DICITUR

Documento di specifiche tecniche di elaborazione

Versione del documento: 1.0

Data: 28/08/2026

Sommario

1. Obiettivi del Data Product 4
2. Origine dei dati ed eventuali elaborazioni 5
3. Tecnologie utilizzate all'interno del Data Product Lab 9
3.1. Servizi abilitanti DPaaS 10
3.2. Eventuali tecnologie esterne a DPaaS 13
3.2.1. Modelli in locale 13
3.2.2. PyAnnote per la diarizzazione e l'identificazione dei parlanti 14
4. Pipeline di elaborazione e trasformazione 16
4.1. Pipeline di generazione di sottotitoli, riassunti e metadati semantici per contenuti audiovisivi 16
4.1.1. Preparazione dell'audio e analisi preliminare 16
4.1.2. Trascrizione e generazione dei sottotitoli 16
4.1.3. Traduzione dei sottotitoli in lingua inglese 17
4.1.4. Generazione di riassunti accessibili 17
4.1.5. Produzione del riassunto in lingua inglese 17
4.1.6. Generazione delle versioni audio dei riassunti 17
4.1.7. Estrazione delle entità e dei concetti 17
4.1.8. Generazione dei tag semantici 18
4.1.9. Produzione dei metadati di interoperabilità 18
4.1.10. Revisione e validazione dei contenuti 19
4.1.11. **Pubblicazione degli artefatti** 19
5. Esposizione e distribuzione 20
5.1. Canali di accesso 20
5.2. Utenti e sistemi destinatari 20
5.3. Autenticazione e autorizzazione 20
5.4. Formati disponibili 21
5.5. Modalità di aggiornamento 21
5.6. Limiti di utilizzo 21

5.7. Livelli di servizio 21
5.8. Accesso alla documentazione 21
5.9. Condizioni di licenza e riutilizzo 22
5.10. Fruizione tramite API 22
5.10.1. Tipologie di informazioni interrogabili 22
5.10.2. Principali operazioni disponibili 22
5.10.3. Ricerca, filtro, ordinamento e paginazione 23
5.10.4. Recupero del dettaglio di una risorsa 23
5.10.5. Servizi di aggregazione e analisi 23
5.10.6. Formati delle richieste e delle risposte 23
5.10.7. Modalità di autenticazione 23
5.10.8. Limiti di chiamata 23
5.10.9. Gestione degli errori 24
5.11. Limitazioni 25
5.12. Nota sulle condivisioni dei dati 26
6. Possibili evoluzioni future 27

1.Obiettivi del Data Product

Il Data Product ICAR ha l'obiettivo di trasformare contenuti audio e video in materiali più accessibili e facili da consultare. A partire da un file multimediale, il sistema genera automaticamente sottotitoli, riassunti, traduzioni e contenuti audio, arricchendo inoltre il contenuto con informazioni semantiche

2. Origine dei dati ed eventuali elaborazioni

1. Sistema, archivio, database, API o dataset di origine

La fonte originaria del Data Product è costituita dalle collezioni digitali e dai relativi contenuti messi a disposizione dall'Istituto Centrale per gli Archivi (ICAR) attraverso il sistema "Ti racconto la storia".

Il dataset sorgente, identificato come DS00000112 ("DS Standard 2 - Catalogo accessibile delle risorse ICAR"), appartiene al dominio archivistico e ha origine remota.

Il sistema sorgente utilizza un repository basato su Fedora Commons, interrogabile mediante servizi REST. La fonte mette a disposizione gli oggetti digitali, i relativi metadati e i diversi file associati a ciascun oggetto, denominati datastream.

Le collezioni presenti nel sistema sorgente costituiscono anch'esse oggetti digitali e dispongono di propri datastream e metadati. Ogni collezione può comprendere ulteriori oggetti, ciascuno composto dai relativi file e contenuti descrittivi.

2. Soggetto responsabile o titolare della fonte

L'Istituto Centrale per gli Archivi (ICAR) è il soggetto responsabile del sistema sorgente, delle collezioni archivistiche digitali e dei contenuti messi a disposizione attraverso il sistema "Ti racconto la storia". ICAR è responsabile della disponibilità della fonte remota, della corretta organizzazione delle collezioni e della presenza dei metadati e dei datastream associati agli oggetti digitali.

La titolarità dei diritti sui singoli video, sulle immagini di anteprima, sui metadati descrittivi, sui descrittori archivistici e sugli altri contenuti associati deve essere ricondotta alle informazioni fornite da ICAR e alla relativa documentazione amministrativa (per gli aspetti di licenza si veda il punto 7).

3. Tipologia e formato dei dati

I dati sorgente messi a disposizione da ICAR sono costituiti principalmente da:

- collezioni archivistiche digitali;
- oggetti digitali appartenenti alle collezioni;
- video della raccolta "Ti racconto la storia";
- immagini di anteprima degli oggetti e delle collezioni;
- metadati descrittivi, inclusi metadati Dublin Core;

- descrittori archivistici in formato XML;
- informazioni EAD;
- informazioni relative alla licenza;
- relazioni di appartenenza tra oggetti e collezioni.

Ogni oggetto digitale appartiene a una collezione ed è composto da più datastream, che possono comprendere un'immagine di anteprima, metadati Dublin Core, descrittori XML e uno o più video. Anche le collezioni possono disporre di un'immagine di anteprima, metadati Dublin Core, un descrittore XML e informazioni EAD e di licenza.

I principali formati della fonte sono:

- **MP4**, formato sorgente principale per i contenuti video;
- **JPEG**, per le immagini di anteprima;
- **XML**, per i descrittori archivistici;
- **FOXML 1.1**, per la rappresentazione e lo scambio degli oggetti digitali del repository Fedora Commons;
- formati strutturati per i metadati Dublin Core, EAD e le informazioni di licenza.

