31 luglio 2026

Finanziato dall'Unione europea NextGenerationEU

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

TRAME DI COMUNITÀ A OLIVETO CITRA – DATASET RELAZIONALE DI PERSONE, LUOGHI ED EVENTI SULL'EMIGRAZIONE OLIVETANA

Creazione di un dataset relazionale contenente persone, luoghi ed eventi legati all'emigrazione olivetana dall'Ottocento, integrando informazioni da documenti d'archivio e dalle 12 storie disponibili.

Le storie dell'emigrazione diventano patrimonio condiviso quando persone, luoghi, eventi e documenti si ricompongono in un unico catalogo relazionale condiviso

Persone, luoghi ed eventi tornano a parlarsi

Fotografie d’archivio, documenti anagrafici, corrispondenza e testimonianze audiovisive conservano la memoria dell’emigrazione da Oliveto Citra.

Queste fonti raccontano storie individuali e familiari, ma sono distribuite tra formati e documenti differenti, rendendo difficile ricostruire i legami tra persone, destinazioni ed eventi biografici.

In che modo il Data Product genera valore?

Il Data Product analizza i contenuti e identifica emigranti, famiglie, luoghi, date, attività ed eventi come nascite, matrimoni, partenze e arrivi.

Le informazioni vengono organizzate in un catalogo relazionale che collega ogni entità ai documenti e alle testimonianze in cui compare.

Diventa così possibile ricostruire percorsi migratori, riunire materiali riferiti alla stessa persona o famiglia e sviluppare nuove narrazioni sulla memoria della comunità.

Per ogni asset, un profilo che ne descrive caratteristiche tecniche e significato culturale

Livello 1
Identità tecnica

Il sistema estrae automaticamente le proprietà oggettive del file:

- nome, formato e dimensione;
- autore e data di creazione;
- numero di pagine per i documenti;
- risoluzione per le immagini;
- durata per audio e video;
- identificativo univoco e impronta SHA-256.

Queste informazioni permettono di conoscere la consistenza e la qualità dell'archivio digitale.

Livello 2
Profilo semantico

L'intelligenza artificiale analizza il contenuto e individua:

- persone ed emigranti;
- famiglie e organizzazioni;
- luoghi di origine e destinazione;
- date e periodi storici;
- attività e temi;
- eventi biografici come nascite, matrimoni, partenze e arrivi;
- titolo e sintesi del contenuto.

Relazioni – La conoscenza che collega i contenuti



I metadati tecnici e semantici vengono riuniti e collegati secondo standard come METS, ArCo e JSON-LD.
Ogni asset diventa così parte di un catalogo relazionale che connette persone, luoghi, eventi e documenti, rendendo ricostruibili le storie dell'emigrazione olivetana.

La pipeline traduce l’eterogeneità in conoscenza condivisa

1

INPUT

01 · Acquisizione

Il sistema importa fotografie, documenti e testimonianze audiovisive dallo storage cloud, verifica eventuali duplicati ed estrae i metadati tecnici.

2

ANALISI

02 · Interpretazione

OCR, trascrizione e servizi di intelligenza artificiale riconoscono persone, famiglie, luoghi, date, temi ed eventi biografici presenti nei contenuti.

3

OUTPUT

03 · Rendere disponibile

Le informazioni vengono normalizzate secondo METS e ArCo, archiviate in PostgreSQL e organizzate in relazioni tra persone, eventi, luoghi e documenti.

Un’architettura dati per analisi, mappe ed esperienze immersive

Il Data Product è un microservizio cloud-native che integra componenti specializzati per trattare documenti, immagini, contenuti audiovisivi, dati strutturati e modelli 3D all’interno di un unico processo.



Acquisizione e accesso

Le API REST gestiscono l’acquisizione degli asset dallo storage cloud, l’avvio delle elaborazioni e la consultazione dei risultati.

Il sistema calcola l’impronta SHA-256 del file per evitare duplicazioni e utilizza processi asincroni per svolgere le analisi più complesse in background.



Estrazione e arricchimento

Apache Tika estrae i metadati tecnici da fotografie, documenti, audio e video. I servizi AI completano il profilo attraverso:

OCR dei contenuti visivi;

- trascrizione delle testimonianze audiovisive;
- riconoscimento di persone, luoghi, organizzazioni e date;
- classificazione dei temi;
- generazione di abstract;
- identificazione degli eventi biografici.



Normalizzazione semantica

Le informazioni estratte vengono ricondotte a concetti e vocabolari condivisi. Persone, luoghi, attività, periodi ed eventi possono così essere riconosciuti e collegati anche quando sono indicati in modo diverso nelle fonti.



Persistenza, analisi e riuso

Metadati e relazioni vengono archiviati in PostgreSQL e resi disponibili attraverso profili, viste analitiche, API REST ed export CSV/JSON.

La modularità permette di aggiungere nuovi contenuti, estendere le entità analizzate e alimentare portali, mappe e percorsi narrativi dedicati alla memoria dell’emigrazione.

Quattro scenari per riattivare la memoria locale

01/Ricerca per persona

Riunire tutti i materiali legati a un emigrante o a una famiglia.

02/ Percorsi migratori

Correlare partenze, arrivi, matrimoni e luoghi nel tempo.

03/Narrazioni digitali

Alimentare storie AR, mappe e percorsi di comunità.

04/Analisi del corpus

Esportare dati per studi, statistiche e visualizzazioni.

Riusabilità

Il Data Product può essere adottato da comuni, archivi, biblioteche, musei e istituzioni culturali che conservano fotografie, documenti, corrispondenza e testimonianze audiovisive legate alla storia di una comunità.

La stessa architettura può essere adattata a patrimoni differenti, definendo le persone, i luoghi, gli eventi e le relazioni più significative per il nuovo contesto.

Condizioni per il riuso: l'ente deve selezionare il corpus, verificare qualità e formati degli asset e definire le entità e le relazioni più significative per il proprio patrimonio.

Riunire fonti distribuite

Contenuti provenienti da fondi, collezioni e repository diversi possono confluire in un unico catalogo relazionale.
Il sistema estrae metadati tecnici e semantici da immagini, documenti, audio e video, rendendo le fonti più facilmente consultabili e confrontabili.

Ricostruire storie e percorsi

Persone, famiglie, luoghi, date ed eventi vengono riconosciuti e collegati ai documenti in cui compaiono.
Ciò consente di ricostruire biografie, percorsi migratori, legami familiari e trasformazioni del territorio, mantenendo sempre il riferimento alle fonti originali.

Supportare ricerca e gestione

Archivisti, curatori e ricercatori possono consultare il profilo completo degli asset, individuare le entità più ricorrenti ed esplorare le relazioni tra persone, eventi e luoghi.
API REST ed esportazioni CSV/JSON consentono inoltre di integrare i dati in strumenti di analisi, sistemi documentali e applicazioni esterne.

Creare nuove narrazioni

Il catalogo può alimentare mappe dei percorsi migratori, portali di comunità, mostre digitali, itinerari territoriali e strumenti didattici.
Le memorie conservate nell'archivio diventano così una base strutturata per nuove forme di ricerca, partecipazione e valorizzazione culturale.

Un archivio di comunità acquista nuovo significato quando ogni documento ritrova le persone, i luoghi e gli eventi di cui conserva memoria.