

EcoDigit

Ecosistema Digitale per la fruizione e la valorizzazione dei beni e delle attività culturali del Lazio

Luigi Asprino
STLab, ISTC-CNR
luigi.asprino@istc.cnr.it

Antonio Budano
INFN Sezione di Roma Tre
antonio.budano@infn.it

Marco Canciani
RilTec, Università di RomaTre
marco.canciani@uniroma3.it

Luisa Carbone
Università della Tuscia
luisa.carbone@unitus.it

Miguel Ceriani
DIAG, Sapienza Università di Roma
ceriani@diag.uniroma1.it

Ludovica Marinucci
STLab, ISTC-CNR
ludovica.marinucci@istc.cnr.it

Massimo Mecella
DIAG, Sapienza Università di Roma
mecella@diag.uniroma1.it

Federico Meschini
Università della Tuscia
fmeschini@unitus.it

Marialuisa Mongelli
ENEA
marialuisa.mongelli@enea.it

Andrea Giovanni Nuzzolese
STLab, ISTC-CNR
andreagiovanni.nuzzolese@cnr.it

Valentina Presutti
LILEC, Università di Bologna STLab,
ISTC-CNR
valentina.presutti@cnr.it

Marco Puccini
ENEA
marco.puccini@enea.it

Mauro Saccone
Università di RomaTre
mauro.saccone@uniroma3.it

Abstract

English. EcoDigit is one of the projects of the “Centro di Eccellenza DTC Lazio”, which intends to aggregate and integrate expertises in the field of technologies applied to cultural heritage. EcoDigit aims at enriching the Anagrafe delle Competenze, another project of the DTC Lazio, with a middleware platform that is able to facilitate the integration of new data sources and to allow the publication and reuse of services for the enhancement and fruition of the cultural heritage in the Lazio region. The project is designed to integrate the collected data through open formats and semantic technologies based on Linked Data model.

Italiano. EcoDigit è uno dei progetti del “Centro di Eccellenza DTC Lazio”, che ha l’obiettivo di aggregare e integrare varie competenze nell’ambito del patrimonio culturale digitale. EcoDigit intende arricchire il sistema dell’Anagrafe delle Competenze, altro progetto del DTC Lazio, con una piattaforma middleware che faciliti l’integrazione di nuove risorse di dati e permetta la pubblicazione e riuso di servizi per la valorizzazione e fruizione del patrimonio culturale nella regione Lazio. Il progetto è finalizzato all’integrazione dei dati censiti attraverso formati aperti e tecnologie semantiche basate sul modello Linked data.

1 Introduzione

Il progetto *Ecosistema Digitale per la fruizione e la valorizzazione dei beni e delle attività culturali del Lazio* (EcoDigit)¹ è una delle iniziative del *Centro di Eccellenza del Distretto Tecnologico per i beni e le attività Culturali* (DTC)², costituito dalle cinque università statali del Lazio - Sapienza, Tor Vergata, Roma Tre, Cassino e Tuscia - in rete con CNR, ENEA e INFN per aggregare e integrare competenze nel settore delle tecnologie per i beni e le attività culturali. Il progetto (luglio 2018 - gennaio 2020) si rivolge alle organizzazioni operanti nel settore della cultura le quali producono e mantengono basi di dati anche di grandi dimensioni archiviati usando formati, modelli e processi diversi. Questo fenomeno comporta la necessità di identificare processi, tecnologie e modelli di integrazione semplici e sostenibili che consentano la fruizione del patrimonio in modo globale e collegato.

