

# L'infrastruttura software per il patrimonio culturale (ISPC) come abilitatore di un Ecosistema digitale nazionale del patrimonio culturale

«DigItalia» 1-2023  
DOI: 10.36181/digitalia-00059

**Luigi Cerullo — Antonella Negri**

*Istituto centrale per la digitalizzazione del patrimonio culturale - Digital Library*

*L'Infrastruttura Software per il Patrimonio Culturale (ISPC) rappresenta il primo spazio dati nazionale della cultura in grado di ospitare in sicurezza tutto il patrimonio digitale del Paese, consentendo agli Enti flessibilità nel modello di adesione e totale autonomia nella scelta di condivisione.*

*Nella ISPC il patrimonio è valorizzato e arricchito per mezzo di algoritmi innovativi e tecniche di Intelligenza Artificiale e viene messo in correlazione con tutte le risorse culturali digitali presenti, sia all'interno di ciascun dominio di appartenenza (storico, artistico, archeologico, demo-etnoantropologico, archivistico, librario) sia tra domini diversi.*

*La ISPC è una infrastruttura con una doppia anima: un sistema di servizi di cooperazione applicativa pensati per potenziare funzionalmente e tecnologicamente i sistemi informativi del Ministero della Cultura, degli enti territoriali e degli istituti culturali in generale, in un'ottica "Business-to- Institution"; un laboratorio di sviluppo per la creazione di nuovi servizi basati sui dati, pensato per sviluppatori, imprese culturali e start-up, in una prospettiva "Business-to-Business".*

*Con un'architettura scalabile e sicura basata su soluzioni Cloud che offrono performance e livelli di disponibilità adeguati, ISPC si candida a sostenere credibilmente il ruolo di abilitatore nell'offerta di servizi digitali per il panorama culturale italiano, anche nella dimensione europea.*

## La Digital Library e il Piano nazionale di digitalizzazione del patrimonio culturale

Il **Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library**, all'interno del Ministero della Cultura (MiC), realizza, coordina e promuove il Piano nazionale di digitalizzazione del patrimonio culturale (PND), che costituisce la visione strategica con la quale il Ministero intende promuovere e organizzare il processo di trasformazione digitale nel quinquennio 2022-2026.

Si rivolge in prima istanza ai musei, agli archivi, alle biblioteche, alle soprintendenze, agli istituti e ai luoghi della cultura pubblici che conservano, tutelano, gestiscono e/o valorizzano beni culturali, ma si configura anche come contesto meto-

dologico, intellettuale e tecnico di riferimento per la realizzazione degli obiettivi del PNRR.

Il PND è frutto di un processo di condivisione e confronto con diverse istituzioni culturali: non è un documento prescrittivo, ma costituisce un utile riferimento metodologico per tutte le istituzioni e per i professionisti che, sia in ambito pubblico che privato, si riconoscono nei valori in esso enunciati. Per le sue caratteristiche si può definire un documento “aperto”, perché liberamente accessibile<sup>1</sup>, “dinamico”, in quanto periodicamente aggiornato, “condiviso”, perché frutto di un processo di partecipazione articolato in più livelli e suscettibile di gradi differenti di applicazione e scalabilità<sup>2</sup>.

### La struttura e la dimensione di processo del PND

Il PND è articolato in tre sezioni, tra loro collegate in una dimensione di processo: la visione, che prefigura la trasformazione e le opportunità del cambiamento, indicando gli obiettivi a lungo termine; la strategia, che definisce il percorso per implementare e conseguire gli obiettivi; le linee guida, quali strumenti operativi che supportano la pianificazione e l'esecuzione delle attività legate alla digitalizzazione del patrimonio culturale.

Gli strumenti che accompagnano il PND indicano e suggeriscono metodologie e procedure per l'attuazione dei processi individuati nella strategia e si possono sinteticamente rappresentare con il processo del ciclo di vita dei dati, ovvero: “produci i dati/gestisci i dati/fai circolare i dati/crea servizi dai dati/valuta la maturità digitale”. Le Linee guida sono pertanto costituite da cinque documenti:

1. *Linee guida per la digitalizzazione del patrimonio culturale*: definiscono approcci e procedure per la creazione, la metadattazione e l'archiviazione degli oggetti digitali del patrimonio culturale analogico;
2. *Linee guida per la redazione del piano di gestione dei dati (Data Management Plan)*: descrivono e analizzano come archiviare, condividere e conservare i dati provenienti dai progetti di digitalizzazione del patrimonio culturale e dalle banche dati esistenti. Sono corredate da FAQ sui dati aperti;
3. *Linee guida per l'acquisizione, la circolazione e il riuso delle riproduzioni dei beni culturali in ambiente digitale*: inquadrano il contesto normativo di riferimento, fornendo strumenti operativi che permettono di individuare gli ambiti legislativi corretti, in funzione delle tipologie di bene culturale e delle modalità di riproduzione e acquisizione di immagini e materiali audiovisivi, in relazione alle diverse finalità di utilizzo e ai diritti d'autore eventualmente gravanti sui beni e sulle ripro-

<sup>1</sup> <https://docs.italia.it/italia/icdp/>.