Per ciascun oggetto o collezione possono essere disponibili: l'identificativo dell'oggetto, l'identificativo della collezione di appartenenza, il titolo, i metadati Dublin Core, il descrittore XML, l'immagine di anteprima, uno o più video, i riferimenti ai singoli datastream, le informazioni EAD, le informazioni di licenza e le relazioni tra la collezione e gli oggetti appartenenti ad essa.

4. Modalità di acquisizione

L'acquisizione dei dati avviene mediante accesso remoto ai servizi esposti dal repository Fedora Commons di ICAR (modalità di origine: "Remoto (URL)"). Non è previsto il conferimento tramite file Excel né il caricamento manuale di un pacchetto di dati.

Il repository Fedora Commons mette a disposizione, mediante credenziali fornite da ICAR:

- **API-A di Fedora Commons**, per l'accesso agli oggetti digitali e ai relativi datastream (contenuti rappresentati in formato FOXML 1.1);
- **Resource Index Search (RISearch)**, per la ricerca delle risorse e delle relazioni;
- **OAI-PMH**, per l'esposizione e la raccolta dei metadati.

L'acquisizione comprende sia i metadati degli oggetti e delle collezioni sia il recupero dei relativi contenuti digitali, comprese le immagini di anteprima e i video.

5. Frequenza di aggiornamento

La fonte ICAR può essere aggiornata quando nel sistema "Ti racconto la storia" vengono pubblicati nuovi oggetti, nuove collezioni o nuovi contenuti associati agli oggetti già presenti. Non risulta definita una cadenza prestabilita (giornaliera, settimanale o mensile) per la pubblicazione di nuovi dati.

I servizi RISearch e OAI-PMH possono supportare l'individuazione di contenuti nuovi o modificati e l'acquisizione differenziale delle risorse.

La frequenza di aggiornamento della fonte può pertanto essere descritta come **non periodica**, dipendente dalla pubblicazione o modifica dei contenuti nel repository ICAR.

6. Volume indicativo dei dati

La dimensione complessiva del dataset sorgente DS00000112 non è attualmente quantificata. Il volume dipende da:

- numero delle collezioni archivistiche;
- numero degli oggetti appartenenti a ciascuna collezione;
- numero dei video associati a ciascun oggetto;
- durata e dimensione dei video;
- numero e dimensione delle immagini di anteprima;
- dimensione dei metadati e dei descrittori XML;
- presenza di datastream multipli per il medesimo oggetto;
- crescita della fonte in caso di pubblicazione di nuovi contenuti.

7. Eventuali vincoli di accesso, utilizzo o licenza

L'accesso alla fonte non è pubblico e indiscriminato, ma avviene attraverso i servizi messi a disposizione da ICAR utilizzando apposite credenziali, che consentono l'accesso tecnico al repository e la lettura degli oggetti, dei metadati e dei relativi datastream.

Le condizioni di utilizzo, pubblicazione e riuso devono invece essere determinate sulla base delle informazioni di licenza associate alle collezioni e ai singoli contenuti (presenti tra i metadati di ciascuna collezione). Non risulta individuata una licenza unica applicabile indistintamente all'intero dataset DS00000112: la licenza da associare al Data Product era ancora oggetto di richiesta all'ente e non poteva essere stabilita autonomamente dall'operatore economico.

Le autorizzazioni tecniche e le condizioni giuridiche di utilizzo vanno quindi considerate separatamente:

- le **credenziali** disciplinano l'accesso tecnico alla fonte;
- i **metadati di licenza** e gli atti dell'ente disciplinano l'utilizzo, la pubblicazione e il riuso dei contenuti.

8. Livello di qualità e completezza riscontrato

La fonte presenta una struttura organizzata per collezioni, oggetti e datastream, con identificativi, metadati descrittivi e relazioni che permettono di ricostruire il contesto archivistico dei contenuti e di individuare i file associati a ciascuna risorsa.

La qualità e la completezza della fonte devono essere valutate in relazione a:

- presenza e validità dell'identificativo dell'oggetto;
- corretta appartenenza dell'oggetto a una collezione;
- presenza dei metadati Dublin Core;
- presenza e validità dei descrittori XML;
- disponibilità dell'immagine di anteprima;
- presenza e raggiungibilità dei video dichiarati;
- corretta associazione tra oggetto e datastream;
- presenza delle informazioni EAD, ove applicabili;
- presenza delle informazioni relative alla licenza;
- coerenza delle relazioni tra collezioni e oggetti.

La fonte risulta strutturalmente idonea all'acquisizione remota, ma non sono disponibili KPI quantitativi, soglie formali di accettazione o una percentuale complessiva di completezza del dataset. La qualità della fonte può pertanto essere descritta come **adeguata** all'acquisizione dei contenuti archivistici e audiovisivi, subordinatamente alla verifica della presenza, raggiungibilità e corretta associazione dei datastream relativi a ciascun oggetto e collezione.

3. Tecnologie utilizzate all'interno del Data Product Lab

Il Data Product si avvale di alcuni componenti applicativi utilizzati per l'orchestrazione dei processi di elaborazione, la gestione dei flussi di lavoro, l'esecuzione delle pipeline di Intelligenza Artificiale e il monitoraggio dello stato di avanzamento delle elaborazioni. Tali componenti operano in integrazione con le funzionalità offerte dalla piattaforma e consentono l'automazione dell'intero ciclo di generazione degli artefatti di accessibilità.

In particolare, il processo è orchestrato tramite Apache NiFi, sul quale sono configurati job schedulati che effettuano periodicamente il monitoraggio delle cartelle S3 associate ai singoli Enti conferitori. I job verificano la presenza di nuovi contenuti e controllano, attraverso appositi meccanismi di tracciamento, se i file siano già stati elaborati dal sistema.

Nel caso in cui un contenuto risulti non ancora processato, NiFi attiva una sequenza di chiamate verso due API REST esposte sul Livello 1 dell'architettura applicativa. Entrambe le API sono state sviluppate in linguaggio Python e costituiscono componenti esterne rispetto alla piattaforma DPaaS.

