

Dalla somiglianza alla sinergia. La descrizione del patrimonio culturale dalla specificità all'integrazione digitale

Laura Moro

Intervento del mattino

Vorrei introdurre la discussione della tavola rotonda con tre blocchi di riflessioni: da dove veniamo (i fondamenti dei sistemi descrittivi degli standard catalografici dell'ICCD), quale lo stato attuale, quali le prospettive per l'integrazione digitale.

I fondamenti metodologici degli standard catalografici dell'ICCD

A rischio di dire cose già note, vorrei ricordare quali sono le tipologie di beni culturali che gli standard catalografici descrivono: il catalogo nasce alla fine dell'ottocento per scavi, monumenti e gallerie (quindi monumenti e reperti archeologici. Architetture, opere d'arte); poi progressivamente si amplia ad altre tipologie di oggetti culturali, con l'allargamento del concetto stesso di patrimonio culturale: fotografie e fondi fotografici; beni demontantropologici materiali e immateriali; beni naturalistici e patrimonio scientifico e tecnologico per le collezioni dei musei scientifici; paesaggio e settori urbani, opere d'arte contemporanee.

Si tratta di oggetti culturali che non trovano nella scrittura il linguaggio di espressione, come avviene invece per i beni archivistici e bibliografici: I sistemi descrittivi che vengono utilizzati devono quindi farsi carico di mediare linguaggi visivi o l'assenza totale di qualsiasi linguaggio.

A questo si aggiunge che fin dalle origini il catalogo, con il suo corredo metodologico, aveva una duplice funzione:

- attestare la valenza culturale del bene (è l'essere nel catalogo che trasforma una "cosa" in un bene culturale), soprattutto per i beni non raccolti in collezioni, e quindi definire il perimetro dell'azione di tutela, nonché la rappresentazione stessa di patrimonio culturale nazionale;
- fornire una guida scientifica alla consultazione/fruizione del bene, soprattutto nel caso dei beni musealizzati.

Per assolvere a queste funzioni era necessario che il catalogo superasse la funzione di mero inventario, per contenere elementi di natura storico critica, tanto nelle descrizioni dei singoli beni che nelle ricostruzioni dei contesti di provenienza.

Con l'avvento della catalogazione informatizzata, questo sistema di conoscenza si è trasformato in un insieme articolato di standard catalografici con un livello molto granulare di organizzazione dei dati, spaziando dalle informazioni sulle caratteristiche materiali del bene alla lettura e interpretazione dei suoi valori iconografici, dal collocamento nel contesto di produzione, alla sua localizzazione e inventariazione; a questo si aggiunge un sistema altrettanto articolato di collegamento a tesauri e liste di autorità esterne.

Il limite di questi primi standard catalografici era senza dubbio la prevalenza dell'approccio storico-artistico, che tendeva a porre al centro della descrizione le caratteristiche materiali del bene e i contenuti visivi dell'opera; inoltre le varie tipologie di schede, derivate concettualmente dalla scheda dell'opera d'arte, sono cresciute come "gemmazione" creando un sistema poco coerente nel suo complesso (in quanto basato appunto su presupposti metodologici quasi esclusivamente "storico artistici"). Per ultimo, la prevalenza del valore storico-critico dei contenuti ha fatto sì che si investissero meno risorse intellettuali nella definizione di tesauri controllati (cosa che invece è stato fatto per esempio nell'*Inventaire* francese), nella convinzione che una scheda di catalogo fosse un piccolo saggio di storia dell'arte, prediligendo così i vocabolari aperti - più difficili da controllare - e rinviando ad un momento successivo la complessa attività di omogeneizzazione terminologica, che ad oggi non è stato ancora possibile realizzare.

Lo stato attuale

Allo stato attuale il sistema descrittivo prodotto dall'ICCD si presenta solido e molto strutturato:

- ampio: 30 standard per altrettante tipologie di oggetti culturali.
- articolato: le varie normative nel loro insieme formano un sistema relazionato e gerarchico; ad esempio per il patrimonio archeologico si ha a disposizione una serie di standard che vanno dalla scheda di descrizione del sito archeologico, a quella del complesso archeologico, al monumento archeologico, al reperto archeologico, all'unità stratigrafica muraria fino ad arrivare alla tabella dei materiali di scavo. Per i

beni fotografici si passa dal fondo fotografico, alla serie fotografica, alla singola fotografia.