<sup>2</sup> <https://partecipa.gov.it/processes/piano-nazionale-digitalizzazione-patrimonio-culturale/f/144/>.

duzioni. L'obiettivo è porre le istituzioni e gli utenti nelle condizioni di distinguere senza equivoci i limiti e le possibilità di riutilizzo delle riproduzioni rese disponibili in rete dagli istituti.

4. *Linee guida per la classificazione di prodotti e servizi digitali, processi e modelli di gestione*: individuano e descrivono le diverse tassonomie e mappature di processi e servizi;

5. *Introduzione alla metodologia per la valutazione della maturità digitale degli istituti culturali*: illustra i modelli (*digital maturity model*) adottabili per gli istituti culturali, evidenziandone le opportunità di applicazione al patrimonio pubblico e proponendo l'utilizzo, a livello nazionale, della metodologia valutativa più appropriata per consentire alle istituzioni di comprendere con chiarezza il proprio livello iniziale e governare più efficacemente i processi di transizione digitale.

Le linee di azione strategica rappresentano gli intenti del Ministero della Cultura per innescare, facilitare e accelerare i cambiamenti strutturali che l'innovazione tecnologica abilita in tutto il settore; sono suddivise e descritte secondo tre ambiti: le tecnologie abilitanti, i processi e le persone. Le tecnologie abilitano i processi, che sono governati dalle persone in una prospettiva di crescita consapevole.

### **Il contesto tecnologico abilitante**

La prima linea di azione è dunque creare un contesto tecnologico abilitante, che si basi sul disegno di una strategia per la gestione dei dati e delle risorse digitali del patrimonio culturale, coerente con le direttive nazionali ed europee: i dati rappresentano il fulcro della trasformazione digitale, per questo sono stati progettati interventi per la creazione di un'infrastruttura cloud hardware e software a disposizione di tutto l'ecosistema digitale della cultura, da realizzare sia attraverso la creazione di piattaforme di servizio a livello nazionale, sia stimolando e sostenendo sviluppatori e imprese per la produzione di servizi digitali da diffondere sul mercato, in grado di connettere il patrimonio culturale con le persone e le industrie creative.

I processi chiave della trasformazione digitale, in coerenza con i cosiddetti principi FAIR (rintracciabilità, accessibilità, interoperabilità e riusabilità), possono essere individuati nella digitalizzazione e nella gestione del ciclo di vita della risorsa digitale, nelle politiche aggiornate per l'accesso, la valorizzazione e il riuso delle riproduzioni digitali dei beni culturali e nel design dei servizi e dei modelli per la creazione di valore culturale, sociale ed economico. Ma per consentire alla transizione digitale del patrimonio di produrre valore culturale e sociale è necessario che al centro del processo di cambiamento ci siano le persone: l'obiettivo può essere raggiunto con la formazione e l'aggiornamento delle competenze attraverso un programma di apprendimento permanente (*lifelong learning*), ma anche e so-

prattutto tramite la costruzione di una rete di scambi attorno al patrimonio culturale, in grado di incentivare il riuso delle risorse digitali e i meccanismi di co-creazione e di produzione di valore.

Nello schema di seguito riportato, nel quale sono evidenziate le relazioni esistenti tra i diversi progetti in cui si articola l'investimento PNRR M1C3 1.1, l'Infrastruttura Software per il Patrimonio Culturale (ISPC) occupa una posizione centrale come piattaforma abilitante di dati, processi e servizi.