La prima API è dedicata alla traduzione automatica dei metadati in lingua inglese, necessaria per garantire la disponibilità delle informazioni descrittive nelle lingue supportate dal Data Product. La seconda API implementa invece la pipeline di generazione degli artefatti, responsabile della creazione dei contenuti derivati quali descrizioni, audiodescrizioni, trascrizioni, sottotitoli, riassunti e ulteriori risorse di accessibilità.

Per ciascun Ente sono stati configurati specifici task di elaborazione, definiti sulla base dei requisiti funzionali e delle esigenze espresse durante la fase di raccolta dei fabbisogni. Ogni task identifica una particolare tipologia di elaborazione da eseguire sui contenuti conferiti dall'Ente e viene gestito in maniera indipendente all'interno della pipeline applicativa.

Lo stato di avanzamento delle elaborazioni viene registrato all'interno di un database PostgreSQL, anch'esso utilizzato come componente di supporto esterno alla piattaforma DPaaS. Per ogni elaborazione sono memorizzati identificativo del task, stato di esecuzione, informazioni di monitoraggio ed eventuali errori rilevati durante il processamento, consentendo il tracciamento completo del ciclo di vita delle lavorazioni.

A livello operativo, Apache NiFi implementa inoltre specifici meccanismi di retry automatico, che consentono di rieseguire le chiamate verso le API in caso di errori

temporanei o malfunzionamenti transitori, aumentando la resilienza complessiva del sistema e riducendo il rischio di perdita delle elaborazioni.

Al completamento di tutti i task previsti per uno specifico Ente, gli artefatti generati e i relativi metadati vengono considerati disponibili per la pubblicazione e possono essere successivamente fruiti dai portali applicativi che accedono al Data Product.

3.1. Servizi abilitanti DPaaS

Di seguito vengono riportati tutti i servizi abilitanti DPaas che sono stati associati dall'ente al Data Product e sono stati utilizzati al fine di generare gli artefatti:

- CPAI – OCR;
- CPAI - Speech2Text;
- CPAI - Text2Speech;
- CPAI – Ner;
- CPAI – Tagging;
- CPAI - Image Enhance;
- CPAI – Abstract;
- CPAI - Estrazione Entità;
- CPAI - Tagging Entità;
- CPAC – Yolo;
- CPAC – WhisperTTS;
- CPAC – Whisper;
- CPAC – Tagging;
- CPAC - Stable Diffusion;
- CPAC - Real ESRGAN;
- CPAC - Qwen2;
- CPAC - P2PaLa;
- CPAC – NERs;
- CPAC - Llama 31;
- CPAC - Kosmos2;
- CPAC – Keras
- CPAC- Gliner
- VM;
- JupyterLab;

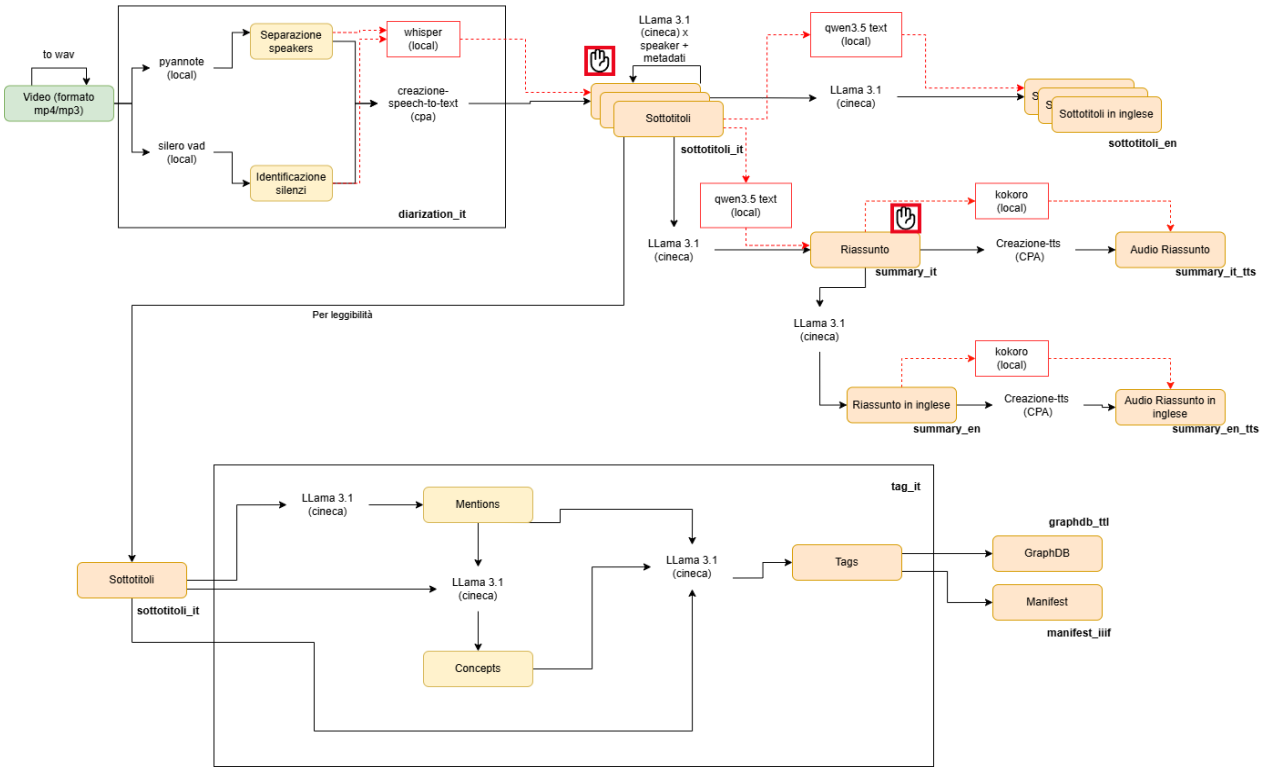
- Neo4j;
- Ontotext Refine;
- OpenSeaDragon;
- NiFi;
- Postgres;
- Elastic Search;
- GraphDB.