- molto analitico: ogni scheda è divisa in paragrafi, campi e sottocampi che sondano tutte le caratteristiche del bene. Sono schede molto ampie, centinaia tra campi e sottocampi (ovviamente in fase applicativa si può scegliere a che livello di analiticità procedere, quindi la presunta eccessiva complessità della scheda di catalogo – fatte salve le obbligatorietà previste dalle varie normative - appare più che altro come un alibi); le informazioni sono raccolte in modo disaggregato così che risulta possibile sia un controllo sintattico capillare in fase di *dataentry* che la mappatura semantica praticamente verso qualsiasi altro standard esistente (possiamo produrre, e di fatto produciamo, dati in qualsiasi formato).

- fortemente relazionato: esistono relazioni verticali (come quelle che intercorrono in un bene complesso, ad esempio un ciclo pittorico, fra l'insieme - l'intera sequenza figurativa - e le singole opere componenti) e orizzontali tra beni diversi (come nel caso di beni appartenenti ad una medesima collezione oppure ad un corredo funerario, o accomunati da uno specifico utilizzo), relazioni con authority file (autore, bibliografia, luogo di conservazione, eventi come le campagne di scavo e di ricognizione per l'ambito archeologico), relazioni con tesauri controllati (questo rappresenta uno dei fronti di miglioramento di cui dirò a breve).

- fortemente omogeneo: nelle schede di ultima generazione esiste una normativa trasversale comune a tutte le tipologie di beni (le schede condividono il paragrafo di definizione dell'oggetto, definizione culturale, relazioni, localizzazione, georeferenziazione, dati patrimoniali, cronologia, condizione giuridica, dati tecnici, conservazione, documentazione), che rende le schede coerenti tra loro, riducendo l'eterogeneità nell'impostazione di cui accennavo prima, e facilita le operazioni di mappatura verso altri standard.

Sottolineo qui la grande attenzione nelle schede di ultima generazione per le informazioni geografiche, sia alfanumeriche (toponimi controllati da uno specifico thesaurus stato/regione/provincia/comune/località che fa riferimento a liste ISTAT) che mediante coordinate geografiche, dal momento che il territorio è uno dei luoghi privilegiati di convergenza dei dati per una visione globale che consente di ricostruire i “cronsistemi culturali”.

Da questi pochi dati che ho fornito, si capisce come non si tratta di un sistema descrittivo che produce “metadati” sulla risorsa materiale dell'opera o dell'oggetto culturale, come fossero semplici dati di natura “amministrativa” sulle risorse digitali

che “rappresentano” il bene o “metadati descrittivi minimi” che richiamano in maniera approssimativa elementi conoscitivi che risiedono altrove: piuttosto è un sistema descrittivo che genera conoscenza (molto spesso nella ricerca scientifica le schede di catalogo sono fonte primaria di conoscenza). Sono dunque standard che producono “dati”, che hanno una loro autonomia scientifica, indipendente dalla risorsa digitale che riproduce il bene culturale a cui possono essere collegati (risorsa, questa sì, invece corredata da metadati amministrativo-gestionali).

Ovviamente non tutto quello che è stato prodotto in termini di catalogazione (che ricordo è un processo partecipato) ha le caratteristiche di fonte primaria di conoscenza, giacché c’è stato un lungo periodo in cui le scelte politiche sono andate nella direzione della catalogazione speditiva, che ha coinciso - per altro - proprio con il *digital turn* nella fruizione dei beni culturali, in cui la scheda di catalogo doveva fornire solo e semplicemente i dati essenziali per mettere una didascalia sotto la riproduzione fotografica del bene.

Potete constatare voi stessi navigando nei siti di consultazione del Catalogo, ma anche nei portali di Cultura Italia e Musei d-Italia, che questo tipo di concezione restituisce dati pressoché inservibili nel nuovo paradigma del web semantico, in quanto poveri, spesso ambigui e approssimativi. E quindi oggi, alla prova dell’integrazione digitale tra le descrizioni di risorse, ciò che da alcuni è stato rappresentato come un limite delle normative catalografiche (la loro ampiezza e complessità) si sta dimostrando invece oggi un valore.

Le prospettive future per l’integrazione digitale

Il quadro che ho delineato non è privo di punti deboli che vorremmo migliorare, anche grazie al lavoro che si potrebbe avviare da questo tavolo di confronto.