Le tecnologie semantiche, e in particolare i Linked Open Data (LOD), sono state ampiamente sfruttate con successo nel campo del patrimonio culturale al fine di migliorare l'accesso e l'esperienza di fruizione dei beni culturali da parte dei cittadini, così come di facilitare la reperibilità, l'integrazione e l'arricchimento dei dati (Dijkshoorn et al., 2018; Daquino et al., 2017; Lodi et al., 2017). Infatti, il paradigma dei LOD è utilizzato per collegare dati provenienti da diverse istituzioni culturali, aumentando così la possibilità di raggiungere i dati culturali disponibili nel Web of Data. L'interconnessione dei contenuti delle organizzazioni collaboratrici ha anche contribuito ad arricchire le informazioni in modo efficace e finalizzato alla valorizzazione del patrimonio culturale (Hyvönen, 2009). La collaborazione tra organizzazioni culturali ha portato anche allo sviluppo collaborativo di ontologie che descrivono il patrimonio culturale a livello internazionale, ad esempio CIDOC-Conceptual Reference Model (CRM) (Doerr, 2003)³, in modo tale che i requisiti di interoperabilità semantica potessero essere soddisfatti all'interno dei loro sistemi. Inoltre, l'uso di ontologie comuni ha facilitato lo scambio di dati e la creazione di enormi biblioteche digitali, ad esempio l'Europeana Data Model (EDM) (Isaac and Haslhofer, 2013)⁴.

EcoDigit, inoltre, non solo ha preso ispirazione da esperienze già consolidate sia degli enti partner e non del progetto sia maturate nell'ambito delle Pubbliche Amministrazioni (PA), ma intende anche instaurare collaborazioni con tali realtà. Ad esempio, il progetto ReCAP⁵ dell'Università Sapienza ha creato una rete di condivisione di conoscenze, strumenti e sperimentazioni che permette di elaborare linee guida e modelli per la costruzione di processi conservativi del patrimonio digitale. Inoltre, il *Sacher Project*⁶ sta affrontando problematiche simili, ma non a livello di progettazione di una piattaforma regionale. Molte istituzioni regionali e nazionali che gestiscono il nostro patrimonio culturale stanno adottando il modello dei dati aperti seguendo le linee guida dell'*Agenzia per l'Italia Digitale* (AgID); tuttavia per le organizzazioni pubbliche e private non è semplice adeguarsi per la mancanza di una piattaforma che ne faciliti il compito. La missione del progetto *Data & Analytics Framework* (DAF)⁷, di cui il network di ontologie OntoPia⁸ è uno dei risultati, è analoga a quella di EcoDigit, ma insiste sull'intero territorio nazionale con focus sull'interoperabilità tra dati tra PA. In tale contesto, EcoDigit può contribuire a rendere efficace e coordinata l'integrazione con i sistemi nazionali.

2 Obiettivi

EcoDigit ha l'obiettivo di arricchire il sistema *Anagrafe delle Competenze*⁹ con una piattaforma middleware che faciliti l'integrazione di nuove sorgenti di dati e consenta la pubblicazione e il riuso di servizi per la valorizzazione e la fruizione del patrimonio culturale del Lazio. Nello specifico, EcoDigit fornisce (i) l'architettura di riferimento per l'integrazione di servizi modulari e per la loro pubblicazione e il riuso; (ii) una componente software sviluppata in forma prototipale per l'orchestrazione dei servizi, l'integrazione