Si evidenzia che **non tutti i servizi, modelli e componenti riportati nell'elenco delle tecnologie disponibili sono stati effettivamente utilizzati nell'ambito della presente fase progettuale**. La loro inclusione è stata richiesta all'Ente al fine di consentire un'attività di confronto tra diverse soluzioni tecnologiche, sia in termini di prestazioni sia di qualità dei risultati ottenuti, nonché per valutare la fattibilità di specifici requisiti funzionali e non funzionali.

L'associazione delle tecnologie al Data Product ha pertanto avuto anche una finalità esplorativa e di assessment, volta a identificare gli strumenti maggiormente idonei per l'evoluzione della soluzione. Alcuni dei servizi elencati potrebbero quindi non essere stati utilizzati nelle pipeline attualmente implementate, ma rappresentano opzioni tecnologiche valutate durante la fase di analisi o candidate per future evoluzioni del sistema.

La presenza di tali componenti nella documentazione non implica pertanto il loro impiego operativo nella versione corrente del Data Product, ma ne documenta la disponibilità e la potenziale adozione in successive fasi di sviluppo, ottimizzazione o ampliamento delle funzionalità previste.

Di seguito viene riportato l'albero elaborativo che descrive, per ogni tipologia di input, il flusso che è stato costruito e i servizi utilizzati per le varie trasformazioni.



Come è possibile vedere dagli alberi elaborativi, al fine di garantire l'accessibilità delle risorse vengono generati:

INPUT: *Video*
OUTPUT :

- Sottotitoli in Italiano e inglese
- Riassunto in Italiano e in inglese
- Audio riassunto in Italiano e in inglese
- Estrazione delle entità per il motore semantico

La descrizione approfondita delle pipeline viene riportata di seguito, nel paragrafo 4.

3.2. Eventuali tecnologie esterne a DPaaS

3.2.1. Modelli in locale

Durante la fase di realizzazione del Data Product sono stati valutati e testati i servizi di Intelligenza Artificiale messi a disposizione dall'ecosistema DPaaS, verificandone l'applicabilità rispetto agli specifici casi d'uso, agli obiettivi progettuali e ai requisiti di accessibilità definiti dagli Enti coinvolti. Per alcune funzionalità, al fine di supportare le attività di sviluppo e collaudo e di assicurare la necessaria continuità nella sperimentazione, si è ritenuto opportuno affiancare ai servizi disponibili modelli eseguiti su infrastrutture esterne gestite dal fornitore. Tale configurazione ha mantenuto la piena integrazione con i workflow di orchestrazione e con i meccanismi di pubblicazione della piattaforma DPaaS.
Per la maggior parte delle funzionalità sono stati adottati modelli funzionalmente equivalenti ai servizi presenti sulla piattaforma DPaaS. L'impiego di istanze locali durante la fase di sviluppo ha consentito di configurare, verificare e ottimizzare i modelli in relazione agli specifici contenuti e requisiti del Data Product. Per la componente di elaborazione multimodale è stato utilizzato **Qwen 3.5 Vision**, mentre la piattaforma DPaaS rende attualmente disponibile la versione **Qwen2**.
La scelta di utilizzare Qwen 3.5 Vision è stata determinata dalla necessità di elaborare contenuti multimodali, in particolare immagini di opere, fotografie, pagine manoscritte e documenti digitalizzati. Rispetto a Qwen2, la versione 3.5 introduce capacità più evolute di interpretazione visiva, una migliore integrazione tra informazioni testuali e contenuti grafici, una maggiore accuratezza nell'estrazione del contenuto documentale e una qualità superiore nella generazione di descrizioni contestualizzate. Tali caratteristiche si sono dimostrate particolarmente rilevanti per la realizzazione degli artefatti di accessibilità previsti dal progetto, quali descrizioni di immagini, audiodescrizioni, trascrizioni di manoscritti e contenuti derivati destinati alla consultazione museale.
Considerata la rilevanza di tale componente per il corretto funzionamento delle pipeline implementate, è stato richiesto ai responsabili della piattaforma DPaaS di rendere disponibile e ospitare una versione del modello **Qwen 3.5 Vision** all'interno dell'infrastruttura ufficiale. L'adozione del medesimo modello consentirebbe infatti di mantenere allineati i risultati prodotti durante la fase di sviluppo con quelli generati successivamente in esercizio.
Si evidenzia pertanto che, a valle della messa in produzione del Data Product, tutti i nuovi contenuti caricati o modificati verranno elaborati mediante i modelli disponibili sulla piattaforma DPaaS. Qualora il modello **Qwen 3.5 Vision** non venga reso disponibile all'interno dell'infrastruttura, le elaborazioni che richiedono capacità di analisi visuale verranno eseguite utilizzando **Qwen2**. In tale scenario, pur mantenendo invariata la funzionalità complessiva del sistema, gli artefatti generati potrebbero presentare un

livello di dettaglio, accuratezza e qualità descrittiva differente rispetto a quello ottenuto durante la fase di realizzazione e collaudo del Data Product, nella quale è stato utilizzato Qwen 3.5 Vision come modello di riferimento.

3.2.2. PyAnnote per la diarizzazione e l'identificazione dei parlanti

Tra le componenti esterne utilizzate dal Data Product è stato adottato il framework **PyAnnote**, sviluppato e mantenuto dalla comunità open source e distribuito tramite il repository Hugging Face. PyAnnote è utilizzato per la **speaker diarization**, ovvero il processo di identificazione e separazione automatica degli interventi dei diversi parlanti presenti all'interno di contenuti audio e video.

