

## 2 Patrimonio culturale, archivi digitali e aggregatori

ANTONIO DAVIDE MADONNA

Ministero della Cultura

DOI: 10.25430/pupb-9788869384394-09

---

Il livello di maturità raggiunto nella metadattazione del patrimonio culturale richiede di bilanciare ricercabilità e ricchezza dei dati. Dal punto di vista della aggregazione è cruciale garantire la ricercabilità tramite standard diffusi e vocabolari controllati, mentre la ricchezza dei dati sorgente, come mostrato dalla piattaforma Phaidra, assicura la completezza delle informazioni. Questo duplice approccio ottimizza l'interoperabilità e il valore scientifico dei dati come dimostrato dai risultati eccellenti di Phaidra in Europeana dove ha ottenuto il 98,72% di conformità al Tier C raggiunto in generale dal 4,60% delle risorse.

*The level of maturity achieved in cultural heritage metadata requires a careful balancing act between searchability and data richness. At the aggregation level, it is of paramount importance to ensure searchability through the implementation of widely adopted standards and controlled vocabularies. Conversely, the richness of the source data, as demonstrated by the Phaidra platform, is essential for ensuring the completeness and accuracy of the information. This dual approach optimises interoperability and the scientific value of the data, as evidenced by the exemplary results achieved by Phaidra in Europeana, where it attained 98.72% Tier C compliance, in comparison to the overall compliance rate of 4.60%.*

Patrimonio culturale digitale, Interoperabilità, Metadati, Archivi digitali, Aggregatori  
*Digital Cultural Heritage, Interoperability, Metadata, Digital repositories, Aggregators*

---

Il grado di maturità raggiunto nell'ambito della metadattazione del patrimonio culturale impone alcune riflessioni.

I tempi in cui lo slogan era “purché si digitalizzi” sono ampiamente finiti, sostituiti da una maggiore consapevolezza di ciò che oggi serve al pubblico di riferimento al fine di ottimizzarne l'esperienza. In tale scenario, la strutturazione delle informazioni ha assunto un valore centrale.

In particolare, la sfida principale che dobbiamo affrontare riguarda la capacità di traslare in maniera efficiente ed efficace le informazioni catalografiche in modo che rispettino i *FAIR data*.

Definiti nel 2016, tali principi hanno avuto ampia diffusione anche in ambito culturale. Generare un metadato, di conseguenza, non è più una mera operazione di duplicazione di informazioni note, ma anche un processo che garantisca al dato di essere ricercabile (*Findable*), accessibile (*Accessible*), interoperabile (*Interoperable*) e riusabile (*Reusable*).

Considerando che i primi due punti riguardano prevalentemente aspetti gestionali (legati ad esempio all'utilizzo di *persistent identifiers*, oppure il rendere disponibili i metadati attraverso applicativi *open*), sono i principi dell'interoperabilità e della riusabilità quelli che più incidono sulla struttura del metadato e che assumono valore centrale nell'ambito dell'aggregazione del patrimonio culturale.

Da questo punto di vista, il primo aspetto da tenere in considerazione riguarda l'utilizzo di standard che siano ampiamente diffusi, sia a livello nazionale che internazionale.

Il mondo digitale ha ampiamente abbattuto frontiere che un tempo sembravano invalicabili, per cui l'interconnessione tra sistemi informativi differenti non può essere considerata un desiderata, ma un *must* a cui non ci si può sottrarre.

Particolarmente significativo è l'impegno, a livello europeo, della Commissione per la creazione di un *data space* comune per il settore dei beni culturali digitalizzati. Tale spazio, infatti, non deve essere inteso come mero luogo di *storage*, ma una piattaforma di ideale condivisione che consenta l'ottimizzazione di una serie di processi che al momento avvengono a livello locale.

Tale processo di condivisione passa anche attraverso l'adozione di standard comuni di metadato. Altrimenti, il rischio di trovarsi di fronte ad una nuova torre di Babele è oggettivamente concreto. Ma anche giungere ad un esperanto di settore è ugualmente pericoloso. La ricchezza dei metadati, infatti, deve essere salvaguardata, al fine di preservarne il valore scientifico.

