

Il futuro digitale di 4CH: il gemello digitale dei beni culturali

«DigItalia» 2-2024
DOI: 10.36181/digitalia-00105

Franco Niccolucci – Achille Felicetti

Polo Universitario di Prato (PIN)

Il presente articolo descrive il modello di gestione dei dati sviluppato nel progetto europeo 4CH e la creazione della base della conoscenza per il progetto. Questa si basa sul concetto di gemello digitale in un modo innovativo, cioè come complesso dei dati relativi ad un bene, sia grafici o visuali come immagini o modelli 3D sia testuali o documentali, e delle loro interrelazioni. Si tratta di un approccio innovativo che consente di gestirne la documentazione in modo unitario, indipendentemente dal formato dei dati e dalla loro provenienza.

Introduzione

Come descritto in un precedente articolo su questa rivista¹, il progetto europeo 4CH² ha avuto come obiettivo la progettazione del centro di competenza europeo per la conservazione del patrimonio culturale. Conclusosi a dicembre 2023, 4CH ha prodotto una serie di documenti in cui sono stati analizzati tutti gli aspetti necessari alla creazione di tale Centro, da quelli istituzionali e organizzativi a quelli economici e finanziari e a quelli tecnologici. Confermando una visione “confederale” del Centro di Competenza, cioè basata su una molteplicità di Centri nazionali o regionali relativamente autonomi, coordinati da un Centro europeo incaricato inoltre di fornire servizi comuni, sono in corso di attuazione iniziative per la creazione di soggetti giuridici autonomi in grado di assumere direttamente questo compito a livello europeo e italiano. A livello europeo si prevede la creazione di una entità giuridica di diritto belga senza fine di lucro e dotata della necessaria autonomia giuridica e operativa. La legislazione belga è infatti una delle poche in Europa a consentire la creazione di un ente di questo tipo³, con sede legale in Belgio ma operativo a livello internazionale e per questo ampiamente utilizzato da molti consorzi europei per la continuazione delle attività di progetto. Per l’Italia si sta invece considerando una tipologia flessibile, prevista dal nostro ordinamento, a cui potranno aderire sia i partner italiani

¹ Franco Niccolucci, *4CH: un progetto per sviluppare le applicazioni e le competenze digitali per la gestione del patrimonio culturale*, «DigItalia. Rivista del digitale nei beni culturali», 17 (2022), n. 2, p. 161-167. <<https://doi.org/10.36181/digitalia-00045>>. L’articolo illustra la progettazione e la creazione di un centro di competenza per la conservazione dei beni culturali basato sui modelli 3D di siti ed edifici storici, in risposta a un bando della Commissione Europea che collegava tale Centro ai principi esposti nella Raccomandazione 1970/2021 sulla creazione di un “Data Space” per i beni culturali, attualmente in corso di costituzione da parte di un consorzio guidato da European.

² 4CH – *Competence Centre for the Conservation of Cultural Heritage* è stato un progetto finanziato all’interno del programma Horizon2020 della Commissione Europea con il Grant 101004468, operativo nel periodo dal 1/1/2021 al 31/12/2023.

³ Si tratta dell’*Association Internationale Sans But Lucratif* (AISBL).

del progetto che altri enti interessati al raggiungimento degli obiettivi progettuali. Queste attività istituzionali sono in corso di attuazione da parte del coordinatore di 4CH, Francesco Taccetti, per conto dell'Istituto nazionale di fisica nucleare (INFN).

Per quanto riguarda la tecnologia 3D e le sue applicazioni ai beni culturali, coordinata nel progetto 4CH da Roberto di Giulio per conto di INCEPTION, uno spin-off dell'Università di Ferrara, è stato di recente pubblicato il bando per la creazione di un centro europeo di competenza "virtuale" (cioè non basato su un'infrastruttura fisica) per il 3D, che dovrà supportare le istituzioni del patrimonio culturale nell'utilizzo di questa tecnologia, collaborando con il centro di competenza di 4CH. Recentemente, il progetto vincente è risultato *3D-4CH - Online Competence Centre in 3D for Cultural Heritage*, coordinato da INCEPTION srl, uno dei coordinatori di 4CH, che inizierà le attività nel febbraio 2025. L'obiettivo di 3D-4CH è la predisposizione di standard, tecnologie e servizi per le applicazioni ai beni culturali e la formazione degli operatori del settore, in collaborazione con altre iniziative collegate alla digitalizzazione e alla creazione di modelli tridimensionali per i beni culturali.

