed è utilizzato da musei e istituzioni artistiche di tutto il mondo. Il sistema Iconclass è disponibile come Linked Open Data (LOD) e accessibile tramite il browser Iconclass: <<http://www.iconclass.org/help/outline>>. Maggiori informazioni all'indirizzo: <<https://rkd.nl/nl/collecties/services-tools/iconclass>>.

<sup>38</sup> L'AAT è stato originariamente fondato alla fine degli anni Settanta dai direttori di biblioteche ed esperti di architettura Toni Petersen, Dora Crouch e Pat Molholt, in risposta a un'esigenza di catalogazione di pubblicazioni e riviste d'arte. Il lavoro editoriale è gestito dal 1983 dal Getty Research Institute. Nel 1987 il Getty ha creato un dipartimento dedicato alla compilazione e alla distribuzione della terminologia, ora noto come Getty Vocabulary Program.

Il Getty Vocabularies ha pubblicato l'AAT come LOD:

<<https://www.getty.edu/research/tools/vocabularies/obtain/download.html>>. Per approfondimenti: <<https://www.getty.edu/research/tools/vocabularies/aat/about.html>>.

legati alla lettura simbolica, teologica, devozionale e pastorale di un bene culturale ecclesiastico emergeranno con difficoltà se non si opera un'attività di analisi, bonifica e pulizia su queste tipologie di dato (come il soggetto, appunto), già rilevate in questi anni, nel corso delle attività di catalogazione del patrimonio ecclesiastico iniziate nel 1996. Il soggetto non è ovviamente il solo campo in cui si inseriscono dati di rilievo per questo tipo di lettura del patrimonio, ma un'attività di *data cleaning* giova in ogni caso all'emersione dei significati, oltre che al potenziamento delle relazioni con alcuni concetti da classificare, come quello di rito, sacramento ed evento religioso, o con le partizioni temporali che durante l'anno hanno precisi significati per il culto cattolico.

## Il processo di conversione del dataset in RDF

L'ultimo aspetto operativo da affrontare, come anticipato, è quello relativo all'operazione di conversione del dataset selezionato in RDF: durante la fase 7, un set di dati viene tradotto in una rappresentazione di dati collegati, generalmente tramite script o altri processi automatizzati<sup>39</sup>. Indipendentemente dall'approccio utilizzato, la conversione dei dati implica la mappatura dei dati di origine in un insieme di descrizioni RDF che possono poi essere serializzate in diversi formati.

I dati esposti su BeWeB risiedono al momento in un database relazionale (RDBMS). I dati presenti nel database relazionale andrebbero strutturati mediante un sistema di mappatura basato sul linguaggio RML<sup>40</sup> (cioè tramite una modellizzazione adatta a rappresentare gli oggetti nel linguaggio RML). Successivamente, con l'adozione di uno strumento come ad esempio SDM-RDFizer<sup>41</sup>, si potrebbe procedere all'elaborazione dei dati di origine<sup>42</sup> presenti nel database relazionale di BeWeB, in modo da poter produrre la serializzazione RDF dei dati nel formato desiderato (ad esempio, JSON-LD)<sup>43</sup>. RDFizer è un interprete di regole di mappatura che consente la trasformazione di dati (anche non) strutturati in grafi RDF. L'attuale versione di RDFizer presuppone che le regole di mappatura siano definite nel linguaggio di mappatura RDF (RML). Sarebbe possibile quindi partire dal database attuale per produrre il dataset in formato JSON (in formato standard RML)

<sup>39</sup> Ad esempio tramite linguaggi di mappatura dichiarativa, o linguaggi come R2RML che eseguono la traduzione dei dati (attraverso mappature personalizzate) da database relazionali a dataset RDF. Maggiori informazioni all'indirizzo: <<https://www.w3.org/ns/r2rml>>.

<sup>40</sup> Maggiori informazioni all'indirizzo: <<https://rml.io/specs/rml/>>.

<sup>41</sup> Enrique Iglesias – Samaneh Jozashoori – David Chaves-Fraga – Diego Collarana – Maria-Esther Vidal, *SDM-RDFizer: An RML Interpreter for the Efficient Creation of RDF Knowledge*, «CIKM '20: The 29th ACM International Conference on Information and Knowledge Management», ACM, ottobre 2020, p. 3039–3046, <<https://arxiv.org/abs/2008.07176>>.

<sup>42</sup> Maggiori informazioni all'indirizzo: <<https://rml.io/specs/rml/#logical-source>>.

<sup>43</sup> Maggiori informazioni all'indirizzo: <<https://github.com/SDM-TIB/SDM-RDFizer>>. Il tool RDFizer permette la conversione di dati in RDF da origini dati (*source*) eterogenee (CSV, JSON, RDB, XML). Anche l'ICCD nel progetto di produzione LOD per il sito del Catalogo dei beni culturali ha adottato il tool RDFizer: in quel caso però i dati di origine erano in formato XML.

e convertire poi JSON tramite il tool SDM-RDFizer, ottenendo la serializzazione di RDF nel formato JSON-LD.