Nell'ambito del Data Product, PyAnnote viene impiegato durante l'elaborazione di file multimediali per supportare la generazione di **trascrizioni strutturate e sottotitoli**, consentendo di associare correttamente le diverse porzioni di parlato ai rispettivi speaker. Le informazioni prodotte dal framework vengono successivamente integrate con i servizi di Speech-to-Text della piattaforma e utilizzate per migliorare la qualità dei sottotitoli e la comprensione del contenuto audiovisivo.

PyAnnote non è una componente della piattaforma **DPaaS** e viene eseguito come servizio esterno all'interno dell'infrastruttura applicativa dedicata al Data Product. Lo scambio dati avviene tramite file audio temporanei e output strutturati contenenti i segmenti temporali associati ai diversi parlanti, che vengono successivamente acquisiti dai workflow di elaborazione orchestrati dalla piattaforma.

L'utilizzo di PyAnnote è stato necessario in quanto, al momento della realizzazione del Data Product, non erano disponibili all'interno dei servizi standard di DPaaS funzionalità equivalenti in grado di fornire un livello di accuratezza adeguato per la separazione automatica degli speaker nei contenuti audiovisivi. L'adozione del framework ha pertanto consentito di integrare una capacità specialistica indispensabile per la corretta produzione degli artefatti di accessibilità richiesti.

Dal punto di vista tecnico, l'utilizzo di PyAnnote introduce una dipendenza dall'ambiente di esecuzione Python e dalle librerie open source necessarie al funzionamento del framework. Inoltre, le performance e l'accuratezza del processo risultano influenzate dalla qualità della traccia audio elaborata e dalla disponibilità delle risorse computazionali necessarie all'esecuzione dei modelli di diarizzazione. L'utilizzo del framework è soggetto alle condizioni di licenza open source previste dai relativi componenti software e richiede il rispetto delle politiche di sicurezza e gestione dei dati adottate dall'infrastruttura che ospita il Data Product.

4. Pipeline di elaborazione e trasformazione

4.1. Pipeline di generazione di sottotitoli, riassunti e metadati semantici per contenuti audiovisivi

La pipeline è progettata per elaborare contenuti audiovisivi in formato **MP4** e **MP3**, trasformandoli in una serie di artefatti testuali, semantici e audio destinati a migliorare l'accessibilità e la valorizzazione del patrimonio culturale. Il processo combina tecnologie di riconoscimento del parlato, identificazione dei parlanti, generazione automatica di sottotitoli, sintesi e arricchimento semantico dei contenuti. Tutti gli artefatti vengono archiviati nel Data Product e resi disponibili attraverso il portale. Gli elementi contrassegnati con l'icona della mano possono essere revisionati e modificati dagli utenti autorizzati prima della pubblicazione.

4.1.1. Preparazione dell'audio e analisi preliminare

La pipeline prende in ingresso contenuti audiovisivi nei formati **MP4** e **MP3**. Nella fase iniziale la traccia audio viene estratta e convertita in formato WAV, rendendola compatibile con i servizi di elaborazione successivi.
Successivamente vengono eseguite due elaborazioni parallele:

- **Speaker Diarization**, mediante il framework **PyAnnote** eseguito localmente, per identificare i diversi parlanti presenti nella registrazione;
- **Voice Activity Detection**, mediante il modello **Silero VAD**, per rilevare automaticamente i periodi di silenzio e le porzioni effettivamente contenenti parlato.

Queste informazioni consentono di segmentare il contenuto audio e di migliorare la qualità della trascrizione prodotta nelle fasi successive.

4.1.2. Trascrizione e generazione dei sottotitoli

Le informazioni ottenute dalla diarizzazione e dall'identificazione dei silenzi vengono elaborate dal servizio **Speech-to-Text (CPA)** basato su **Whisper**, eseguito localmente. Il risultato è una trascrizione temporizzata del contenuto parlato, arricchita con informazioni relative ai parlanti identificati e ai metadati della registrazione.
A supporto del processo interviene inoltre il modello **Llama 3.1**, eseguito presso l'infrastruttura **CINECA**, che contribuisce alla normalizzazione e all'organizzazione del testo estratto.
Da questa elaborazione vengono generati i **sottotitoli in italiano** (*sottotitoli_it*), che rappresentano l'artefatto principale della pipeline.

I sottotitoli possono essere revisionati e corretti dagli operatori autorizzati tramite il portale, in modo da garantire accuratezza terminologica, correttezza dei nomi propri e qualità complessiva del contenuto.

4.1.3. Traduzione dei sottotitoli in lingua inglese

I sottotitoli italiani vengono successivamente elaborati dai modelli **Llama 3.1** e **Qwen 3.5 Text** per produrre una versione inglese dei contenuti.
Da questa fase deriva l'artefatto:

- **sottotitoli_en**: sottotitoli sincronizzati in lingua inglese.

L'obiettivo è consentire la fruizione del materiale audiovisivo anche da parte di utenti internazionali mantenendo la coerenza semantica con il contenuto originale.

4.1.4. Generazione di riassunti accessibili

A partire dai sottotitoli italiani, la pipeline genera automaticamente un **riassunto del contenuto audiovisivo** (*summary_it*).
Il riassunto viene prodotto mediante **Llama 3.1**, con il supporto di **Qwen 3.5 Text**, e sintetizza i principali argomenti affrontati nella registrazione.
Poiché la qualità di questo contenuto è particolarmente importante per la comprensione rapida del materiale audiovisivo, il sistema consente agli operatori di modificare e validare il riassunto direttamente dal portale.

4.1.5. Produzione del riassunto in lingua inglese

Dalla versione italiana viene generata una corrispondente versione inglese:

- **summary_en**: riassunto in lingua inglese.

Il contenuto mantiene le informazioni essenziali presenti nella registrazione e garantisce una
consultazione agevole anche da parte di utenti non italofoni.

