

Verso uno standard di catalogazione su software e dati aperti: il progetto VRC

Fabrizio Giudici^{*}, Alick Macdonnel McLean[†],
Augusto Palombini[‡],

SOMMARIO. Il progetto Visual Resource Collaborative (VRC), della Syracuse University in Florence (SUF) è un Database relazionale studiato per la catalogazione e la fruizione on-line del patrimonio culturale. La sua realizzazione si basa sull'idea di una condivisione di codici, dati e metadati secondo la quale, a valle della realizzazione del software open source, ciascuna istituzione può fruire liberamente di esso, mettendo a propria volta a disposizione i propri dati e divenendo un nodo del sistema. Le istituzioni e i singoli interessati possono sin d'ora contattare il progetto VRC scrivendo agli autori.

ABSTRACT. *The Syracuse University in Florence “Visual Resource Collaborative Project” (VRC), is a relational database conceived for the classification and on-line diffusion of Cultural Heritage. The project core is an extension of the concept of source code sharing: thanks to the use of Open Source software, each institution can use the VRC to share its data and, thus, becoming a system 'node'. Institutions and people who are interested in the project may contact VRC project, writing to the authors.*

1. Le origini di VRC

Le origini del progetto Visual Resource Collaborative (VRC), della Syracuse University in Florence (SUF), aiutano a comprendere la struttura e le potenzialità di uso allargato di questo strumento. Il Media Lab della Syracuse Uni-

versity in Florence iniziò a sviluppare Visual Resource Collaborative nel 2003, in rapporto alle necessità interne dei propri docenti di passare dagli strumenti analogici per la gestione di diapositive a quelli digitali. Si concluse che la via più conveniente per portare dall'analogico al digitale una parte quanto più ampia possibile del materiale, nel minor tempo, era quella di proporre un "do ut des" che è poi divenuto il nucleo del lavoro e della filosofia della SUF. Ogni docente mise a disposizione una parte della propria collezione di slides elaborate alla Syracuse University, e le ricevette indietro su CDs, in directories distinte e con i dati di catalogazione incorporati nel formato dei files.

Tutte le slides di ciascun professore vennero quindi messe a disposizione, a turno, di colleghi e studenti di tutte le classi. Lo scambio era e resta tuttora semplice: in cambio della possibilità di ricevere, da parte dell'università, il servizio di digitalizzazione e catalogazione delle proprie slides, per uso individuale, ciascuno deve consentire all'intera comunità universitaria l'accesso ai propri dati.

La rapida adozione di tali servizi di catalogazione e digitalizzazione da parte di svariate discipline ha poi costretto a sviluppare strumenti di catalogazione molto flessibili, in grado di adattarsi anche a nuovi campi della conoscenza, con il crescere degli interessi interdisciplinari.

Nessuna soluzione già pronta soddisfaceva tali esigenze, e si è quindi deciso di sviluppare un database in autonomia, optando per il DB relazionale Open Source PostgreSQL, che è stato usato per offrire uno strumento di catalogazione semplice e tuttavia infinitamente estendibile, consistente in entità, date, valori e modificatori di relazione. Tutti i termini sono basati su standard e tutti i dati immessi dispongono di backup bibliografico.

2. VRC e la catalogazione del Patrimonio Culturale

La proposta di VRC, al di là del suo uso contingente e delle indubbie utilità pratiche, può essere vista anche sotto la prospettiva di un valore metodologico che ha una precisa collocazione nel contesto della storia delle discipline cui si riferisce. Vi è un evidente analogia fra l'idea originale di VRC e gli interessi di archeologi e storici: la consapevolezza della pubblica utilità dei dati, da sottolineare, e la trasparenza dei processi. L'idea di una catalogazione flessibile, grazie alla quale è possibile ridisegnare costantemente i rapporti tassonomici fra gli attributi di un oggetto, rappresenta infatti un potente strumento di ricerca.

La storia dei processi di classificazione in archeologia è strettamente legata all'evoluzione stessa della disciplina, delineando i diversi approcci che si sono succeduti allo studio del passato. Già per gli studi antiquari del XVII secolo la classificazione (almeno nelle intenzioni) rappresentava una forma basilare di conoscenza e di attribuzione di valore. La svolta rappresentata dagli studi di Winckelmann si pone proprio su questo piano.¹ Le crono-tipologie della prima metà del XX secolo, prima che il radiocarbonio risolvesse su un altro piano il problema della determinazione cronologica, rappresentano ulteriori sforzi nella

medesima direzione. Ma il denso dibattito sull'effettiva capacità delle classificazioni di contribuire a reali passi in avanti della ricerca è un problema tutt'oggi aperto (per una trattazione dettagliata della questione si veda: Palombini 2001).