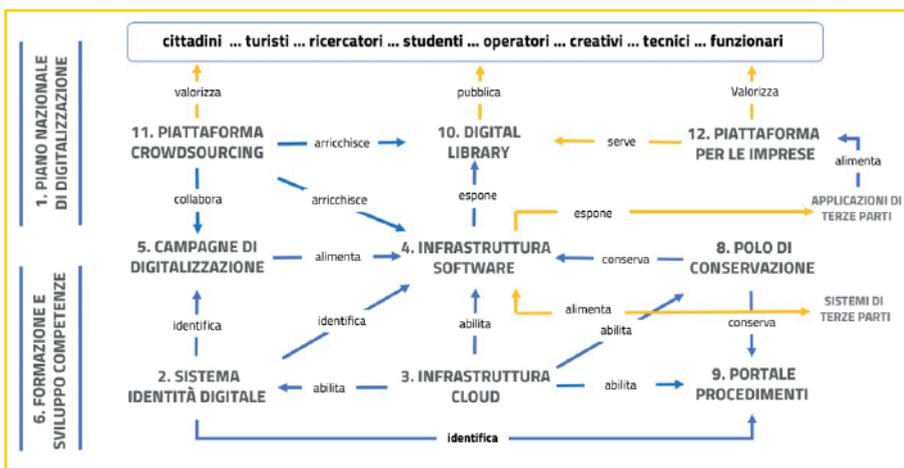


Figura 1. Relazioni tra i progetti/sub-investimenti in cui si articola l'Investimento PNRR M1C3 1.1

### L'infrastruttura software per il patrimonio culturale (ISPC)

L'ISPC è un complesso sistema di servizi basati sulle più moderne ed avanzate tecnologie orientate al Cloud che:

- supporta e potenzia i sistemi per la valorizzazione del patrimonio che operano a livello nazionale e territoriale;
- offre funzioni relative alla gestione e all'arricchimento delle risorse digitali, basate sia su modelli e schemi predefiniti (motori a regole e ontologie) sia su algoritmi di intelligenza artificiale (AI);
- espone un vasto catalogo di API di cooperazione applicativa (in lettura e scrittura) relative sia a dati di dominio sia a viste cross-dominio;
- mette a disposizione una piattaforma di sviluppo che consente di ottenere nuovi servizi complessi, personalizzati e adattabili a esigenze diversificate mediante l'orchestrazione di servizi di base.

La ISPC è destinata a rappresentare il primo spazio dati nazionale della cultura in

grado di ospitare in sicurezza tutto il patrimonio digitale del paese, consentendo agli enti che vi cooperano flessibilità nel modello di adesione e totale autonomia nelle proprie scelte progettuali.

Il soggetto che sottoscrive una convenzione con la ISPC può candidare, infatti, uno o più sistemi di produzione del dato e/o di accesso, scegliendo tra due modalità di adesione:

- *integrazione*: quando un sistema conferisce tutti i propri dati che vengono, quindi, “ospitati”, memorizzati e processati all’interno della ISPC, continuando però a gestire interamente i *workflow* di catalogazione dei beni/entità culturali<sup>3</sup>;
- *federazione*: quando un sistema condivide solo i dati descrittivi e mantiene al proprio interno le risorse digitali, ma referenziandole<sup>4</sup>.

## Il Catalogo dei Servizi della ISPC

ISPC mette a disposizione una serie di servizi abilitanti, descritti in uno specifico documento denominato “Catalogo dei Servizi”, per l’interoperabilità e la cooperazione applicativa nei confronti dei sistemi e soggetti ad essa aderenti nell’ambito dell’ecosistema digitale del patrimonio culturale delineato dal Piano nazionale di digitalizzazione del patrimonio culturale<sup>5</sup>. Il Catalogo è articolato in quattro macro-aree di servizi (Fig. 2):

1. Gestione e processamento Asset digitali
2. Grafi di conoscenza di dominio
3. Grafi di conoscenza cross-dominio
4. Platform as a Service (PaaS) come Data Fabric



Figura 2. *Catalogo dei Servizi – Livello I e II*

<sup>3</sup> In tal caso l’Infrastruttura funge sia da repository per la memorizzazione in sicurezza delle proprie risorse digitali, sia come fornitore di servizi per la gestione e fruizione di quanto acquisito.

<sup>4</sup> In questo scenario, ISPC agisce almeno come fornitore di tutti i servizi di content processing avanzato e di accesso ai Grafi di conoscenza arricchiti.

<sup>5</sup> Sull’ecosistema digitale si rimanda ai vari riferimenti riportati nel PND:

[https://digitallibrary.cultura.gov.it/wp-content/uploads/2023/04/PND\\_V1\\_1\\_2023\\_v2.pdf](https://digitallibrary.cultura.gov.it/wp-content/uploads/2023/04/PND_V1_1_2023_v2.pdf).