4.1.6. Generazione delle versioni audio dei riassunti

I riassunti vengono successivamente convertiti in tracce audio mediante il servizio **Text-to-Speech (CPA)** basato sul motore **Kokoro**, eseguito localmente.
Sono pertanto prodotti i seguenti artefatti:

- **summary_it_tts**: audio del riassunto in italiano;
- **summary_en_tts**: audio del riassunto in inglese.

Questi contenuti consentono modalità alternative di accesso alle informazioni attraverso l'ascolto.

4.1.7. Estrazione delle entità e dei concetti

La pipeline prevede una fase di arricchimento semantico finalizzata alla valorizzazione dei contenuti e alla loro integrazione nei sistemi di ricerca e navigazione.
Partendo dai sottotitoli italiani, il modello **Llama 3.1** identifica innanzitutto le **menzioni** presenti nel testo, ovvero riferimenti a persone, luoghi, eventi, organizzazioni o altre entità rilevanti.
Le menzioni individuate vengono successivamente utilizzate per estrarre un insieme di **concetti** che rappresentano gli argomenti principali trattati nel contenuto audiovisivo.
Questa elaborazione consente di costruire una rappresentazione semantica del documento e di migliorare significativamente le funzionalità di ricerca, classificazione e correlazione dei contenuti.

4.1.8. Generazione dei tag semantici

Le informazioni derivate dall'analisi delle menzioni e dei concetti vengono ulteriormente elaborate da **Llama 3.1** per produrre un insieme strutturato di tag descrittivi.
L'artefatto risultante:
tag_it
contiene parole chiave e classificazioni tematiche che descrivono il contenuto audiovisivo.
I tag vengono utilizzati per facilitare la navigazione del patrimonio digitale, supportare ricerche avanzate e creare collegamenti con altri contenuti correlati presenti nel sistema.

4.1.9. Produzione dei metadati di interoperabilità

Dall'insieme dei tag generati vengono creati due ulteriori artefatti destinati all'integrazione con sistemi esterni e piattaforme semantiche.

- **Generazione del GraphDB**

La pipeline produce una rappresentazione strutturata del contenuto sotto forma di grafo semantico:

- **graphdb_ttl**

Tale artefatto consente l'integrazione con sistemi RDF e knowledge graph, favorendo attività di interrogazione semantica e correlazione tra contenuti appartenenti a differenti collezioni.

- **Generazione del Manifest**

- Parallelamente viene generato un:
  - **manifest_iiif**

contenente i metadati necessari per la pubblicazione e l'interoperabilità secondo gli standard di fruizione digitale adottati dal patrimonio culturale.

4.1.10. Revisione e validazione dei contenuti

La pipeline prevede la supervisione umana sui contenuti maggiormente rilevanti per l'esperienza utente. In particolare possono essere modificati e validati:

- i sottotitoli in italiano (*sottotitoli_it*);
- il riassunto in italiano (*summary_it*).

Questa attività garantisce la qualità finale degli artefatti pubblicati e consente di integrare i controllo umano nei punti più critici del processo di generazione automatica.

4.1.11. Pubblicazione degli artefatti

Al termine dell'elaborazione il Data Product rende disponibili:

- Sottotitoli in italiano (*sottotitoli_it*);
- Sottotitoli in inglese (*sottotitoli_en*);
- Riassunto in italiano (*summary_it*);
- Riassunto in inglese (*summary_en*);
- Audio del riassunto in italiano (*summary_it_tts*);
- Audio del riassunto in inglese (*summary_en_tts*);
- Tag semantici (*tag_it*);
- Grafo semantico RDF (*graphdb_ttl*);
- Manifest IIIF (*manifest_iiif*).

Attraverso questa pipeline il contenuto audiovisivo viene trasformato in un insieme strutturato di informazioni accessibili, ricercabili e interoperabili, migliorando sia la fruizione da parte degli utenti finali sia le capacità di catalogazione e valorizzazione del patrimonio culturale digitale.

5. Esposizione e distribuzione

Il Data Product viene pubblicato nel catalogo della piattaforma DPaaS e reso fruibile attraverso un portale web e mediante API REST in sola lettura. Consente la consultazione strutturata dei contenuti audiovisivi conferiti, dei relativi metadati e degli artefatti di accessibilità generati dalla pipeline.

L'esposizione è articolata su due livelli funzionali:

- fruizione tramite portale web, destinata agli utenti finali e agli operatori autorizzati;
- fruizione machine-to-machine tramite API, destinata al portale, al Backend for Frontend e agli altri sistemi consumer autorizzati.

Il portale consente di accedere ai contenuti audiovisivi, ai metadati, ai sottotitoli, ai riassunti, alle versioni audio e agli elementi semantici. La read-only API costituisce il canale di interoperabilità: espone oggetti, metadati, classificazioni, raggruppamenti, relazioni e artefatti generati, senza consentire ai consumer di modificare direttamente i dati.

5.1. Canali di accesso

L'accesso ai contenuti e ai servizi del Data Product è reso disponibile attraverso i seguenti canali:

- portale web dedicato alla consultazione;
- API REST read-only tramite HTTPS.

5.2. Utenti e sistemi destinatari

Il Data Product è rivolto alle seguenti categorie di utenti e sistemi, in funzione delle diverse modalità di consultazione e riutilizzo dei contenuti:

- cittadini, studiosi e visitatori interessati al patrimonio culturale;
- persone che beneficiano di sottotitoli, sintesi, contenuti audio o modalità alternative di consultazione;
- utenti internazionali, grazie ai contenuti in inglese;
- operatori dell'Ente e operatori incaricati della revisione e validazione;
- portale applicativo e relativo Backend for Frontend;
- sistemi applicativi autorizzati che interrogano contenuti e metadati tramite API;
- enti culturali, ricercatori e sviluppatori interessati al riutilizzo, nel rispetto delle licenze applicabili.