- il primo margine di miglioramento risiede nell’incrementare i tesauri rispetto a quelli oggi presenti (ad esempio nel caso del vocabolario per la materia e la tecnica, che andrebbe organizzato gerarchicamente - a partire da categorie materiche generali - e omogeneizzato per l’applicazione sia in ambito storico artistico sia in ambito archeologico, con riferimento anche a quanto già formalizzato nel settore del restauro/conservazione), magari adottando vocabolari controllati già esistenti presso altre realtà. Considerando la vastità dei vocabolari che si sono prodotti in anni di catalogazione partecipata risulta oggi assai oneroso procedere a un allineamento *ex-post* tra tesauri diversi (soprattutto se multilingue); potrebbe invece essere più utile procedere per ambiti omogenei e in modo incrementale, laddove esista una comunità

scientifico che ha concreto interesse su uno specifico aspetto, utilizzando anche gli strumenti relativi all'allineamento tra significati che già da qualche anno sono stati sperimentati nell'ambito del semantic web.

- la seconda prospettiva di lavoro riguarda la definizione di ontologie per la modellazione del sistema di standard descrittivi messi a punto nel tempo dall'ICCD. Data la vastità del sistema che prima vi ho sinteticamente descritto, si comprende come non sia per noi pienamente soddisfacente adottare uno standard esistente a livello internazionale (ad esempio CIDOC o Europeana Data Model) ed adattarlo – o, peggio, “forzarlo” - alla bisogna, questo perché:

1. uno degli scopi dei LOD dovrebbe essere quello di rappresentare il dato nella maggior profondità semantica possibile; è opportuno quindi che ogni dominio di conoscenza produca i modelli descrittivi che maggiormente riescano a far emergere i contenuti culturali del dominio stesso (ricordando che non si tratta di modellare metadati ma dati di conoscenza primaria): una delle tendenze – anche nella principale letteratura scientifica del settore - è quella di produrre ontologie “complete” relative ad un determinato dominio e lavorare successivamente sugli allineamenti semantici tra modelli ontologici diffusi o conosciuti o comunque ritenuti significativi (come è stato fatto per l'ontologia Cultural-ON).

2. Con questi strumenti a disposizione, appare anacronistico sviluppare sistemi che sacrificino le specificità dei vari domini e che costringano sistemi nati per la descrizione analitica del patrimonio ad “appiattare” le proprie descrizioni sulla base di tracciati o standard predefiniti, dal momento che con l'allineamento sopra descritto si ottengono risultati adeguati in fase di costruzione degli indici (ad esempio per definire le chiavi di ricerca “comuni”) restituendo al contempo all'utente tutta la ricchezza informativa di cui i sistemi attuali dispongono.

Poiché gli standard messi a punto dall'ICCD negli anni sono fortemente strutturati ed estremamente analitici, e l'esperienza nel settore della formalizzazione della conoscenza è oramai un patrimonio consolidato, ci aspettiamo che la “traduzione” dei suddetti standard in ontologie (attività già avviata con il supporto del CNR) sarà un'operazione senza ostacoli; la pubblicazione del catalogo secondo gli standard del *semantic web* impatterà così positivamente sulla condivisione delle conoscenze desumibili dal Catalogo.

- la terza prospettiva di lavoro, e qui arrivo finalmente all'ultima questione posta dal documento di presentazione di questo seminario, sta nel trovare una modalità di integrazione tra sistemi descrittivi differenti del patrimonio culturale. Per tutto quello

che prima ho cercato di raccontare del nostro vasto sistema di standard, è assolutamente insoddisfacente per noi procedere all'integrazione con altri domini disciplinari attraverso la strada del set minimo di dati comuni ai diversi tracciati descrittivi (per esempio lo standard Dublin Core), come avviene nei principali aggregatori oggi esistenti; questo produce un appiattimento dei contenuti descrittivi di ciascun dominio, determinando sommatorie di dati e non integrazioni di contenuti. Per altro ciò determina un'ampia possibilità di errore nelle ricerche integrate dal momento che un elemento potrebbe avere significati completamente diversi nei vari domini (vedi per esempio il concetto - volutamente ambiguo - di dc:creator del Dublin Core con cui nei sistemi che finora hanno descritto trasversalmente il patrimonio culturale dovrebbero teoricamente essere pubblicati sia i soggetti produttori di archivio, che gli autori di fotografia – e qui anche stampatori, editori - che gli artisti, ovvero entità la cui descrizione comporta necessari e irrinunciabili distinguo).