Fra i servizi centrali che 4CH dovrà assicurare restano quindi da definire gli aspetti che riguardano la gestione dei dati, un altro pilastro del progetto originario, che gli autori del presente contributo hanno coordinato durante la sua realizzazione. All'interno di questo aspetto puramente digitale possiamo distinguere un primo gruppo di contenuti che riguardano una serie di servizi offerti dal centro in modo virtuale agli operatori e alle istituzioni culturali. Si tratta in sostanza di una biblioteca digitale che contiene i riferimenti a prodotti interni o esterni utili a questa comunità, come ad esempio strumenti di formazione prodotti all'interno del progetto o da terzi, manuali, documentazione, e così via, sia in formato testuale che multimediale. Da un punto di vista informatico l'organizzazione di questi dati è relativamente semplice: si tratta di un catalogo, organizzato per parole chiave, collegato a un archivio interno dove sono depositati i relativi contenuti oppure con link ad archivi esterni. Resta quindi da progettare e realizzare la gestione dei dati veri e propri, ovvero quella che è stata denominata la *Base della Conoscenza*, indicata con l'acronimo *KB* corrispondente alle iniziali del termine inglese corrispondente⁴.

1. La Base della Conoscenza e il suo contenuto

Come in molti altri settori della ricerca, anche nel campo della conservazione e restauro dei beni culturali la mole di dati prodotti dalle attività di operatori e ricercatori è in continua crescita. Se l'utilizzo di tecnologie digitali produce dati e risultati facilmente condivisibili, almeno in via teorica, il loro deposito in sistemi isolati l'uno dall'altro, veri e propri silos informativi, ne rende spesso l'accesso problematico e di fatto irrealizzabile. Si tratta invece di un'esigenza fortemente sentita, come dimostrano esperienze analoghe in settori affini. Per esempio, il sistema ARIADNE⁵, anch'esso prodotto di un progetto europeo e attualmente gestito da una AISBL costituita per questo scopo, si è rivelato di grande potenzialità nel campo della ricerca archeologica. Con una disponibilità di circa quattro milioni di contenuti, il catalogo di ARIADNE si è rivelato di

⁴ Una *Base della Conoscenza* è internazionalmente denominata con il termine tecnico *Knowledge Base*. Questo in qualche modo richiama il concetto più familiare di *Data Base*, da cui differisce sostanzialmente in quanto utilizza una tecnologia più avanzata, più flessibile e più aderente ai contenuti che sarà sinteticamente descritta nel seguito dell'articolo. Per questo motivo si utilizzerà l'acronimo *KB* per indicare la Base della Conoscenza.

⁵ Sono stati pubblicati numerosi articoli sul progetto ARIADNE, fra cui: Sara Di Giorgio – Paola Ronzino, *ARIADNE, l'infrastruttura europea per l'archeologia*, «*Digitalia. Rivista del digitale nei beni culturali*», 11 (2017), n. 1/2, p. 124–135; Franco Niccolucci, *ARIADNEplus: l'avventura continua*, «*Digitalia. Rivista del digitale nei beni culturali*», 15 (2020), n. 2, p. 88–95. <<https://doi.org/10.36181/digitalia-00016>>.

estrema utilità come testimoniano⁶ ricercatori di ogni parte d'Europa e numerose nuove istituzioni che chiedono di esservi incluse.

È ragionevole quindi ritenere che un sistema di dati relativi alla conservazione e restauro avrebbe lo stesso successo. Progetti in tal senso sono ad esempio in corso in Belgio e nel Regno Unito, finanziati dai rispettivi governi. La KB di 4CH intende quindi agire in modo proattivo ed essere pronta a collegare e indicizzare questi sviluppi. Allo stesso tempo, la ricerca sull'organizzazione ottimale della KB può contribuire a migliorare l'organizzazione dei dati stessi oltre che a strutturarne in modo efficiente il catalogo. In effetti, il lavoro sulla KB di 4CH ha anche contribuito a sviluppare aspetti teorici che riteniamo molto rilevanti per quanto riguarda i dati stessi.