5.3. Autenticazione e autorizzazione

La consultazione pubblica tramite portale non richiede autenticazione. Revisione, modifica e validazione sono riservate agli operatori autenticati e autorizzati in base ai ruoli applicativi. Per l'accesso machine-to-machine, le API utilizzano le credenziali applicative previste dall'ambiente di esercizio e il codice identificativo del Data Product. Durata delle credenziali ed eventuali restrizioni di rete sono configurate sul livello infrastrutturale di esposizione.

5.4. Formati disponibili

I dati e i metadati restituiti dalle API sono in formato JSON. Le risorse comprendono:

- contenuti video in formato MP4 e audio in formato MP3;
- sottotitoli e riassunti in italiano e inglese;
- tracce audio dei riassunti;
- tag, metadati tecnici e descrittivi e informazioni semantiche strutturate;
- cataloghi, esposizioni, gruppi tematici, geolocalizzazioni e relazioni tra oggetti, ove presenti.

5.5. Modalità di aggiornamento

L'aggiornamento avviene mediante pipeline schedulate. Le pipeline verificano periodicamente la presenza di nuovi contenuti o contenuti non ancora elaborati, attivano i task previsti e registrano avanzamento, esito ed eventuali errori. Gli artefatti vengono associati all'oggetto sorgente e pubblicati al completamento delle elaborazioni.

Sono previsti meccanismi di ripetizione in caso di errori temporanei e modalità di ripresa controllata. Se l'aggiornamento riguarda solo dati o artefatti, senza modificare modello dati, struttura delle API o pipeline, non è necessaria una nuova versione. Modifiche significative a modello informativo, contratti API, elaborazioni o modalità di esposizione richiedono invece un nuovo versionamento.

5.6. Limiti di utilizzo

Le ricerche utilizzano paginazione per limitare la dimensione delle risposte. Limiti su frequenza delle chiamate, concorrenza, dimensione delle risposte e timeout sono definiti nella configurazione dell'ambiente di esercizio.

5.7. Livelli di servizio

Il servizio dispone di funzioni tecniche per monitorare stato applicativo e disponibilità. I livelli quantitativi di servizio sono definiti nell'accordo di esercizio della piattaforma.

5.8. Accesso alla documentazione

Il Data Product è corredato dalla documentazione prevista per il catalogo DPaaS: descrizione pubblica, manuale utente, manuale operativo, specifiche tecniche di elaborazione e documentazione delle API. La documentazione pubblica descrive la fruizione a livello funzionale; il dettaglio delle operazioni, dei parametri e dei modelli informativi è disponibile secondo il processo di richiesta di consumo previsto dal catalogo.

5.9. Condizioni di licenza e riutilizzo

Dati, metadati e artefatti sono riutilizzabili secondo la licenza indicata nella scheda del Data Product. Restano applicabili i diritti e i vincoli associati ai contenuti conferiti dall'Ente e agli artefatti derivati. La licenza definitiva, gli obblighi di attribuzione e le eventuali restrizioni sono quelli pubblicati dal titolare nel catalogo.

5.10. Fruizione tramite API

L'API fornisce accesso programmatico, in sola lettura, agli oggetti audiovisivi e ai relativi contenuti accessibili. Consente alle applicazioni consumer di ricercare, recuperare e navigare i contenuti, insieme ai metadati descrittivi e tecnici, alle classificazioni e agli artefatti generati.

5.10.1. Tipologie di informazioni interrogabili

Attraverso le API è possibile interrogare le seguenti tipologie di informazioni associate alle risorse pubblicate nel Data Product:

- oggetti digitali e relativi identificativi;
- contenuti audio e video e riferimenti ai file associati;
- nome, descrizione, autore o soggetto responsabile, ove presenti;
- datazione, lingua, tipologia, sottotipo e stato;
- cataloghi, esposizioni, gruppi tematici, tag e geolocalizzazioni, ove presenti;
- relazioni tra oggetti e tra contenuti originali e artefatti derivati;
- sottotitoli, riassunti, versioni audio e metadati semantici, secondo disponibilità.

Nel loro insieme, tali informazioni consentono ai sistemi consumer di ricostruire il contesto descrittivo, tecnico e semantico delle risorse disponibili e dei relativi artefatti.

5.10.2. Principali operazioni disponibili

Le principali operazioni di consultazione rese disponibili dall'API sono le seguenti:

- ricerca degli oggetti con filtri e paginazione;
- recupero del dettaglio di un oggetto e dei relativi metadati;
- consultazione degli oggetti collegati e delle relazioni;
- consultazione di stati, lingue, tipi, sottotipi, tag e tipi di relazione;

- consultazione aggregata di cataloghi, esposizioni, gruppi tematici e geolocalizzazioni.

Le operazioni sono esclusivamente di consultazione. Revisione e validazione sono gestite dai servizi riservati agli operatori e non fanno parte dell'interfaccia pubblica read-only.

5.10.3. Ricerca, filtro, ordinamento e paginazione

La ricerca combina i criteri supportati dal modello informativo, tra cui identificativo o codice, nome, lingua, stato, tipologia e sottotipo, data di caricamento, catalogo, esposizione, gruppo tematico, geolocalizzazione e tag. I tag possono essere combinati richiedendo la corrispondenza con almeno uno o con tutti i valori indicati. I risultati possono essere ordinati e sono restituiti in forma paginata.

5.10.4. Recupero del dettaglio di una risorsa

Il dettaglio viene recuperato tramite l'identificativo univoco. La risposta comprende le informazioni principali e, tramite operazioni dedicate, metadati, cataloghi, esposizioni, gruppi tematici, tag, geolocalizzazioni e relazioni disponibili. Le relazioni permettono di individuare contenuti collegati e artefatti derivati dalla risorsa originaria.