### 1. Gestione e processamento Asset digitali

Questa prima macro-area mette a disposizione (dei soli Sistemi integrati) gli strumenti necessari per la conservazione, l'elaborazione e la presentazione degli asset digitali, tenendo conto della loro natura, del loro formato e delle politiche di protezione che ciascuno di essi vorrà definire per la fruizione. Per la gestione degli asset digitali i sistemi cooperanti potranno estendere le proprie funzionalità sia attraverso l'invocazione diretta delle API primitive di questa componente funzionale sia disponendo di grafiche utente (Single Page App e Widget di frontend) direttamente e facilmente integrabili nelle piattaforme abilitate (nello schema di figura 2 rappresentato dal SaaS Teca multimediale).

Di seguito, si rende una sintetica panoramica dei principali macro-servizi esposti come API standard:

*1.1 Funzioni di Digital Asset Management (DAM):* servizi che offrono, ai Sistemi integrati, funzioni CRUD sulle risorse digitali appartenenti alla propria Tenancy<sup>6</sup>.

*1.2 Content processing di base (o media processing):* servizi che offrono ai Sistemi integrati un insieme di tecnologie che consentono l'avvio automatico delle pipeline di rendition, ovvero la trasformazione dei bitstream originari orientata alla fruizione e alla distribuzione online dei contenuti digitali (in dettaglio: 1.2.1 Renditioning di testi elettronici, 1.2.2 Renditioning di immagini raster, 1.2.3 Renditioning di audio, 1.2.4 Renditioning di video, 1.2.5 Renditioning di file immagine 3D, 1.2.6 Renditioning di file immagine 3D fotosferico, 1.2.7 Cutting/cropping file immagine, 1.2.8 Cutting/cropping file audiovisivo, 1.2.9 Image/audio/video enhancement, 1.2.10 Mirroring file immagine, eccetera);

*1.3 Content processing avanzato:* servizi che permettono l'elaborazione delle risorse digitali originariamente acquisite sia per generare nuovi media sia entità individuate nel contenuto dei media (in dettaglio: 1.3.1 Tagging automatico dei media, 1.3.2 Named Entity Recognition (NER) per classificazione automatica, 1.3.3 Optical Character Recognition (OCR), 1.3.4 Intelligent Character Recognition (ICR), 1.3.5 Geocoding, 1.3.6 Speech to Text, 1.3.7 Text to Speech, 1.3.8 Language Translation, 1.3.9 AI image/audio/video enhancement);

*1.4 Media player:* servizi offerti ai sistemi di produzione del dato e/o accesso, che permettono di riprodurre le risorse digitali del bene in base allo specifico ruolo dell'utente, consentendogli di sfruttare protocolli internazionali quali International Image Interoperability Framework (IIIF), per operare in diversi modi sui media facilitandone l'interoperabilità e il riuso (in dettaglio: 1.4.1 Viewer audiovisivo, 1.4.2 Player 3D, 1.4.3 Viewer 3D fotosferico, 1.4.4 Viewer testo elettronico, 1.4.5 Viewer IIIF);

<sup>6</sup> Area logica di pertinenza (visibilità e operatività sul dato) associata ai singoli Content Owner/Tenant, creata automaticamente nel Data Lake per consentire la gestione diretta delle Risorse digitali/Contentitori.

*1.5 Media streaming server:* servizi offerti ai sistemi di produzione del dato e/o accesso, che permettono di processare (encoding, trascodifica, cifratura e distribuzione) e di fruire in particolare dei media audiovisivi (in dettaglio: 1.5.1 Acquisizione, 1.5.2 Sottocampionamento ed encoding, 1.5.3 Trascodifica, 1.5.4 Dynamic Adaptive Streaming, 1.5.5 Rendition, 1.5.6 DRM Encryption, 1.5.7 Distribuzione).

## *2. Grafi di conoscenza di dominio*

Questa seconda macro-area è relativa ai servizi per la gestione dei grafi di conoscenza di dominio, ovvero rappresentazioni di relazioni semantiche tra entità di uno specifico dominio della conoscenza e descrizione del patrimonio (basati sugli standard guida, di struttura e di contenuto elaborati dalle principali agenzie internazionali e nazionali per i domini descrittivi archivistico, bibliografico e ABAP).

ISPC offre, infatti, ai sistemi integrati e/o federati la possibilità di poter interagire - a seconda del dominio descrittivo di appartenenza del sistema - con i grafi di conoscenza di dominio secondo due modelli di cooperazione applicativa:

1° livello di cooperazione: modalità di colloquio che include solo servizi di interrogazione e recupero delle informazioni dal grafo di interesse (saranno esposti almeno i seguenti macro-servizi: 2.1.1 Visualizza, 2.1.2 Cerca, 2.1.3 Recupera);

2° livello di cooperazione: modalità di colloquio che include sia servizi di interrogazione e recupero delle informazioni dal grafo sia servizi di scrittura diretta delle entità e delle relazioni che lo costituiscono (saranno esposti almeno i seguenti macro-servizi: 2.2.1 Crea, 2.2.2 Modifica, 2.2.3 Elimina, 2.2.4 Fonde, 2.2.4 Importa, 2.2.5 Allinea).