L'implementazione del sistema di ricerca è stata fornita dal CNAF⁷, il centro di calcolo dell'INFN. Non entreremo qui nella descrizione tecnica di tale sistema, che è riportata in dettaglio nella documentazione⁸ del progetto, per due motivi. Il primo è determinato dall'impostazione del presente contributo, che intende presentare gli aspetti strutturali della KB nell'organizzazione dei contenuti, anziché quelli strettamente tecnici. Il secondo consiste nel fatto che la strumentazione tecnologica su cui un'implementazione è basata è in continuo progresso e quindi nuovi e più efficienti motori di ricerca possono essere utilizzati senza alterare le caratteristiche del sistema, anzi migliorandole, né cambiare i suoi contenuti. Ad esempio, il motore del sistema ARIADNE è stato sostituito in maniera indolore dopo alcuni anni di funzionamento con uno più efficiente senza nessun cambiamento evidente nell'interfaccia utente, se non una più veloce e più efficiente esecuzione. Basterà quindi accennare che il sistema è basato su un database *NoSQL*, cioè non strutturato in tabelle rigide ma invece organizzato secondo relazioni flessibili fra le entità che rappresentano le informazioni.

Il sistema risultante è stato valutato molto positivamente dagli esperti nominati dalla Commissione Europea in sede di revisione del progetto, benché tuttora sviluppato solo a livello di prototipo.

2. La Base della Conoscenza di 4CH e i modelli 3D

Benché lo scopo del progetto riguardasse programmaticamente, come da bando, i monumenti, gli edifici storici e i siti, cioè il patrimonio culturale materiale e inamovibile, ci si è presto resi conto che tale limitazione era immotivata, illogica e disfunzionale. Non esiste infatti un bene culturale che non abbia una componente immateriale e non c'è motivo di fare distinzioni fra una statua collocata in un museo, mobile anche se con difficoltà, e una posta in una piazza o in una fontana, invece inamovibile. Anche privilegiare in modo eccessivo la tecnologia tridimensionale e farne il fulcro della documentazione appariva riduttivo, ed era forse condizionato da una tendenza presente nella ricerca, influenzata dall'indubbia potenzialità dei modelli tridimensionali. Le prime applicazioni di questa tecnica ai beni culturali nascono negli anni Novanta del secolo scorso in campo archeologico, spesso uno dei più sensibili all'innovazione tecnologica digitale, e sono utilizzate per narrare i risultati della ricerca in un modo di forte impatto e di facile comprensione da parte del pubblico, sostituendo ad esempio nei musei i plastici e i diorami già uti-

⁶ La pubblicazione più recente a questo proposito è: *International Data Aggregation for Archaeological Research and Heritage Management: the ARIADNE experience*, ed. by Edeltraud Aspöck, Julian D. Richards, «Internet Archaeology», 64 (2023), <<https://intarch.ac.uk/journal/issue64/index.html>>.

⁷ L'acronimo CNAF sta per Centro Nazionale Analisi Fotogrammi, un nome un po' anacronistico che risale al 1962 ma che oggi designa il più potente centro di calcolo italiano e uno dei maggiori in Europa.

⁸ La documentazione tecnica è descritta in uno dei *deliverable* del progetto: 4CH Project, *Deliverable 3.3 - Final report on services and tools*, 2023, p. 11-21, <<https://doi.org/10.5281/zenodo.11204151>>.

lizzati nelle ricostruzioni del passato. Il campo di applicazione si estende progressivamente agli oggetti e agli edifici, documentandone la forma ora anche con finalità di gestione, di analisi e di conservazione. Ci si rende presto conto però che specialmente per quest'ultima è necessario aggiungere informazioni di altro genere, relative ad esempio ai materiali, al grado di conservazione o di degrado, agli interventi di restauro e così via. Il modello 3D diventa così anche il supporto su cui posizionare queste informazioni aggiuntive, come nel caso del restauro della fontana del Nettuno di Bologna⁹ in cui vengono individuate le aree della statua da restaurare e se ne documentano le condizioni "appendendole" al suo modello 3D.