Giungere ad un compromesso non è semplice, ma l'esperienza consolidata nel corso degli anni, soprattutto nel campo dell'aggregazione, consente oggi di distinguere, idealmente, due differenti livelli di metadato che si integrano naturalmente.

Da un lato, quindi, avremo i metadati utili ai fini della ricercabilità dei dati; dall'altro avremo i metadati che ne garantiranno la ricchezza e la completezza.

Prendendo in prestito dal mondo anglosassone la regola delle cinque W (ma sfruttandone solo quattro), la utilizzeremo per definire i metadati che forniscono risposta alle domande "Why", "What", "Where" e "When" come quelli in grado di aumentare in maniera notevole la loro ricercabilità.

L'esperienza complessiva maturata nel settore dell'aggregazione ha infatti permesso di stabilire, in maniera chiara, che il comportamento degli utenti (e la conseguente metodologia di ricerca) è indirizzato in maniera decisa verso uno o più di questi criteri.

Definito questo aspetto e il *core* delle informazioni da condividere, ci troviamo di fronte ad un quesito di fondamentale importanza: in che modo possiamo garantire l'ottimizzazione del processo di condivisione delle informazioni per incrementare la ricercabilità dei dati ad un livello più ampio?

Si è già discusso dell'importanza relativa all'utilizzo di formati standard. Ma esiste anche un altro modo per incrementare la ricercabilità delle risorse in contesti più ampi.

L'utilizzo di vocabolari controllati rappresenta un valore aggiunto di notevole rilevanza. Definire in maniera univoca un nome, un luogo, una data o la tipologia di bene che stiamo descrivendo implica una serie di notevoli vantaggi.

Basti pensare ai LOD e alle possibilità di interconnessione delle risorse se si hanno dei valori univoci da condividere, oppure al vantaggio di rendere immediatamente "leggibile" l'informazione anche a chi non conosce una determinata lingua.

Una volta stabilite le strategie e gli strumenti per migliorare la ricercabilità del dato, resta aperto il fronte riguardante la ricchezza delle informazioni da veicolare tramite metadati.

Per questo aspetto, la soluzione è indubbiamente più semplice. Se consideriamo i portali di aggregazione il punto di arrivo di un processo di condivisione, allora dobbiamo identificare le fonti dati quali luoghi deputati a mostrare i metadati nella loro completezza. Ciò non significa che in fase di condivisione sia opportuno fornire esclusivamente campi obbligatori e raccomandati, ma consente di strutturare un *mapping* tra sistemi informativi diversi in base a logiche di ottimizzazione del flusso di lavoro.

La piattaforma Phaidra, da questo punto di vista, rappresenta un'eccellenza che potrebbe assumere il ruolo di *case study*. In particolare, la collezione *Archivio storico dell'Orto botanico* contiene una serie di informazioni in grado di soddisfare sia i criteri della ricercabilità che della ricchezza dei metadati.

La scelta di adottare un doppio livello di metadattazione (Dublin Core e il modello Phaidra) si dimostra una scelta vincente ed in linea con quanto auspicato in precedenza. La dimostrazione più lampante, da questo punto di vista, è il raggiungimento del metadata Tier C all'interno di Europeana, a dimostrazione dell'elevato grado di completezza delle informazioni veicolate tramite metadati.

Per comprendere il valore del lavoro svolto dall'Università di Padova è sufficiente analizzare i numeri. All'interno di Europeana, solo il 4,60% delle risorse raggiunge il Tier C, mentre Phaidra si attesta al 98,72%.

Una differenza abissale, frutto anche di scelte strategiche che non guardano solo al breve, ma anche al medio-lungo periodo.

Una lungimiranza, in un'epoca di trasformazioni sempre più veloci, che può fare la differenza e che rappresenta una base solida per continuare ad essere, anche in futuro, un punto di riferimento per il mondo della cultura digitale.