## *3. Grafi di conoscenza cross-dominio*

I grafi di conoscenza cross-dominio sono sviluppati e alimentati attraverso l'adozione e la sperimentazione di più logiche e tecnologie di integrazione dei dati. Le viste integrate, indipendente dai domini della descrizione standardizzati dagli Istituti centrali del MiC, saranno tante quante le tecnologie di integrazione sperimentate nella prima fase di implementazione della ISPC.

- un motore semantico basato su tecnologie NLP (Natural Language Processing);
- un motore semantico ibrido basato su ontologie e ML (Machine Learning).

In relazione a questi grafi, ISPC offre principalmente servizi d'interrogazione e recupero delle informazioni (rappresentati dal solo 1° livello di cooperazione), ai fini di una più ampia condivisione di dati arricchiti per la creazione di nuovi percorsi di ricerca, sfruttando in primo luogo le tecnologie strumentali abilitanti presenti nella

PaaS come Data Fabric della ISPC. Tali servizi di interrogazione e recupero saranno potenziati da modelli di AI linguistico-generativi pre-addestrati, tutt'ora in fase di sperimentazione su porzioni di dati di dominio provenienti dai sistemi informativi degli Istituti centrali.

#### 4. PaaS come Data Fabric

Il complesso funzionale denominato PaaS come Data Fabric rappresenta una vera e propria piattaforma nella piattaforma progettata per fornire agli sviluppatori tecnologie e servizi strumentali utili alla creazione di nuovi servizi basati sui dati (Data Product), sfruttando, da un lato, ciascuna delle tecnologie di Content processing avanzato (concepito nel contesto della Data Fabric ISPC come singoli servizi invocabili ovvero orchestrabili in nuove pipeline di elaborazione complesse) e, dall'altro, le descrizioni standardizzate e i modelli semantici dei beni culturali definiti e usati nella alimentazione dei grafi di conoscenza di dominio e cross-dominio presenti nello spazio dati della ISPC.

All'interno della visione della ISPC, tali servizi sono stati pensati per essere raggruppati nelle seguenti macro-categorie:

1. *creazione di Data Product e virtualizzazione*: insieme di funzionalità che consentiranno di definire nuovi Data Product attraverso criteri di selezione operanti sui dati disponibili - sia interni allo spazio dati della ISPC sia acquisibili dall'esterno - consentendo di connettere, combinare e orchestrare dati eterogenei provenienti da molteplici fonti in modo da ottenerne una vista unificata;
2. *elaborazione e visualizzazione di dati orientata alla Business Intelligence (BI) e alla reportistica*: insieme di servizi utili nella definizione di prospetti e report sui dati. I dataset così ottenuti possono essere resi disponibili sia come Data product accessibili e scaricabili via API che attraverso widget grafici integrabili in portali di terze parti;
3. *intelligent workflow*: attraverso queste funzionalità la piattaforma fornirà la possibilità di definire pipeline di trasformazione sui dati personalizzate, sfruttando i servizi di *Content processing avanzato* presenti nel catalogo dei servizi "standard" ISPC;
4. *training di modelli di Machine Learning/Deep Learning*: la piattaforma fornirà strumenti per il training di nuovi modelli di Machine Learning per poter riconoscere ed estrarre in modo automatico entità di interesse dai media digitali e dalle descrizioni dei beni e delle entità collegate, al fine di poter specializzare i grafi di conoscenza di dominio e cross-dominio standard presenti nella ISPC.