Questo metodo, di grande efficacia visiva e documentativa, presenta un rischio intrinseco, già evidenziato anni fa nel progetto europeo 3D-COFORM¹⁰: nei necessari processi di trasformazione della nuvola di punti si possono modificare o addirittura eliminare parti di modello (un procedimento chiamato decimazione) per ridurne il "peso" digitale e quindi inavvertitamente modificare o perdere delle informazioni associate alle regioni dell'oggetto digitale modificate o eliminate. Occorre quindi un'estrema cautela nella manipolazione del modello 3D.

Se in questi casi il modello 3D consiste in una nuvola di punti ottenuta mediante tecniche di scansione laser o fotogrammetria, in campo architettonico prende invece piede una metodologia mutuata dalla progettazione. I modelli 3D in questo caso sono ottenuti in CAD¹¹ e le altre informazioni vengono aggiunte utilizzando un sistema derivato dalla progettazione di nuovi edifici e denominato BIM (*Building Information Modeling*), che prevede la possibilità di unire alle informazioni grafiche di progetto quelle relative ai materiali da utilizzare, ai servizi presenti nelle pareti, alla struttura dell'edificio e altre ancora. Il BIM è codificato in uno standard ISO e possiede software specializzato per il suo utilizzo. Esso nasce quindi per documentare edifici nuovi, e richiede un adattamento e informazioni aggiuntive quando invece s'intende documentare quelli storici; non si presta inoltre ad oggetti come quelli presenti in un museo. Occorrono comunque delle aggiunte, che danno luogo a un sistema chiamato HBIM, dove H sta per *Heritage*, patrimonio culturale. Tuttavia, tali estensioni non-standard dello standard BIM risultano difficili da gestire con il software esistente.

Ma al di là di queste problematiche – tutte le soluzioni digitali ne presentano diverse – c'è un aspetto che rende questi approcci inadatti a una base della conoscenza come quella di 4CH: essendo focalizzati su un singolo oggetto, quello rappresentato dal modello 3D, essi guardano all'albero e non vedono la foresta. Fuor di metafora, essi non consentono una visione d'insieme del patrimonio culturale, ad esempio per individuare gli oggetti fatti con lo stesso materiale, oppure quelli di una determinata epoca o soggetti a uno specifico intervento di conservazione. In altri termini, non è possibile (o è troppo complicato) determinare ad esempio tutti i monumenti equestri prodotti in un materiale soggetto a degrado per via delle piogge acide; non si può stabilire quali interventi di restauro conservativo siano stati compiuti con successo su dipinti prodotti usando determinati materiali; e così via. Il modello 3D è poi irrilevante (o forse addirittura inesistente) per documentare beni immateriali come il canto a tenore sardo, la transumanza o la falconeria, tutti inseriti nella lista UNESCO del patrimonio immateriale dell'umanità. Si è reso

⁹ Fabrizio Ivan Apollonio et al., *A 3D-centered information system for the documentation of a complex restoration intervention*, «Journal of Cultural Heritage», 29 (2017), p. 89-99.

¹⁰ 3D-COFORM è un progetto sviluppato negli anni 2008-2012 e dedicato alla creazione di collezioni digitali di modelli 3D di beni culturali. Informazioni sul progetto e sui suoi risultati si possono reperire sul sito della Commissione Europea: <<https://cordis.europa.eu/project/id/231809>>.

¹¹ Anche se questo acronimo è di uso frequente, ricordiamo che il suo significato è *Computer-Aided Design*, cioè la parte grafica della progettazione supportata da computer mediante software specializzato.

quindi necessario sviluppare un approccio diverso alla documentazione del patrimonio culturale, basato su un punto di vista olistico che integri forma (se esistente), sostanza, condizione e interventi deliberati o accidentali su di esso. Il paradigma che integra tutti questi aspetti e potenzialmente altri ancora è il concetto di gemello digitale, anche noto come *digital twin*¹².