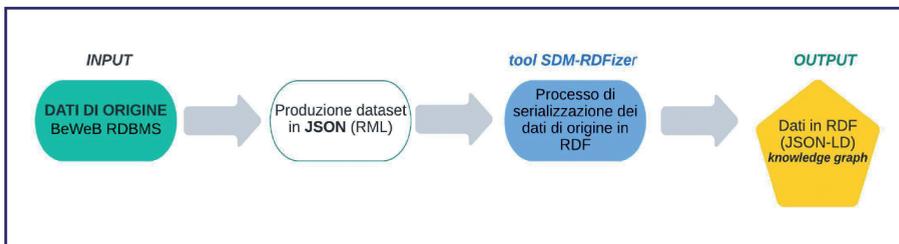


Figura 4. Fasi operative del processo di serializzazione dei dati di origine in RDF

Un’attenzione particolare andrebbe poi riservata al problema della conservazione e manutenzione nel tempo dei linked open data<sup>44</sup> pubblicati (anche al fine di consentirne il riuso): negli ambienti digitali, infatti, occorre progettare e attuare sin dalla fase di creazione e gestione degli oggetti digitali il processo a garanzia della tenuta e del mantenimento delle relazioni, dei metadati, dei dati e della loro corretta interpretazione. In questo senso, dunque, la dimensione gestionale e quella conservativa si configurano come un *continuum*, e andranno affrontate simultaneamente.

## Conclusioni e prossimi obiettivi

In sintesi, i passi operativi su cui abbiamo condotto un approfondimento al fine di delineare una prima valutazione per un’ipotesi di progetto sperimentale di produzione LOD per il portale BeWeB sono:

- *selezione di un set di dati*: l’ipotesi è partire dal dataset degli istituti culturali (musei, archivi, biblioteche conservatori dei beni culturali);
- *modellazione dei dati*: l’ipotesi è di riusare le ontologie
  - Cultural-ON per gli istituti culturali;
  - ArCo per la modellazione dei beni storici e artistici (e a breve fotografici) e dei beni architettonici. In potenza ArCo potrebbe essere usato anche per la modellazione dei beni immateriali e dei beni musicali;
  - BIBFRAME per la modellazione dei beni librari;
  - RiC-O per la modellazione dei beni archivistici.

Dall’analisi delle ontologie già esistenti per i beni culturali è emerso che il patrimonio culturale ecclesiastico può essere modellato riusando una buona parte di ontologie disponibili. Le peculiarità del patrimonio culturale ecclesiastico potrebbero però essere maggiormente valorizzate con l’adozione (o meglio con la costruzione) di vocabolari controllati, piuttosto che con la predisposizione di nuovi modelli ontologici.

<sup>44</sup> Margherita Bartoli – Fabiana Guernaccini – Giovanni Michetti, *Preservation of Linked Open Data*, «JLIS.it», 11 (2020), n. 2, p. 20-44: <<https://www.jlis.it/index.php/jlis/article/view/43>>.

- *adozione di vocabolari standard*: l'ipotesi è di usare SKOS per costruire dei sistemi di classificazione di concetti intrinsecamente legati al patrimonio culturale ecclesiastico. In particolare, potrebbe essere utile creare un vocabolario controllato per tutti gli istituti di vita consacrata/società di vita apostolica che sono proprietari degli istituti culturali conservatori di beni culturali: in questo modo si potrebbe ricostruire fedelmente la mappa di questi enti, che hanno competenza su diversi istituti culturali sul territorio a seconda della provincia religiosa di appartenenza (che non trova corrispondenza con i livelli di geografia amministrativa ecclesiastica "diocesi" e "regione ecclesiastica"<sup>45</sup>). L'utilizzo di SKOS potrebbe aumentare le potenzialità espressive dei campi di tipo *Soggetto* che finora non si sono avvalsi dell'uso di vocabolari controllati e di precise regole sintattiche nella costruzione delle stringhe di soggetto, che contribuirebbero invece a una migliore comprensione di simboli e significati sottesi alle descrizioni dei beni culturali ecclesiastici, potenziando la lettura pastorale, teologica e devozionale per questa peculiare tipologia di risorsa;
- *conversione dei dati*: l'ipotesi è di partire dal sistema RDBMS di BeWeB e adottare un tool di conversione come RDFizer per serializzare i dati in JSON-LD.

La conversione dei dati in RDF non è però sufficiente per raggiungere l'obiettivo della piena interoperabilità semantica e di un riuso efficiente dei dati da parte di altre applicazioni informatiche: per produrre e pubblicare linked open data di qualità attraverso un processo omogeneo e sostenibile nel tempo, è necessario uno sforzo progettuale preliminare, definendo tutto il processo di produzione con una chiara governance. Produrre linked open data non significa esclusivamente pubblicare dataset aperti del proprio patrimonio informativo, ma significa ripensare e riprogettare le modalità di produzione e governo dei dati dal momento della loro produzione. Devono essere individuate precise responsabilità, e tutte le fasi di produzione devono essere integrate continuamente, poiché si devono prevedere attività relative a un costante aggiornamento, monitoraggio e coinvolgimento degli utenti finali dei dataset messi a disposizione per il riuso.