### Le logiche di processo della Data Fabric ISPC

Attraverso la combinazione delle quattro macro-classi di servizi descritte, l'ambiente-laboratorio della PaaS come Data Fabric abilita i macro-processi di business sintetizzati nella seguente infografica (Fig. 3) e di seguito elencati:

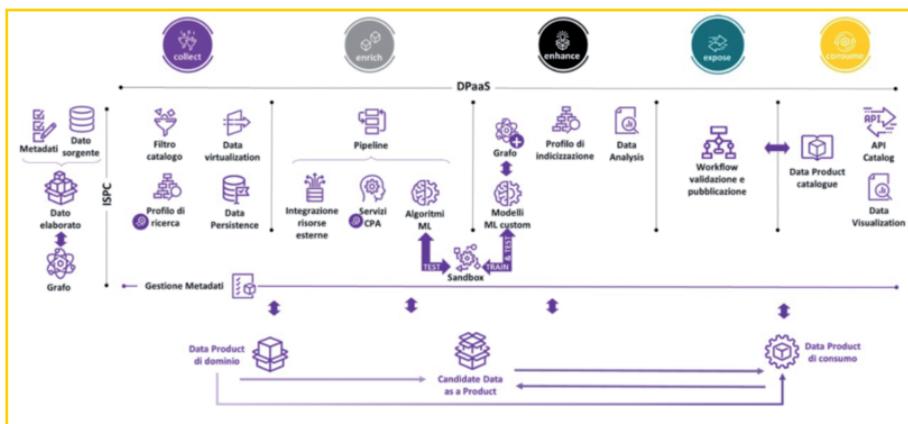


Figura 3. Capabilities del modello di progetto della PaaS ISPC come Data Fabric

- Collect – rappresenta il processo che consente di accedere al Data Lake della ISPC per creare collezioni personalizzate di risorse, configurando profili di ricerca personalizzati, basati su grafi di dominio e cross-dominio;
- Enrich – rappresenta il processo che consente di arricchire oppure estendere i Data Product ottenuti con pipeline di elaborazione, integrando in tal modo modelli di Machine Learning (ML), dataset esterni e servizi di processamento avanzato (cfr. Content processing avanzato) della ISPC;
- Enhance – rappresenta il processo che consente di arricchire/estendere/specializzare il Data product ottenuto (i.e. un grafo di conoscenza riferito a un dominio di interesse), addestrandolo i modelli di ML con le classi e le tassonomie specifiche; per poi esplorarlo configurando e dotarlo di profili di ricerca specializzati;
- Expose - è il processo cross-piattaforma che consente, una volta validati i Data product ottenuti attraverso una o più delle fasi precedenti, di esporre i servizi anche via API;
- Consume – è il processo che abilita l'accesso ai Data Product come API, dashboard/widget, report ad hoc, eccetera.

### L'ISPC e l'Ecosistema digitale del patrimonio culturale

L'ISPC, quale infrastruttura di servizi di arricchimento funzionale e tecnologico destinati ai sistemi di produzione dei dati e di accesso (in generale, sistemi informativi di terze parti), si colloca in una posizione centrale all'interno dell'Ecosistema digitale del patrimonio culturale, con un ruolo abilitante di servizi e processi per un numero quanto più possibile ampio di utenti/sistemi. I dati da essa esposti, dopo essere stati rielaborati e arricchiti in termini di semantica e tecnologie, vengono utilizzati dalle applicazioni pubbliche di front-end attraverso specifiche API (prodotte anche grazie ai servizi di sviluppo rappresentati dalle funzionalità di Data Fabric integrate nella ISPC) e resi fruibili all'intera collettività.

In particolare, l'ISPC si relaziona con le piattaforme previste dall'investimento PNRR M1C3 1.1 (Stream 1.1.10 – Piattaforma di accesso integrata della Digital Library, Stream 1.1.11 – Piattaforma di co-creazione e crowdsourcing e Stream 1.1.12 – Piattaforma dei servizi digitali per sviluppatori e imprese culturali), ma soprattutto con i sistemi informativi dei quattro Istituti centrali del MiC (ICAR, ICBSA, ICCD e ICCU), oltre a enti di vario tipo (pubblici territoriali e/o ecclesiastici, istituti culturali e di ricerca, università, piccole e medie imprese) nonché ad ulteriori soggetti attivi nelle diverse aree di dominio (Fig. 4).

Le interazioni tra ISPC e gli altri soggetti possono essere diverse a seconda delle modalità di adesione, del livello di cooperazione applicativa, del dominio della conoscenza e della sfera di appartenenza MiC o non MiC. I principali interlocutori restano comunque gli Istituti centrali con i loro sistemi di produzione<sup>7</sup> e di accesso<sup>8</sup> al dato, per cui le descrizioni del patrimonio culturale istituzionalmente ad essi riconducibili restano il punto di riferimento più valido e autorevole all'interno dell'ecosistema.