3. I gemelli digitali nella documentazione del patrimonio culturale

Il concetto di *digital twin*, cioè di gemello digitale, è ampiamente diffuso in campo industriale. Esso designa un complesso di dati e processi digitali che rappresentano e simulano il comportamento di un apparecchio collocato nel mondo reale. È quindi composto da un insieme di dati che definiscono e descrivono complessivamente il modello digitale, e da connessioni con il mondo reale che permettono di simulare le reazioni dell'oggetto reale rappresentato dal suo gemello digitale rispetto ad eventi che accadono (o possono accadere) nel mondo reale. Tralasciando momentaneamente l'aspetto di simulazione, il concetto di gemello digitale corrisponde esattamente alle nostre esigenze: un oggetto digitale composto da tutte le informazioni pertinenti a un bene culturale, collegate da relazioni che ne indicano le reciproche corrispondenze e interdipendenze. Quindi il gemello digitale di un edificio storico sarà composto dal modello della sua forma, come un modello 3D, ma anche dalla documentazione dei suoi materiali costruttivi, delle relazioni fra i suoi interni, da informazioni storico-architettoniche, dalla storia del suo utilizzo, dalla documentazione degli interventi di restauro e così via, compreso il suo valore civico presente e passato. Non è necessario che tutte queste informazioni siano collocate nello stesso archivio digitale, in quanto il sistema di documentazione può aggregare e catalogare informazioni accessibili in rete anche se poste in archivi informatici diversi. Per esempio, si potrebbe trattare di uno (o più) modelli 3D depositati in sistemi diversi a seguito di attività di digitalizzazione compiute separatamente; di fotografie e filmati; di documentazione ufficiale quale quella contenuta in schede ministeriali; di articoli scientifici di ricerca storico-architettonica pubblicati in riviste; di rapporti di interventi di restauro conservati presso gli istituti che li hanno effettuati; di documentazione disparata sull'uso dell'edificio in passato e nel presente; e di storie relative all'edificio stesso, che contribuiscono a crearne il valore culturale per la sua comunità di riferimento e in generale per tutti. Il complesso di tutte queste informazioni relative a un bene culturale è definito come il *gemello digitale culturale* di tale bene, in inglese *Heritage Digital Twin* (in sigla: HDT). Il gemello digitale culturale di un bene comprende quindi tutte le informazioni disponibili per il bene in questione e le loro interrelazioni. Il sistema che organizza tali dati, che prende com'è noto il nome tecnico di *ontologia*, cioè la «rappresentazione formale, condivisa ed esplicita di una concettualizzazione di un dominio di interesse»¹³, è stato denominato HDTO, dove O sta per ontologia. Quindi HDTO è formata dai concetti generali necessari e pertinenti alla documentazione dei beni culturali (tecnicamente denominati *class*) e dalle loro interrelazioni (chiamate

¹² La bibliografia sui *digital twin* e le loro applicazioni in campo industriale è sterminata. Forse la prima applicazione nel campo del patrimonio culturale si può far risalire al progetto *Gemini* (2018-2022) curato dal *Centre for Digital Built Britain* dell'Università di Cambridge, che ha sviluppato una serie di principi alla base di una progettazione del costruito finalizzata al bene pubblico e alla condivisione. Anche se il progetto si è concluso si possono trovare informazioni sul sito ad esso dedicato: <<https://www.cdcb.cam.ac.uk/what-we-did/national-digital-twin-programme>>. Per quanto invece riguarda le applicazioni di questo concetto ai beni culturali si può fare riferimento agli articoli sul tema citati nella nota 15. Anche se di contenuto tecnico, queste pubblicazioni illustrano la necessità e l'uso del concetto di *digital twin* con una serie di esempi relativi al patrimonio culturale rivolti agli operatori del settore. I contenuti di questi articoli sono comunque riassunti nel seguito del presente contributo.

¹³ La definizione è tratta da Wikipedia, <[https://it.wikipedia.org/wiki/Ontologia_\(informatica\)](https://it.wikipedia.org/wiki/Ontologia_(informatica))>.

proprietà). Per ciascun oggetto culturale, ogni informazione ad esso relativa apparterrà a una classe, di cui le singole informazioni saranno esempi – o, come si dice tecnicamente, *istanze*¹⁴.

Vale la pena di osservare che questo approccio riduce la centralità del modello 3D nell'organizzazione dei dati, rendendolo solo uno degli elementi, anche se fra i più importanti, che descrivono l'oggetto.

Per garantire l'interoperabilità di questa organizzazione dei dati con quelle esistenti, l'ontologia HDTO è compatibile con lo standard per la documentazione dei beni culturali che è, com'è noto, il CIDOC *Conceptual Reference Model* (CRM). HDTO è infatti un'estensione compatibile del CRM, di cui specializza alcuni concetti, utilizzando anche altre estensioni, già esistenti, dello stesso CRM, come quella destinata agli oggetti digitali in generale, denominata CRMdig; quella per le analisi scientifiche, CRMsci; e altre ancora.