¹<http://ecodigit.dtclazio.it/>

²<https://dtclazio.it/>

³<http://cidoc-crm.org/>

⁴<https://pro.europeana.eu/page/edm-documentation>

⁵<http://digilab.uniroma1.it/attivita/recap>

⁶<http://www.sacherproject.com/progetto>

⁷<https://teamdigitale.governo.it/it/projects/daf.htm>

⁸<https://github.com/italia/daf-ontologie-vocabolari-controllati>

⁹<http://dtc.si.cnr.it/anagrafe-delle-competenze>

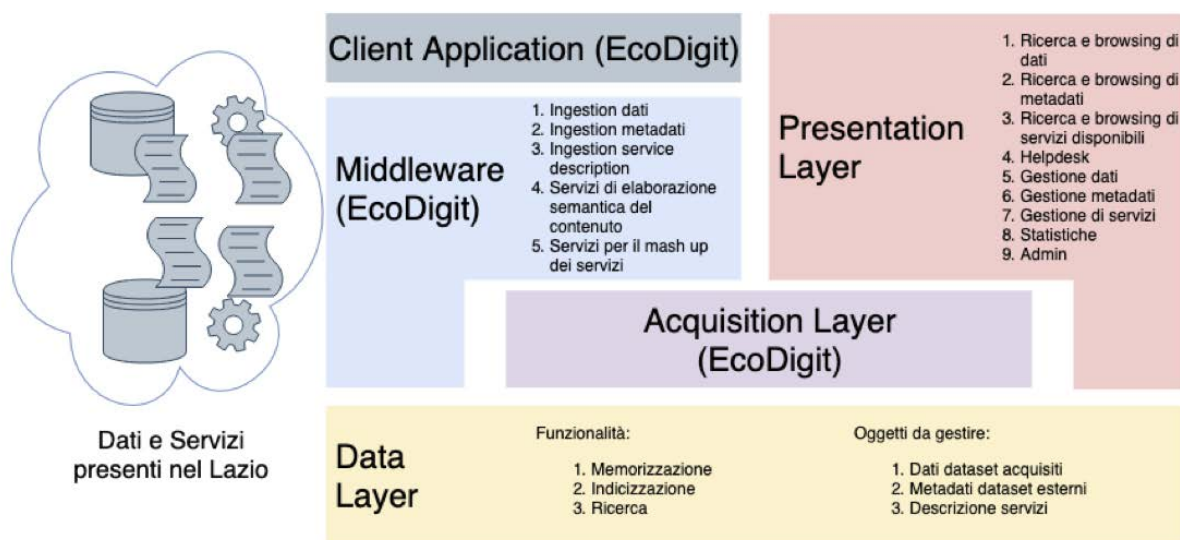


Figura 1: Schema architetturale in cui si inserisce il Middleware di EcoDigit.

e l'aggregazione delle interfacce e dei dati; (iii) la versione prototipale di servizi orientati alla fruizione e valorizzazione del patrimonio culturale.

Con queste caratteristiche EcoDigit si configura come un progetto di ricerca e trasferimento tecnologico collegato in maniera significativa agli altri progetti del DTC Lazio. Con Anagrafe è in relazione fornendo un livello software intermedio, detto middleware, che consente al sistema di aggregare nuove sorgenti di dati, servizi, strumenti innovativi sia industriali che accademici. In questa prospettiva EcoDigit estende il sistema Anagrafe e gli conferisce la capacità di evolvere ed essere esteso. Per gli altri progetti di ricerca, EcoDigit svolge il ruolo di mediatore nei confronti del sistema Anagrafe. Ciò significa che i risultati dei vari progetti potranno essere integrati grazie all'interfacciamento con il middleware di EcoDigit.

In Figura 1 è rappresentato uno schema architetturale di alto livello in cui si inserisce il middleware di competenza di EcoDigit. Sulla sinistra sono rappresentate le risorse, ovvero le banche dati (ciascuna col il proprio formato sintattico e modello concettuale di rappresentazione dei dati) e i servizi messi a disposizione dai vari enti operanti sul territorio che, però, mancano di un punto di accesso unico. Ciò ostacola potenziali utenti interessati a consultare le risorse o creare applicativi basati su esse.