In sintesi, il problema nasce da un duplice uso possibile delle classificazioni: come griglie rigide, in cui incasellare oggetti e definirli, cioè attribuire loro una precisa identità, o come strumenti di lavoro flessibili, che tratteggiano gli aspetti salienti di un oggetto solo in funzione di una momentanea operazione euristica, che può essere rielaborata alla luce della creatività del ricercatore nel formulare nuove ipotesi di lavoro. Questo aspetto è già stato affrontato da altre discipline, e cruciale nell'epistemologia moderna², con la distinzione “fra *aspetto ontologico* e *aspetto epistemologico* dell'eidos (l'eidos è un dato o un posto? lo trovo nella cosa o lo applico alla cosa per renderla intelligibile?)” (ivi, 258, corsivo originale), ma tale piano di riflessione non è sempre chiaro agli archeologi, tuttavia è indubbia la possibilità infinitamente maggiore del secondo tipo di approccio dal punto di vista delle potenzialità di ricerca.

Va inoltre osservato che a fianco di alcuni parziali vantaggi (diffusione di modelli standard, possibilità di query meccaniche e facilitate), una tipologia intesa come rigida griglia classificatoria presenta molti inconvenienti: l'adattamento di modelli rigidi a realtà multiformi (si pensi ai dibattiti sulla tipologia litica), la difficoltà di traduzione e *switching* delle definizioni nei diversi sistemi di misura (si pensi ai diversi calendari o alle definizioni territoriali succedutisi nella storia dell'umanità).

Inoltre, e soprattutto, la definizione di relazioni rigide e precostituite fra gli attributi di un elemento ostacola in modo spesso fatale il mutamento dello schema nel tempo, l'autocorrezione dello schema stesso, la messa in discussione delle strategie utilizzate, il riconoscimento di nuovi criteri degni di considerazione, etc.

La difficoltà a riconoscere il valore di una classificazione dinamica, in grado di riformulare continuamente se stessa, di arricchirsi di nuovi attributi e di riformulare le relazioni fra gli attributi stessi, è oggi superata dalle possibilità offerte dai database di ultima generazione, ed in questo contesto VRC rappresenta una novità assoluta in grado di aprire alla ricerca archeologica e storico-artistica delle prospettive assolutamente innovative su diversi piani:

- Sul piano della divulgazione: facilitando la diffusione di informazioni sul patrimonio culturale, e insegnando che la creatività è elemento essenziale nell'attività di ricerca per mettere a fuoco relazioni nuove fra le informazioni.
- Sul piano metodologico della ricerca: come potente strumento dinamico per costruire e riorganizzare archivi senza rischi di obsolescenza, ma anche aprendo la via a un nuovo pensiero classificatorio, e spingendo gli studiosi a strategie di comunicazione e collaborazione.
- Sul piano finanziario: grazie all'approccio Open Source che consente filiere di lavoro fruibili per un uso diffuso anche nei paesi in via di sviluppo, dove

l'uso di software commerciale presenta insormontabili problemi economici. Inoltre, come l'Unesco ha spesso sottolineato³ i file realizzati in formati aperti, nell'ambito dei Beni Culturali, offrono una maggiore garanzia di non divenire obsoleti o andare perduti, in quanto è molto più diffusa, frequente e semplice la loro riproduzione e trasmissione, anche attraverso il web.

3. VRC: gli aspetti tecnici

Il progetto VRC è basato sul concetto di catalogazione di *entità generiche* (oggetti d'arte, reperti, agenti umani, luoghi, materiali, geo-dati, etc...) che possono essere legate da relazioni arbitrarie (ad esempio "L'oggetto X *si trova alle coordinate X,Y*", "L'oggetto Y *è stato scoperto* da Tizio" etc). Questa impostazione prevede la gestione di una serie di campi (le *entità*: nomi geografici, di persone, date, etc., in modo analogo ad un DB tradizionale, ma anche la gestione di una serie di *relazioni*, che possono essere trattate a propria volta come campi. In questo modo è possibile non solo modificare dinamicamente e in qualunque momento l'intera struttura della definizione di un elemento:
es. da:

"l'oggetto x *è stato scoperto* da Paolo ed *è conservato* nel museo di Torino dall'anno 2000"

a:

"l'oggetto x *è stato scoperto* da Mario ed *è conservato* nel museo di Firenze dall'anno 2002")

ma di creare catene di campi, personalizzando costantemente dimensioni e definizioni dei record:

es. da:

"l'oggetto x *è stato scoperto* da Paolo ed *è conservato* nel museo di Torino"

a:

"l'oggetto x *è stato scoperto* da Mario, la scoperta è *stata erroneamente attribuita* a Paolo, *è conservato* nel museo di Firenze dal 2001 ed *è stato conservato* nel museo di Torino dal 2000 al 2001"

Particolare attenzione è infatti rivolta alla rappresentazione di relazioni mutevoli (la storia degli spostamenti di un manufatto o i cambi di nome) e

alla rappresentazione delle date, attraverso un sistema pensato appositamente per le esigenze archeologiche e storico-artistiche in grado di mantenere precisione di giorno e ora, come pure rappresentare datazioni archeologiche e paleontologiche con vari gradi di approssimazione, dal giorno-mese-anno all'era geologica.