5.10.5. Servizi di aggregazione e analisi

L'API rende disponibili viste aggregate per cataloghi, esposizioni, gruppi tematici e geolocalizzazioni, con conteggio degli oggetti associati quando previsto. Le funzioni supportano la navigazione e non costituiscono un sistema di Business Intelligence né forniscono KPI o reporting multidimensionale.

5.10.6. Formati delle richieste e delle risposte

Le richieste utilizzano HTTPS, parametri di percorso per identificare la risorsa e parametri di query per filtri, lingua, ordinamento e paginazione. Le risposte informative sono in JSON; il recupero dei contenuti binari è gestito dai servizi applicativi di esposizione. La struttura completa è descritta nella documentazione tecnica delle API.

5.10.7. Modalità di autenticazione

L'accesso machine-to-machine usa le credenziali applicative previste dall'ambiente di esercizio e il codice del Data Product. Le credenziali dei componenti interni non sono trasferite ai consumer. Le funzioni riservate agli operatori richiedono autenticazione e ruolo autorizzato.

5.10.8. Limiti di chiamata

La paginazione controlla la quantità di dati restituita. Frequenza, quote, concorrenza, dimensione della pagina, timeout e dimensione massima delle risposte sono definiti nell'ambiente di esercizio.

5.10.9. Gestione degli errori

L'API utilizza codici di stato standard. Le categorie di errore comprendono parametri mancanti o non validi, risorsa non esistente, autenticazione assente o non valida, accesso non autorizzato, superamento dei limiti configurati, indisponibilità temporanea ed errori interni. Le risposte di errore sono in JSON. Gli errori delle pipeline sono registrati separatamente per consentire monitoraggio e ripresa.

5.11. Limitazioni

Il Data Product è costituito dai contenuti audiovisivi e dai metadati conferiti nel perimetro progettuale. La completezza dipende dalla disponibilità, qualità e correttezza dei materiali ricevuti.

La qualità di sottotitoli, traduzioni, riassunti, versioni audio e metadati semantici può risentire della qualità della registrazione, del rumore, della sovrapposizione tra voci, della chiarezza del parlato, della correttezza dei metadati e della presenza di terminologia specialistica o nomi propri. Per questo è prevista la revisione degli artefatti principali da parte di operatori autorizzati.

Il Data Product ha natura non real-time. Gli aggiornamenti avvengono tramite caricamento dei contenuti e successive esecuzioni schedulate. Tra disponibilità della fonte e pubblicazione può intercorrere un intervallo legato alla frequenza di esecuzione, alla durata delle elaborazioni e alle attività di revisione.

5.12. Nota sulle condivisioni dei dati

Pubblicazione, condivisione e riutilizzo sono regolati dalla licenza e dalle condizioni indicate nella scheda del Data Product. L'accesso pubblico riguarda esclusivamente i contenuti dichiarati pubblicabili dal titolare; revisione e gestione sono riservate agli operatori autorizzati e l'accesso machine-to-machine è subordinato all'abilitazione prevista dall'ambiente di esercizio.

Restano in capo al titolare la verifica di diritti, vincoli, dati personali o limitazioni associati ai contenuti audiovisivi e ai metadati. La condivisione può riguardare il contenuto completo, i soli metadati o gli artefatti derivati, secondo le condizioni definite dal titolare e pubblicate nel catalogo.

In particolare, si segnala che il materiale Liber Liber e' vincolato da diritti di riproduzione, pertanto non possono essere condivisi senza l'avallo di chi ne detiene la proprietà intellettuale.

6. Possibili evoluzioni future

L'architettura realizzata è stata concepita in modo modulare, al fine di consentire future evoluzioni sia dal punto di vista tecnologico sia dal punto di vista funzionale. Una prima area di miglioramento riguarda la possibilità di sostituire o aggiornare i modelli di Intelligenza Artificiale attualmente impiegati con versioni più recenti o con modelli differenti, qualora disponibili sulla piattaforma DPaaS o ritenuti maggiormente performanti rispetto agli specifici casi d'uso. Tale approccio consentirebbe di beneficiare dei continui progressi tecnologici nel campo dell'AI senza richiedere modifiche significative ai processi di orchestrazione e pubblicazione del Data Product.

Un'ulteriore evoluzione riguarda la trasformazione delle pipeline attualmente sviluppate nell'ambito del progetto in veri e propri **servizi CPA (Common Processing Activities)** integrati all'interno dell'ecosistema DPaaS. In questo scenario, le logiche di generazione degli artefatti verrebbero rese disponibili come servizi condivisi e riutilizzabili, consentendo agli Enti di invocare un unico servizio applicativo per ottenere direttamente gli artefatti di accessibilità richiesti. Questa evoluzione permetterebbe di semplificare i processi di integrazione, ridurre la complessità operativa e favorire il riuso delle componenti sviluppate anche in altri Data Product o progetti futuri.

Dal punto di vista funzionale, potrebbe inoltre essere introdotto un meccanismo di **aggiornamento dinamico del Data Product**, che consenta l'acquisizione automatica di nuovi contenuti successivamente alla pubblicazione iniziale. In tale scenario, l'aggiunta di nuovi file di input da parte degli Enti attiverebbe automaticamente le pipeline di elaborazione e la successiva produzione degli artefatti derivati, mantenendo il Data Product costantemente aggiornato senza richiedere interventi manuali di riconfigurazione.

Tale funzionalità non è stata implementata nella presente versione della soluzione a causa di alcune limitazioni tecnologiche e operative riscontrate nell'attuale piattaforma DPaaS, che non consentono una gestione completamente automatizzata dell'evoluzione del contenuto del Data Product una volta completato il ciclo di pubblicazione iniziale. Qualora tali limitazioni venissero superate in future versioni della

piattaforma, sarebbe possibile estendere l'architettura per supportare scenari di aggiornamento continuo e incrementale dei contenuti, migliorando ulteriormente la capacità del sistema di adattarsi all'evoluzione del patrimonio digitale degli Enti aderenti.