Il sistema, a cui stanno lavorando congiuntamente i gruppi di EcoDigit e Anagrafe, ha l'obiettivo di superare questa frammentarietà. Esso è composto da vari livelli: subito al di sopra del *Data Layer*, c'è (i) l'*Acquisition Layer*, a cui lavorano congiuntamente Anagrafe ed EcoDigit, che si occupa di creare dei flussi di dati i quali saranno acquisiti, uniformati secondo uno schema comune e memorizzati dal sistema. Questi flussi dati verranno creati dal (ii) *Middleware*, che offre alle applicazioni client delle funzionalità per l'elaborazione semantica dei contenuti, facilities per il mashup dei servizi indicizzati dal sistema, oltre a definire le linee guida che servizi e dataset esterni dovranno seguire per garantire l'interoperabilità con il sistema. Nella parte più alta, (iii) il *Client Application* rappresenta una qualunque applicazione che intende utilizzare i dati acquisiti e offerti dal sistema. Essa potrà, attraverso il Middleware, interrogare il sistema per raccogliere dati, recuperare i metadati di dataset esterni o descrizioni dei servizi disponibili.

3 Metodologia

Gli obiettivi sovraesposti sono stati realizzati attraverso l'esecuzione di attività svolte in maniera collaborativa e sinergica tra i partner del progetto e sintetizzabili in tre fasi: 1. Censimento delle risorse relative ai beni culturali presenti nella regione Lazio (cf. Sezione 3.1); 2. Elaborazione del modello di integrazione delle sorgenti nel Middleware (cf. Sezione 3.2); 3. Realizzazione di prototipi volti a verificare la validità dei risultati e dell'approccio seguito (cf. Sezione 3.3).

Tali attività sono state svolte considerando una prospettiva almeno quinquennale di sostenibilità.

3.1 Attività di censimento delle risorse presenti nel Lazio

Fondamentali nel corso del progetto sono state tre tipi di attività preliminari di censimento consistenti in:

- (a) un'analisi dei requisiti del sistema, con relativa ricognizione degli attori e dei casi d'uso, e dello stato dell'arte su tecnologie middleware e architetture per l'integrazione di servizi;
- (b) l'individuazione delle sorgenti dei dati presenti nel Lazio, e dei modelli ontologici e tecniche per la loro integrazione e standardizzazione basata su formati aperti e semantici. In questo contesto, tramite la formalizzazione e disseminazione è stato possibile anche individuare potenziali stakeholder del progetto;
- (c) una valutazione dei tool allo stato dell'arte per la creazione e la gestione di ambienti virtuali 2D/3D.

3.2 Modello di integrazione di una sorgente

Il modello di dati e metadati, che le sorgenti devono rispettare per poter essere acquisite dal middleware EcoDigit, si basa su metodi e tecniche di "metadattazione" e di Semantic Web, comprendendo anche tecnologie proprie delle openAPI e tutto ciò che viene comunemente classificato come Open Data.

Per la sua elaborazione, si è proceduto a un'analisi delle sorgenti censite nel Lazio al fine di evidenziare tanto i domini di conoscenza coperti dalle sorgenti quanto il dettaglio dei vari campi che il modello deve rappresentare. Per ognuno di essi sono stati ricercati gli schemi concettuali considerati standard di riferimento per la modellazione dei dati inerenti a un certo dominio, quali FOAF¹⁰, DOAP¹¹, Org Ontology¹², OntoPia network, SPAR Ontologies¹³, ArCo¹⁴, ecc. Il modello deve essere pensato come l'unione di standard tecnologici, schemi concettuali e di metadattazione allo stato dell'arte per i vari domini di conoscenza, relativi in particolare all'ambito dei beni culturali.

Successivamente, quando i modelli esistenti allo stato dell'arte sono stati ritenuti non in grado di rappresentare semanticamente campi peculiari presenti nei dataset di input, è stata effettuata una modellazione ex novo, utilizzando una metodologia di ingegneria ontologica (Blomqvist et al., 2016), basata su un'estensione di *eXtreme Design* (XD) (Blomqvist et al., 2010). XD è un metodo di progettazione agile di ontologie che si basa sul riuso di *Ontology Design Patterns* (ODP) (Gangemi and Presutti, 2009) al fine di risolvere problemi di modellazione ontologica noti e ricorrenti.