Nel progetto 4CH, la creazione dell'ontologia HDTO è stato il primo passo per organizzare i dati e aggregare le informazioni provenienti da fonti diverse. Nella base della conoscenza (KB) di 4CH organizzata secondo tale ontologia non sono contenuti quindi i dati, ma solo i metadati, con il riferimento agli archivi in cui i dati sono depositati. Una ricerca impostata sulla KB da un utente sarà così effettuata sui metadati che permetteranno la selezione dei dati rilevanti e, a richiesta, l'accesso ai dati originari.

Questa ontologia è stata accolta molto favorevolmente dalla comunità scientifica, tanto che il primo articolo in cui veniva descritta¹⁵, anche se in modo da sviluppare ulteriormente, ha raggiunto rapidamente una significativa popolarità ed ha ricevuto un premio come il migliore nella sua categoria. Importanti centri di restauro si sono dichiarati disponibili a fornire documentazione da usare come test per il sistema. Ulteriori sviluppi dell'ontologia sono in corso, ad esempio, per prendere in considerazione la presenza di sensori che rilevino in modo automatico la condizione dei beni documentati nel sistema e l'eventuale presenza di situazioni di rischio per la loro conservazione¹⁶.

L'utilizzo dell'ontologia HDTO rende la nostra Base della Conoscenza (KB) interoperabile con tutti i sistemi basati sul CRM o comunque con esso compatibili, permettendone quindi l'integrazione diretta nella KB attraverso i metadati. Il contenuto della KB è gestito attraverso un database NoSQL, cioè non strutturato in tabelle ma basato invece sulle classi definite nell'ontologia e sulle relazioni fra di esse, sviluppando così un sistema estremamente articolato e ricco di collegamenti interni. Per esempio, edifici storici posti nella stessa città sono collegati dal luogo comune in cui si trovano; ciascuno di essi sarà poi collegato ad altri edifici dello stesso stile architettonico oppure con la stessa funzione (chiesa, edificio civico, e così via). Un altro possibile collegamento riguarda il materiale di costruzione, e collega a ciascuno di questi tutti gli edifici che lo utilizzano. Sono invece specifici di ogni bene le informazioni che lo riguardano direttamente, come il nome e quelle visive: foto, modelli 3D, video e altro ancora. Ricompare qui il modello tridimensionale, an-

¹⁴ Il termine *istanza* con questo significato è assente dai dizionari della lingua italiana. Si tratta infatti di un anglicismo tecnico abbastanza recente, calco della parola *instance* che in inglese significa esempio.

¹⁵ Franco Niccolucci – Achille Felicetti – Sorin Hermon, *Populating the Digital Space for Cultural Heritage with Heritage Digital Twins*, «Data», 7 (2022), <<https://doi.org/10.3390/data7080105>>. Questo articolo è stato accolto con grande interesse dalla comunità scientifica e premiato come il migliore pubblicato sul tema nel 2022. Più di recente, è stato pubblicato: Franco Niccolucci – Achille Felicetti, *Digital Twin Sensors in Cultural Heritage Ontology Applications*, «Sensors», 24 (2024), <<https://doi.org/10.3390/s24123978>>, che illustra l'utilizzo di sensori nel campo dei beni culturali.

¹⁶ Questo tipo di sistemi prendono il nome di *Internet of Cultural Things* (IoCT), applicazione ai beni culturali del concetto di IoT (*Internet of Things*), dove sensori dialogano fra loro attraverso la rete e svolgono in modo automatico operazioni sugli oggetti presenti nel sistema. Esempi tipici sono le applicazioni di domotica e quelle di sicurezza.

