31 luglio 2026

Finanziato dall'Unione europea NextGenerationEU

MiC MINISTERO DELLA CULTURA

DIGITAL LIBRARY Istituto centrale per la digitalizzazione del patrimonio culturale

Ecomic ecosistema digitale per la cultura

DPaaS Piattaforma data product as a service

DATASET, METADATI E RELAZIONI PERFORMANCE–OPERE–SPAZI DOUBLE REALITY

Sistema di raccolta, normalizzazione e arricchimento semantico dei metadati di immagini, video e documenti delle collezioni digitalizzate della Fondazione Brescia Musei.

Cinque musei, un patrimonio digitale connesso

Dai contenuti dei cinque musei a una conoscenza connessa

Le collezioni digitali della Fondazione Brescia Musei comprendono immagini, video, audio e documenti provenienti da Santa Giulia, Brixia, Pinacoteca Tosio Martinengo, Museo delle Armi Luigi Marzoli e Museo del Risorgimento.

Gli asset documentano opere, performance, allestimenti, luoghi ed esperienze immersive.

Tuttavia, la varietà dei formati e la distribuzione dei contenuti rendono complesso ricostruire le connessioni tra ciò che viene rappresentato, il museo di appartenenza e lo spazio in cui si svolge l'esperienza.

In che modo il Data Product genera valore?

Il Data Product analizza ogni file e ne costruisce un profilo unificato, composto da metadati tecnici e informazioni culturali.

Il sistema identifica opere, persone, luoghi, organizzazioni e soggetti, mettendo in relazione:

- la performance documentata;
- le opere coinvolte o rappresentate;
- gli spazi museali ed espositivi;
- il museo di riferimento;
- gli asset digitali che testimoniano l'esperienza.

Il risultato è un dataset relazionale interrogabile ed esportabile, nel quale i contenuti dei cinque musei diventano una rete di informazioni pronta per alimentare analisi, esperienze virtuali e nuovi servizi digitali.

Ogni asset viene collocato nel suo contesto museale

Livello 1
Profilo tecnico

Il sistema estrae automaticamente le proprietà oggettive dell'asset:
- formato e dimensione;
- autore e date di creazione o modifica;
- risoluzione per le immagini;
- durata e codec per audio e video;
- numero di pagine per i documenti;
- data di acquisizione nel sistema.

Queste informazioni permettono di conoscere consistenza, qualità e composizione tecnica del patrimonio digitale.

Livello 2
Profilo culturale

Attraverso servizi di intelligenza artificiale, il Data Product analizza ciò che il contenuto rappresenta o documenta, producendo:
- titolo e abstract;
- persone e organizzazioni;
- opere e beni culturali;
- luoghi e spazi espositivi;
- soggetti e temi;
- performance ed eventi;
- parole chiave normalizzate.

Le informazioni vengono organizzate secondo standard condivisi, come METS, ArCo e JSON-LD.

Relazioni – La conoscenza che collega i contenuti



I metadati tecnici e culturali confluiscono in un profilo unitario e vengono collegati al contesto di appartenenza.
L'archivio digitale diventa così una rete navigabile di contenuti, opere, esperienze e luoghi museali.

Il sistema può associare:

- un video alla performance che documenta;
- la performance alle opere coinvolte;
- le opere allo spazio espositivo;
- lo spazio al museo di riferimento;
- persone, luoghi e soggetti agli asset in cui compaiono.

Dalla fonte alla risposta contestualizzata

1

INPUT

01 · Acquisizione

Il file arriva dallo storage, viene riconosciuto e controllato contro i duplicati.

2

ANALISI

02 · Arricchimento

AI estrae o trascrive contenuti e riconosce entità e relazioni culturali.

3

OUTPUT

03 · Rendere disponibile

Database, viste, API ed export espongono profili e aggregazioni.

Un’architettura cloud-native trasforma gli asset in dati riutilizzabili

Il Data Product utilizza una pipeline asincrona: restituisce rapidamente le informazioni tecniche essenziali e prosegue in background con le elaborazioni più complesse di intelligenza artificiale.

In questo modo, acquisizione, gestione e valorizzazione dei contenuti confluiscono in un unico processo.



Acquisizione e accesso

Asset acquisiti dallo storage cloud attraverso API REST. Il sistema identifica il formato, calcola l’impronta SHA-256 e assegna un codice alla lavorazione. Le API permettono di:

- avviare l’elaborazione;
- controllarne lo stato;
- consultare il profilo degli asset;
- interrogare le viste analitiche;
- esportare i dati in CSV o JSON.



Normalizzazione e organizzazione

Le informazioni estratte vengono ricondotte a concetti e vocabolari condivisi. Termini provenienti da file differenti possono così essere riconosciuti come riferimenti alla stessa opera, persona, performance o spazio museale.

Il risultato è un grafo semantico interoperabile, utilizzabile anche da altri sistemi culturali.



Estrazione tecnica e arricchimento semantico

Apache Tika estrae i metadati tecnici. I servizi di IA analizzano il contenuto:

- OCR per riconoscere il testo nelle immagini;
- Whisper per trascrivere audio e video;
- NERS per identificare persone, luoghi, organizzazioni e opere;
- Tagging per classificare i temi;
- Llama 3.1 per generare abstract e strutturare le relazioni tra performance, opere e spazi.



Persistenza, analisi e riuso

Metadati tecnici, informazioni semantiche e relazioni vengono archiviati in PostgreSQL. Sette viste pre-elaborate rendono disponibili:

- il profilo completo di ogni asset;
- le relazioni tra performance, opere e spazi;
- le entità più frequenti;
- le statistiche per tipologia di file;
- la distribuzione temporale delle acquisizioni;
- i principali autori;
- il grafo delle relazioni semantiche.

Quattro scenari applicativi

01/ Esplorazione delle collezioni

Consultazione del profilo completo di un asset e delle opere, persone, performance e spazi associati.

02/ Percorsi tra i musei

Creazione di esperienze trasversali basate sulle relazioni tra opere, luoghi e contenuti dei cinque musei.

03/ Analisi e qualità dei dati

Monitoraggio di formati, autori, acquisizioni e completezza dei metadati, con export CSV e JSON.

04/ Esperienze digitali

Sviluppo di visite virtuali, motori di ricerca, mappe interattive e strumenti di visualizzazione.

Riusabilità

Un modello riutilizzabile per collezioni distribuite

Il Data Product può essere adottato da musei, fondazioni e reti culturali che gestiscono patrimoni digitali provenienti da più collezioni, sedi o spazi espositivi.

Condizioni per il riuso: l'ente deve definire asset, collezioni, luoghi ed eventi da mettere in relazione, adattando il modello semantico al proprio patrimonio.

Connettere patrimoni diversi

Riunire immagini, video, audio e documenti in un unico modello informativo, collegandoli a opere, eventi, luoghi e istituzioni.
Ogni asset viene descritto attraverso metadati tecnici e semantici, permettendo di superare la frammentazione tra formati, repository e sistemi documentali differenti.

Gestire reti museali

Organizzare i contenuti di più sedi e rendere esplorabili le relazioni tra collezioni, spazi e iniziative culturali.

Migliorare qualità e accesso

Uniformare i metadati, monitorare la consistenza del corpus e rendere i dati disponibili tramite API ed export CSV/JSON.

Sviluppare nuovi servizi

Alimentare portali, motori di ricerca, percorsi digitali, esperienze virtuali e strumenti di analisi destinati a pubblici e professionisti.

Arte fisica e patrimonio digitale si incontrano in una stessa rete di dati: opere, spazi ed esperienze diventano parti connesse di un'unica esperienza museale.