Il workspace del progetto EcoDigit è disponibile su Github¹⁵. Di seguito, si elencano le ontologie definite ex-novo nel corso del progetto:

- *Ontologia delle Organizzazioni*¹⁶ lo scopo di definire un vocabolario condiviso di termini per la descrizione delle organizzazioni che partecipano al Centro di Eccellenza-DTC Lazio. Essa estende: OntoPiA-COV, FOAF, W3C's Organization Ontology.
- *Ontologia delle Valutazioni*¹⁷ ha lo scopo di definire un vocabolario di termini per la descrizione di qualsiasi cosa abbiamo una valutazione associata che è espressa rispetto a una certa scala.
- *Ontologia delle Esperienze e Competenze*¹⁸ ha lo scopo di definire un vocabolario condiviso di termini per la descrizione delle esperienze e competenze di una persona. Estende: OntoPiA-CPV, OntoPiA-COV, FOAF, BIBO; importa l'Ontologia delle Valutazioni.

¹⁰<http://xmlns.com/foaf/spec/>

¹¹<http://usefulinc.com/ns/doap#>

¹²<https://www.w3.org/TR/vocab-org/>

¹³<http://www.sparontologies.net/ontologies>

¹⁴<http://w3id.org/arco>

¹⁵<https://github.com/ecodigit/workspace>

¹⁶<https://w3id.org/ecodigit/ontology/organization>

¹⁷<https://w3id.org/ecodigit/ontology/grade>

¹⁸<https://w3id.org/ecodigit/ontology/eas/>

3.3 Prototipi

A seguito delle attività sovraespresse di censimento dei dataset, degli schemi concettuali e dei tool allo stato dell'arte per la fruizione del patrimonio culturale, sono in fase di elaborazione due prototipi:

- una *Proof-of-Concept*, volta a mostrare la validità del modello attraverso l'integrazione di alcune delle sorgenti identificate nel task di censimento. L'obiettivo principale è quello di creare una best practice che mostri sia la semplicità dell'approccio di integrazione, e la sua scalabilità nel tempo, sia gli strumenti hardware e software che una sorgente deve adottare per aderire al modello di ingresso di EcoDigit;
- un prototipo di un servizio avanzato per la fruizione dei beni culturali nel dominio della formazione. Si sta lavorando ad una soluzione che permetta di raggiungere e fruire da una singola interfaccia i dati acquisiti dal sistema. I partner del progetto, sulla base delle loro expertise, hanno collaborato alla definizione di una tassonomia di categorie che popolino l'interfaccia utente a fini didattici. Anche questo prototipo fa uso di tecnologie semantiche per migliorare la ricerca delle risorse e collegarle ai contenuti di carte tematiche (GIS based) e ricostruzioni 3D per permettere la fruizione in ambiente virtuale dei beni mostrati dall'educatore. In particolare, sarà possibile gestire completamente modelli 3D attraverso l'implementazione ad hoc (cfr. il Workspace GitHub di EcoDigit¹⁹) del software 3DHOP²⁰ e di mostrare dati cartografici tematizzati grazie all'uso della libreria Openlayer²¹.

I due prototipi utilizzano e rappresentano due viste diverse sugli stessi contenuti ai quali applicare le tecniche semantiche di integrazione dei dati, indicandone la provenienza: (i) il dataset della S&TDL-Science & Technology Digital Library del CNR; (ii) il dataset della Sapienza Digital Library; (iii) il modello 3D di Porta Latina (Mura Aureliane), implementato da Roma Tre; (iv) la mappatura GIS con modelli 2D e 3D di alcune chiese della città di Viterbo, curata dalla Tuscia; (v) la ricostruzione 3D del Trono Corsini e del busto in Terracotta di Alessandro VII Chigi, fornita da ENEA; (vi) i Linked Open Data del progetto Arco-Architettura della Conoscenza²², progettati dal CNR-ISTC in collaborazione con il MiBAC-ICCD.