che se con una funzione ancillare rispetto agli approcci accennati in precedenza, in cui esso riveste solo il ruolo di gemello digitale, che invece nella nostra impostazione ha un carattere più globale e, come già detto, olistico, comprendendo tutti i dati relativi a uno stesso bene culturale. Per rappresentare il contenuto della KB sono disponibili diversi linguaggi tecnici, che hanno la caratteristica della leggibilità da parte di un utente umano oltre che, ovviamente, da parte di una macchina. La descrizione così codificata è però estremamente verbosa, e quindi si preferisce, specialmente negli esempi, utilizzare una descrizione grafica in cui le classi sono rappresentate da caselle e le relazioni che collegano le classi da frecce che uniscono le caselle corrispondenti. Le frecce recano un'etichetta che specifica il nome della relazione. Si ottiene così un *grafo semantico*, cioè la rappresentazione visiva del sistema di documentazione. Lo standard richiede inoltre che tutti gli elementi del grafo riportino il codice (sigla più nome) delle classi e delle proprietà, convenzionalmente formati con parole inglesi. Nel caso di un esempio, come nella sezione seguente, si indicano inoltre i valori specifici delle istanze nelle rispettive caselle.

4. Un esempio di applicazione: il Battistero di San Giovanni a Firenze

Per illustrare un'applicazione del sistema sopra descritto e le sue potenzialità descriviamo di seguito l'esempio relativo al Battistero di San Giovanni a Firenze. Oltre agli aspetti architettonici e storico-artistici, facciamo notare come sia possibile integrare nella documentazione la bibliogra-

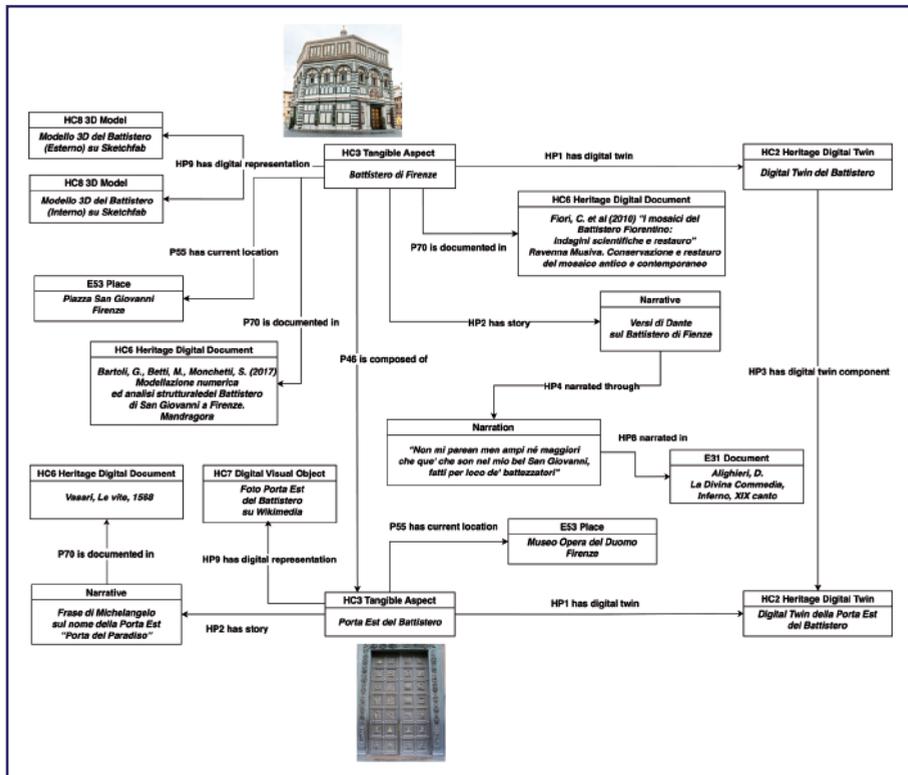


Figura 1. La documentazione del Battistero di San Giovanni a Firenze

fia relativa, gli aspetti letterari (la citazione di Dante nella *Divina Commedia*) e la particolarità di questo edificio di contenere un bene descritto separatamente e dotato di un proprio gemello digitale (non riportato in dettaglio nell'esempio), cioè la Porta est, anche nota come Porta del Paradiso perché così chiamata da Michelangelo, realizzata da Lorenzo Ghiberti. Del gemello digitale fanno parte anche documentazioni visive, fra cui un modello 3D.