Complessivamente ogni attore opera nell'ecosistema tramite propri sistemi di produzione del dato e/o di accesso, fornendo numerosi servizi alla sua utenza composta sia da professionisti e operatori del settore che da utenti generalisti, studen-

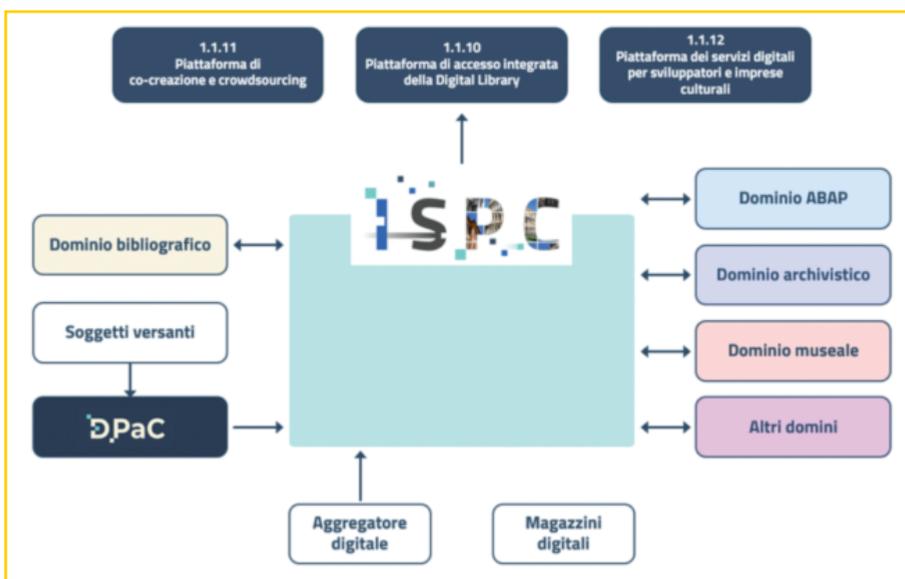


Figura 4. *Ecosistema digitale del patrimonio culturale*

<sup>7</sup> Sistemi di back office che fungono da ambienti di produzione dei dati propri ad un dominio della cultura e che erogano servizi a operatori di settore.

<sup>8</sup> Portale o in generale app che consuma dati provenienti dalla ISPC, che abbiano o meno utilizzato le funzionalità di PaaS per creare il proprio middleware di post-processamento di tali dati in ISPC.

ti, associazioni, turisti e imprese culturali/creative operanti nella filiera produttiva. ISPC, con i propri servizi abilitanti, potenzia e arricchisce i servizi e processi di base offerti dai sistemi di accesso e dai sistemi di produzione del dato aderenti in modalità integrata o federata.

### Un cenno all'architettura dell'informazione dell'Ecosistema

Nell'ecosistema l'architettura delle informazioni è rappresentata da tre aree distinte ma connesse tra loro: ambienti di produzione del dato, ambiente ISPC di manipolazione e arricchimento, ambienti di accesso e fruizione.

a) La prima area, esterna a ISPC, rappresenta gli ambienti di origine dei dati, ovvero i sistemi di produzione degli Istituti centrali, degli enti pubblici territoriali ed ecclesiastici, istituti culturali e di ricerca, università, piccole e medie imprese, che operano nei diversi ambiti della conoscenza. Questi ambienti aderendo alla rete dell'ecosistema inviano i propri contenuti di dominio alla ISPC.

b) L'Infrastruttura è l'ambiente in cui i dati, una volta ricevuti, vengono controllati, processati e arricchiti. In una fase iniziale, a partire da tali dati, si generano grafi di conoscenza locale per ciascuna fonte, che rispecchiano i reticoli propri dei sistemi di produzione. Con i dati acquisiti in ogni dominio per ciascun sistema (MiC e/o non MiC), si ottengono altrettante rappresentazioni a grafo, composte di nodi e relazioni proprie di ogni sistema di produzione, modellato su un tracciato descrittivo base<sup>9</sup>.

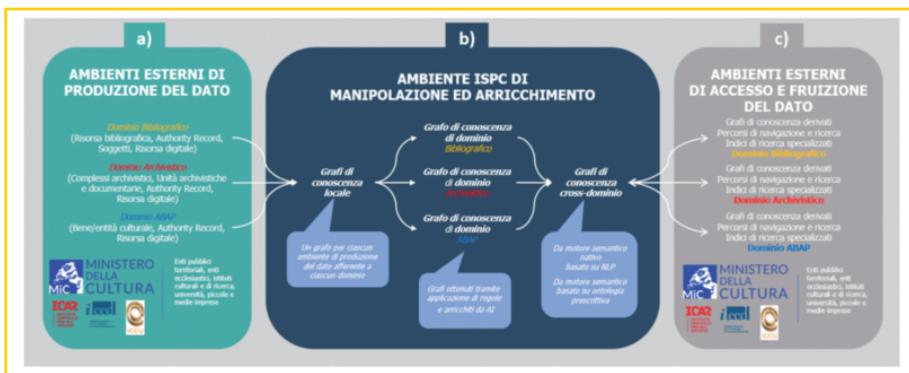


Figura 5. Architettura dell'Informazione (schema semplificato)