Come osservazione di carattere generale, sembra evidente che il punto di partenza per esprimere il massimo potenziale di un qualsiasi progetto di produzione LOD risieda nella qualità dei dati: l'investimento nel *data cleaning* (identificazione, analisi, pulizia e bonifica dei dati) assume quindi un valore prioritario, o meglio propedeutico, per l'avvio di un progetto di produzione di LOD, poiché soltanto dati di qualità possono consentire di rispondere coerentemente alle domande del futuro. Attribuire significati *ex post* è infatti un'operazione foriera di errori che non per-

<sup>45</sup> La geografia ecclesiastica italiana prevede 16 regioni ecclesiastiche, suddivise in 226 diocesi, ognuna delle quali vede il proprio territorio suddiviso in parrocchie (in totale circa 26.000). Tra il livello diocesi e il livello parrocchia è possibile che ci siano altri livelli amministrativi ecclesiastici intermedi, come ad esempio zone pastorali, vicariati/decanati e unità pastorali. Per la denominazione ufficiale di questi livelli amministrativi intermedi si fa riferimento all'annuario diocesano.

mette di sfruttare al meglio le opportunità dei LOD: l'attenzione verso la semantica dovrebbe iniziare dalle modalità di inserimento dei dati di origine nei sistemi di catalogazione, per far emergere importanti relazioni con alcuni importanti concetti che, nell'ambito del patrimonio ecclesiastico, potrebbero essere ad esempio quello di rito, sacramento ed evento religioso, o con le partizioni temporali che durante l'anno hanno precisi significati per il culto cattolico. Anche i tipi di strumento e gli oggetti utilizzati in determinati momenti delle celebrazioni religiose veicolano un significato preciso che non può essere comunicato attribuendo esclusivamente un nome all'oggetto descritto. Non c'è alcun automatismo che possa consentire di attribuire questo genere di significati ai dati inseriti, nell'ottica di veicolare non soltanto il contesto d'uso degli oggetti culturali, non soltanto l'identità di una comunità o di un territorio trasmesse da una risorsa culturale, ma più profondamente quell'insieme valoriale e di senso che non è affatto scontato far emergere attraverso una scheda descrittiva di catalogo, poiché spesso celato da strati simbolici sovrapposti e giustapposti. È quindi sin nella fase di descrizione del bene che dovrebbero essere individuate le relazioni con i contesti, in modo da rendere possibile in seguito, attraverso i LOD, l'interrogazione simultanea di dati provenienti da silos diversi, ma collegati semanticamente.

Ancora una volta l'attività umana si riconferma essere la prima e l'unica in grado di poter conferire significati in maniera puntuale e profonda, per consentire anche a una macchina di poter interpretare un flusso di bit.

*Many cultural heritage institutions like libraries, archives and museums (LAMs) have been applying semantic web technologies and are beginning to adopt linked open data (LOD) as a way to organize and disseminate the catalogues describing their holdings: the challenge is to represent hidden knowledge in data, making it immediately understandable to machines. Through the delicate phase of unambiguous definition of concepts and relationships representing a certain portion of knowledge (data modelling), the semantic web acts as a tool that brings out meanings and produce new mutual relationships.*

*This paper aims at exploring the possibilities offered by LOD to enhance the visibility of ecclesiastical cultural heritage data presented on BeWeB, a cross-domain web portal dedicated to the cultural and artistic heritage of the Italian Catholic Church, by providing an integrated view of data regarding churches, artwork and items of worship they contain, library and archive holdings of diocesan and religious institutes, with a special focus on cultural, scientific and pastoral care aspects, in order to supply an authoritative tool to know more about these extraordinary assets.*

*The digital inventory of ecclesiastical cultural and artistic heritage, that the Italian Bishops' Conference (CEI) urged Italian dioceses – with the support of the National Office for Ecclesiastical Cultural Heritage (BCE National Office) acting as coordinator – to implement since 1997, has led in the last twenty years to the creation of a massive digital data bank. The inventory documents an extremely important, sizeable and peculiar cultural heritage, and testifies to the sense of responsibility and spirit of service of dioceses and religious institutes.*

*BeWeB offers an overview of the ecclesiastical cultural heritage preserved in Italy, where the authority data becomes a hub to weave relationships between cultural resources of different nature (historical-artistic, architectural, bibliographic, archival, photographic, cultural institutes).*

*The project aims to achieve full semantic interoperability, through the analysis of the already existing ontologies in the cultural domain, to encourage their reuse.*

*BeWeB is therefore a highly representative case study, with precious insights about the process behind the elaboration of knowledge graph projects and the production of LOD in general.*