Inoltre, la catalogazione è studiata per la coesistenza di criteri multipli prevedendo classificazioni di località secondo criteri amministrativi, geografici o geopolitici e persino storici (le province dell'Impero Romano o di quello Cinese). Ad ogni oggetto sono associabili *documenti* (fotografie, planimetrie, modelli 3D, etc.) per cui definire politiche di accesso in base al copyright. Un motore di ricerca consente selezioni complesse e facenti uso di relazioni.

Oltre a realizzare nodi di database, fruibili sia attraverso un'interfaccia web che applicazioni specializzate, VRC consente di costruire reti "peer to peer" che scambiano informazioni grazie a tecnologie di Web Semantico. L'uso della lingua franca RDF rende anche possibile interfacciamenti con altri archivi.

VRC è un progetto modulare, composto da parti che possono anche integrarsi a componenti diversi:

- (1) il nucleo di gestione della base di dati;
- (2) un'interfaccia Web;
- (3) moduli per l'integrazione con applicazioni indipendenti "rich client";
- (4) un modulo di elaborazione fotografica.

L'interfaccia web è studiata con criteri di semplicità ed efficienza e richiede semplicemente l'uso di un browser con capacità Ajax, senza particolari requisiti hardware (ad esempio mantenendo la compatibilità con dispositivi economici come il laptop da \$100 del M.I.T.).

Per la parte fotografica, VRC può estrarre informazioni dai metadati presenti nei formati di immagine più diffusi (EXIF, XMP) per facilitare l'attività di catalogazione. Nella roadmap del progetto è inclusa la capacità di archiviare immagini in formato "camera raw" proprietario o aperto, e convertirle al volo, in toto o in parte, in formati più facilmente fruibili sul web (JPEG, PDF) o per scopi editoriali. Nella roadmap di VRC sono incluse le capacità di elaborare immagini HDR, panoramiche, in realtà virtuale e CAD 2D/3D. Infine, l'elaborazione di immagini si basa su una libreria di digital imaging ricca di funzionalità (usata anche dalla NASA) prestandosi all'implementazione di sofisticati algoritmi, qualora necessari per scopi specialistici. È anche supportato il Grid Computing.

Syracuse University in Florence è orientata ad offrire VRC come un progetto *open source* in due sensi: *code* e *data* chiedendo alle altre istituzioni la stessa disponibilità alla condivisione sia del software che dei dati. Il progetto è basato su Java e altre tecnologie open source, garantendo la portabilità su qualsiasi piattaforma (Linux, Windows, Mac OS X).

Aggiornamento (2013)

Gli sviluppi del progetto VRC successivi alla realizzazione di questo lavoro non sono stati ulteriormente seguiti dagli scriventi a seguito dell'interruzione del rapporto del lavoro di uno di essi con la Syracuse University of Florence. Il tema dell'uso delle tecniche di web semantico per forme di archiviazione on line di elementi del Patrimonio è stato peraltro affrontato nell'ambito di altri progetti quali ad esempio *blueOcean* (Giudici e Palombini 2010).

Notes

* Tidalwave s.a.s.

† Syracuse University in Florence.

‡ Virtual Heritage Lab, ITABC-CNR, Roma.

¹ Bianchi Bandinelli 1976, p. 12.

² Eco 1968.

³ Bradley, Lei e Blackall 2007.

Riferimenti bibliografici

Bianchi Bandinelli, R. (1976). *Introduzione all'archeologia*. Roma Bari: Laterza.

Bradley, K., J. Lei e C. Blackall (2007). *Towards an Open Source Repository and Preservation System*. UNESCO.

Eco, U. (1968). *La struttura assente*. Milano: Bompiani.

Giudici, Fabrizio e Augusto Palombini (2010). «blue Ocean: un framework open source per la realizzazione di CMS semantici». In: *ARCHEOFOSS Atti del IV Workshop: "Open Source, free formats, open software nei processi di ricerca archeologica*. A cura di P. Cignoni, A. Palombini e Pescarin S. Archeologia e Calcolatori, supplemento. Firenze: All'Insegna del Giglio, pp. 245-249.

Palombini, Augusto (2001). «Miti e pregiudizi nell'uso dei sistemi di classificazione in archeologia». In: *Rassegna di Archeologia* 18, pp. 127-144.