Nella descrizione precedente ci siamo limitati a indicare alcuni esempi del contenuto della KB relativi al Battistero di San Giovanni per dare un'indicazione della funzionalità del sistema. Si potrebbero naturalmente aggiungere moltissime altre informazioni, per esempio quelle relative al restauro delle formelle della "Porta del Paradiso" danneggiate dall'alluvione del 1966, oppure quelle sulla conservazione dei mosaici oggetto di un recente intervento di conservazione, e altre ancora. L'efficienza del sistema si basa sulla sua capacità di aggregare informazioni da fonti disparate (purché disponibili online in formato digitale) e di stabilire così collegamenti fra contenuti digitali diversi ma tutti pertinenti allo stesso bene, il Battistero. Questa operazione ha l'ulteriore vantaggio di non duplicare i contenuti, di cui sono solo raccolti i metadati.

5. Prospettive per il futuro

Per arrivare a un sistema effettivamente operativo e utilizzabile saranno necessari numerosi passaggi. Il primo richiede la collaborazione di ricercatori, operatori culturali e istituzioni che possiedono i dati necessari. Se abbiamo infatti già a disposizione dei dati reali di test, va però ulteriormente sviluppata la collaborazione con questi soggetti, in Italia e all'estero, per ampliare la base da aggregare nella KB. Queste collaborazioni permetteranno di raffinare e precisare ulteriormente il nostro approccio semantico a fronte di problematiche reali.

Il secondo riguarda assicurare le risorse organizzative e finanziarie per supportare il lavoro. Anche su questo aspetto il gruppo di ricerca che ha lavorato in 4CH sta attivamente ricercando soluzioni.

Uno dei problemi, anzi forse il problema che si presenta è la presenza e la disponibilità di dati e metadati adeguati. Anche assumendo che la documentazione sia conservata in archivi potenzialmente accessibili per l'aggregazione, ci si può aspettare che spesso i metadati disponibili siano modesti, essendo probabile che tali archivi siano stati concepiti più come depositi statici che come oggetto di ricerche strutturate. La creazione di metadati apparentemente non indispensabili è un'attività noiosa, che richiede tempo e non produce effetti tangibili nell'immediato. Nella fase iniziale ci siamo quindi limitati a compilarli manualmente per testare il funzionamento del sistema. Per una fase successiva di funzionamento a regime sono allo studio sistemi automatici che possono essere attivati contestualmente alla creazione del documento, oppure utilizzati su documenti già esistenti, basati sull'intelligenza artificiale. Un esperimento iniziale¹⁷ di qualche tempo fa su un piccolo insieme di documentazione archeologica ha permesso di verificare la fattibilità di questo procedimento, estraendo automaticamente da rapporti di scavo archeologico le parole rilevanti (ad esempio anfora, ceramica, e così via) e permettendo così l'indicizzazione automatica del testo analizzato.

Questo procedimento di arricchimento automatico dei metadati è destinato a migliorare rapidamente in efficienza e nei tempi di "apprendimento" grazie ai progressi recenti in questo settore.

¹⁷ Achille Felicetti – Daniel Williams – Ilenia Galluccio – David Tudhope – Franco Niccolucci, *NLP Tools for Knowledge Extraction from Italian Archaeological Free Text*, in: *3rd Digital Heritage International Congress (DigitalHERITAGE) & 2018 24th International Conference on Virtual Systems & Multimedia (VSMM 2018)*, San Francisco, CA, USA, 2018, p. 1-8, DOI: <<https://doi.org/10.1109/DigitalHeritage.2018.8810001>>.

Si tratta di evitare le cosiddette “allucinazioni”, cioè risultati clamorosamente errati derivanti dall'utilizzo di informazioni non pertinenti. Ciò è consentito da alcuni modelli di intelligenza artificiale che, a differenza di quelli più noti, non pescano a caso fra una grande quantità di informazioni ma invece utilizzano solo un insieme predefinito di partenza. Si tratta comunque di un'attività di ricerca da impostare per il futuro. Sarà poi necessario sviluppare una descrizione sistematica dei protocolli di analisi scientifica, producendo thesauri e dizionari controllati da affiancare a quelli largamente diffusi per le informazioni storico-artistiche quale, ad esempio, lo Art & Architecture Thesaurus (AAT) sviluppato dal Getty Institute.

The present article describes the data management model developed in the 4CH European project and the creation of the project knowledge base. This is based on the digital twin concept in an innovative way, namely as the collection of data pertaining to an asset, both graphical or visual data, such as images or 3D models, and textual or documental data, as well as their interrelations. It is an innovative approach enabling the management of the documentation in a unitary way, independently of the data format and their provenance.