Dal momento che, nella logica della rete, il valore non è dato dalla quantità di oggetti digitali ma dalle relazioni tra essi, ci sembrerebbe invece più interessante individuare le entità comuni ai vari sistemi descrittivi e lavorare ad una modellazione condivisa di tali entità all'interno delle ontologie di dominio. Penso in prima istanza ai LUOGHI, alle PERSONE, e alle CRONOLOGIE, ma anche alle ORGANIZZAZIONI e agli EVENTI, indipendentemente dal significato che queste entità possono assumere nei vari domini (cronologia di produzione, di formazione, di ritrovamento, di inventariazione, ecc; persona come autore dell'opera, come soggetto produttore di un archivio, come tipografo di un libro a stampa ecc.). Queste sono alcune delle entità – forse le principali - che concorrono a ricostruire i contesti, certamente più che – ad esempio - i dati sulle caratteristiche materiali degli oggetti culturali, e che ci consentiranno di giungere a una visione del patrimonio realmente integrata.

Per cui sono molto interessata al dibattito che si svilupperà oggi; dal momento che abbiamo appena iniziato a lavorare sulla definizione delle ontologie dell'ampio sistema di standard che governiamo, questo è il momento giusto per avviare scambi e confronti.

* * *

Tavola rotonda del pomeriggio

Come ho accennato nell'intervento di questa mattina, l'ICCD è pronto ad uniformare ed allineare il proprio sistema di authority a quanto già fatto in ambito archivistico e bibliografico; ugualmente per quanto riguarda i sistemi di datazione.

Si tratta di un lavoro forse lungo, per la quantità di pregresso da gestire, ma sicuramente non dovrebbe avere sorprese dal punto di vista metodologico.

ICCD è disponibile inoltre a condividere con gli altri settori del dominio dei beni culturali la parte di ontologie relative alle entità comuni, così da facilitare nella fase successiva alla pubblicazione dei LOD, l'interlinking dei dati e – attraverso opportune mappature - la costruzioni di indici.

Non sono invece d'accordo sul fatto che l'integrazione dei dati tra settori diversi possa avvenire partendo da un software. Piattaforme come il citato MetaFad, che hanno risolto solo tecnologicamente l'integrazione tra dati provenienti da sistemi descrittivi diversi, necessitano di un'approfondita e condivisa metodologia che può nascere anzitutto dalla condivisione e formalizzazione delle complesse relazioni semantiche tra i dati: la risposta non può essere trovata nei sistemi, ma nei dati stessi e la nuova sfida per gli Istituti centrali dovrebbe essere l'esplicitazione dei modelli formali e delle relazioni tra i modelli, da mettere a disposizione per le comunità sviluppatrici di piattaforme.

Il modello che intendiamo proporre, anche nel Piano nazionale di digitalizzazione, è un modello che parte dai dati, che come ha detto giustamente Anna Maria Marras, sono l'unica cosa che non invecchia.

Ciascuna amministrazione è tenuta a pubblicare il proprio patrimonio informativo in formato aperto, possibilmente secondo il paradigma dei *Linked Open Data*, possibilmente su un unico *End Point*. Al sito dei LOD, modellati sulla base di ontologie di dominio e non su ontologie semplificate, si affiancheranno le teche digitali, ovvero i molteplici repository esistenti degli oggetti digitali, opportunamente metadati per garantirne l'accesso e la conservazione a lungo termine.

Su queste due componenti, si innesteranno le tecnologie atte a consentire la mappatura delle varie ontologie utilizzate e a ottimizzare le prestazioni delle interrogazioni sui sistemi di origine mediante la creazione di indici. Sul "motore" così configurato, si potranno innestare le piattaforme di servizi, prima tra tutte la consultazione integrata, sviluppati dal Mibact o da qualsiasi altro soggetto abilitato.

Ancora una volta ribadisco che il senso dell'integrazione tra sistemi descrittivi non è solo avere il punto di accesso unico, che pure ovviamente ha i suoi vantaggi, ma creare relazioni di qualità tra dati di qualità. Non tutto quello che a oggi è stato fatto dal Mibact va purtroppo in questa direzione, per crescere bisogna essere convinti che innovare significa pensare a qualcosa che ancora potrebbe non esistere.