4 Avanzamento tecnologico e impatto del progetto

Il modello di integrazione delle sorgenti nel quale riveste un ruolo fondamentale lo sviluppo delle tecniche di interoperabilità semantica dei vari dataset censiti nel Lazio consente l'arricchimento delle informazioni dei dati esistenti secondo una fruizione integrata. I risultati di questo studio possono essere applicati a qualsiasi progetto che potenzialmente abbia necessità di integrare sorgenti eterogenee. Data la rilevante diversità dei tipi di dati rinvenuti finora, la riusabilità delle tecniche ideate e sperimentate in EcoDigit rappresenta uno degli aspetti principali di avanzamento tecnologico prodotto. Ciò comporta la possibile partecipazione di stakeholder eterogenei alla fruizione ed evoluzione di servizi e contenuti, dando supporto all'inclusione di patrimoni già esistenti e alla loro filiera di gestione.

L'utilizzo di formati e strumenti aperti garantisce tanto la sostenibilità nel tempo quanto la diffusione massiva dei contenuti. Inoltre, il progetto stimola la creazione di nuove figure professionali e la conseguente richiesta formativa sull'uso e sulla divulgazione delle tecnologie semantiche basate sul modello dei Linked Data per la fruizione e la valorizzazione del patrimonio culturale.

Ringraziamenti

Si ringrazia i seguenti enti per il loro supporto istituzionale: la Galleria Corsini, l'Associazione CIVITA, l'Istituto per il Catalogo e la Documentazione (ICCD) del MiBAC, la Sovrintendenza xx.

¹⁹<https://github.com/ecodigit/3dhop-react>

²⁰<http://vcg.isti.cnr.it/3dhop/>

²¹<https://openlayers.org/en/latest/doc/>

²²<http://dati.beniculturali.it/arco/>; <http://dati.beniculturali.it/>

Bibliografia

- Eva Blomqvist, Karl Hammar, and Valentina Presutti. 2016. Engineering ontologies with patterns - the extreme design methodology. In Pascal Hitzler, Aldo Gangemi, Krzysztof Janowicz, Adila Krisnadhi, and Valentina Presutti, editors, *Ontology Engineering with Ontology Design Patterns*, IOS Press, volume 25 of *Studies on the Semantic Web*.
- Eva Blomqvist, Valentina Presutti, Enrico Daga, and Aldo Gangemi. 2010. Experimenting with extreme design. In *Proc. of EKAW 2010*. Springer, volume 6317, pages 120–134.
- Marilena Daquino, Francesca Mambelli, Silvio Peroni, Francesca Tomasi, and Fabio Vitali. 2017. Enhancing semantic expressivity in the cultural heritage domain: Exposing the zeri photo archive as linked open data. *JOCCH* 10(4):21:1–21:21.
- Chris Dijkshoorn, Lora Aroyo, Jacco van Ossenbruggen, and Guus Schreiber. 2018. Modeling cultural heritage data for online publication. *Applied Ontology* 13(4):255–271.
- Martin Doerr. 2003. The CIDOC conceptual reference module: An ontological approach to semantic interoperability of metadata. *AI Magazine* 24(3):75–92.
- Aldo Gangemi and Valentina Presutti. 2009. Ontology design patterns. In (Staab and Studer, 2009), pages 221–243.
- Eero Hyvönen. 2009. Semantic portals for cultural heritage. In (Staab and Studer, 2009), pages 757–778.
- Antoine Isaac and Bernhard Haslhofer. 2013. Europeana linked open data - data.europeana.eu. *Semantic Web* 4(3):291–297.
- Giorgia Lodi, Luigi Asprino, Andrea Giovanni Nuzzolese, Valentina Presutti, Aldo Gangemi, Diego Reforgiato Recupero, Chiara Veninata, and Annarita Orsini. 2017. *Semantic Web for Cultural Heritage Valorisation*, Springer, pages 3–37.
- Steffen Staab and Rudi Studer, editors. 2009. *Handbook on Ontologies*, International Handbooks on Information Systems. Springer.