Da tali grafi locali, applicando operazioni di deduplica, clusterizzazione, merge e mash-up dei dati provenienti dalle stesse fonti o da fonti diverse, ISPC ottiene i grafi di conoscenza di dominio. All'interno di questi grafi vi confluiscono tutti i dati MiC e non MiC di uno specifico dominio, secondo un tracciato descrittivo target<sup>10</sup>.

<sup>9</sup> Rappresentazione dello schema che descrive per ciascun dominio e per ciascuna fonte (MiC e non MiC) le entità e le relative relazioni presenti nei sistemi di origine.

<sup>10</sup> Rappresentazione dello schema unico per ciascun dominio che descrive le entità e le relative relazioni a valle di un processo di normalizzazione, ottimizzazione e deduplica.

Nella fase successiva a partire dal grafo di dominio, il reticolo è incrementato di nuove entità e/o relazioni, derivanti dall'analisi dei contenuti digitali e/o dei record per mezzo di servizi di *Content processing avanzato* (quali, ad esempio, l'estrazione automatica di entità nominali dai testi (NER), la soggettazione automatizzata e la Image Detection), ottenendo quindi un grafo di conoscenza di dominio arricchito tramite AI<sup>11</sup>.

Infine, questi ultimi alimentano i grafi di conoscenza cross-dominio, le cui entità e relazioni sono trasversali e comprendono tutte le informazioni provenienti dai diversi ambiti.

c) La terza area dell'architettura dell'informazione, esterna come la prima, è costituita dagli ambienti di accesso e fruizione del dato degli Istituti centrali e delle altre tipologie possibili di enti. In questi ambienti esterni si può, quindi, accedere a un'informazione ottimizzata, arricchita, trasversale sia a più fonti sia a più domini della conoscenza, resa tale grazie ai servizi abilitanti e ai macro-processi di business presenti sia nell'ambiente dei servizi standard<sup>12</sup> della ISPC sia nell'ambiente di sviluppo – sua naturale estensione – rappresentato dalla piattaforma PaaS come Data Fabric, che consentirà, infatti, in primo luogo ai sistemi informativi centrali di dominio (come "Alphabetica" e il "Catalogo generale dei beni culturali") di recuperare i dati arricchiti dal grafo di conoscenza di dominio (e crossdominio, in ragione dei percorsi di ricerca di interesse), per poi modellare – grazie alle funzioni specializzate di sviluppo disponibili nella Data Fabric ISPC – nuovi percorsi di ricerca basati su Data Product derivati, specializzati e orientati al particolare modello di classificazione, presentazione e navigazione dei contenuti progettato per lo specifico portale o app mobile.

<sup>11</sup> Ottenuto a partire dal grafo di conoscenza locale che viene arricchito tramite regole e/o tramite AI elaborate in ISPC.

<sup>12</sup> L'insieme dei servizi collegati alle prime tre macro-aree: 1. Gestione e processamento degli asset digitali; 2. Grafi di conoscenza di dominio; 3. Grafi di conoscenza cross-dominio.

*The Software Infrastructure for Cultural Heritage (ISPC) represents the first national cultural data space capable of securely hosting the entire digital heritage of the country, allowing Institutions flexibility in the membership model and complete autonomy in the choice of sharing.*

*In the ISPC, heritage is enhanced and enriched through innovative algorithms and artificial intelligence techniques, and it is correlated with all the available digital cultural resources, both within each domain of belonging (historical, artistic, archaeological, demo-ethno-anthropological, archival, library) and across different domains.*

*The ISPC is an infrastructure with a dual nature: a system of application cooperation services designed to enhance functionally and technologically the information systems of the Ministry of Culture, territorial entities, and cultural institutions in general, from a "Business-to-Institution" perspective; a development laboratory for the creation of new data-based services, designed for developers, cultural enterprises, and start-ups, from a "Business-to-Business" perspective.*

*With a scalable and secure architecture based on Cloud solutions that offer adequate performance and availability levels, ISPC aims to credibly support the role of enabler in providing digital services for the Italian cultural landscape, also in the European dimension.*

L'ultima consultazione dei siti web è avvenuta nel mese di giugno 2023