

Un authority file per i luoghi e le istituzioni: l'Atlante Storico Istituzionale dell'Italia unita

Stella Di Fazio

ICAR - Istituto Centrale per gli Archivi

Abstract. Il contributo vuole evidenziare il processo di recupero e valorizzazione di dati pubblici e autorevoli rinvenibili in fonti diverse per età e formato alla base della realizzazione dell'Atlante Storico Istituzionale dell'Italia unita (ASI), strumento di valenza generale che si è posto l'obiettivo di documentare le evoluzioni nel tempo delle circoscrizioni territoriali dell'Italia postunitaria, riconducendo progressivamente questo patrimonio informativo eterogeneo ad un modello ontologico unitario, facilmente riusabile in qualunque contesto grazie alla sua pubblicazione in modalità Linked Open Data. L'attuale ontologia ASI, alimentata ad oggi dai dati relativi agli enti territoriali dello stato italiano, sarà gradualmente estesa con l'introduzione nel modello di ulteriori tipologie di circoscrizioni territoriali, nonché del riferimento alle istituzioni che vi hanno operato e alle carte da queste prodotte, in modo da realizzare una dorsale informativa basilare per la descrizione e la contestualizzazione del patrimonio archivistico nazionale.

Keywords. Interoperabilità, Linked Open data, Semantic Web, Ontologie geografiche

1. Introduzione

Authority files e vocabolari controllati costituiscono architetture informative di supporto alla descrizione e alla contestualizzazione del patrimonio culturale, che agevolano la normalizzazione e il controllo d'autorità dei dati evitando la frammentazione degli interventi e quindi la dispersione o la duplicazione delle attività, ed innalzando al contempo il livello qualitativo dell'informazione prodotta e resa disponibile all'utenza. Per la loro centralità nel processo descrittivo è opportuno che tali strumenti vengano formalizzati in modo tale da risultare facilmente integrabili e/o referenziabili nei diversi contesti informativi: questa è una delle principali motivazioni che hanno indotto l'Istituto centrale per gli archivi (ICAR) a procedere verso la realizzazione progressiva di questo tipo di infrastrutture informative ricorrendo ai formalismi e alle tecnologie del Semantic Web, con l'obiettivo di arrivare gradualmente a costituire una "dorsale informativa" integrata che possa fungere da essenziale punto di riferimento e di snodo per la produzione dell'informazione descrittiva relativa al cultural heritage, e al patrimonio archivistico in primis, come pure per la sua fruizione sul Web.

Alcuni di questi strumenti sono nati all'interno di contesti di fruizione molto ben delimitati, come i vocabolari controllati confluiti nel Tesoro in formato SKOS del Sistema Archivistico Nazionale (SAN) o, sempre in relazione al SAN, il dataset in formato LOD dei soggetti conservatori "d'autorità" della documentazione archivistica in esso censita e descritta. L'Atlante storico istituzionale dell'Italia unita (ASI) si evidenzia invece come uno

strumento di valenza generale, realizzato dall'ICAR con l'obiettivo di rendere disponibile e fruibile in un contesto di uso aperto e interoperabile una grandissima quantità di dati storico-territoriali provenienti da fonti autorevoli di natura pubblica, riconducendola ad un modello ontologico unitario, che consente di rappresentare sia la dimensione diacronica che una visione sincronica delle porzioni di territorio prese in esame.

Censendo le circoscrizioni territoriali dell'Italia postunitaria, a partire dalle circoscrizioni amministrative (base informativa della prima release del maggio 2018) e registrandone i cambiamenti nel tempo, l'ontologia ASI costituisce come esito basilare un repertorio di autorità di nomi di luogo storicamente attestati e georeferenziati, concreto e autorevole supporto per denominare, descrivere e localizzare correttamente le istituzioni collegate a tali luoghi, e le fonti archivistiche da queste prodotte, ma che può allo stesso modo essere utilmente preso a riferimento in qualunque contesto descrittivo, laddove sussista l'esigenza di fornire informazione relativa all'evoluzione storica del territorio. La base di conoscenza in formato RDF OWL è stata alimentata attraverso un lungo processo di studio, analisi e riuso di almeno due diverse tipologie di fonti.

2. Il riuso di dati aperti

La principale fonte di alimentazione è costituita dal SISTAT (Sistema Informativo Storico delle Amministrazioni Territoriali), i cui dati sono stati resi disponibili come open data in formato excel dall'ISTAT. Il SISTAT raccoglie e gestisce le informazioni relative agli enti territoriali definiti nell'ordinamento costituzionale italiano (regioni, province, comuni), a partire dal 1861 ai giorni nostri; ne vengono forniti dati quali il nome, la data di istituzione, e quella di ciascuna variazione occorsa nel tempo fino all'eventuale estinzione (ciascuna corredata dai relativi riferimenti normativi), la dipendenza amministrativa da enti gerarchicamente superiori, ecc. Altra fonte di particolare rilevanza, cui si è fatto ricorso al momento solo parzialmente, è costituita da una considerevole quantità di dizionari storici dei comuni, compilati su mandato governativo a seguire l'Unità d'Italia, che a partire dal livello amministrativo comune raccolgono per specifici anni (quindi in tagli sincronici) dati relativi all'assetto istituzionale del territorio (per ogni Comune i riferimenti a Mandamento, Provincia o Prefettura, Circondario amministrativo o Sotto-prefettura, Tribunale di circondario, Corte d'appello, Diocesi).

A partire da tali dati si è avviato un processo di rielaborazione, integrazione, validazione e arricchimento informativo (principalmente conseguito integrando i livelli circondario e mandamento assenti dal SISTAT) che ha portato alla definizione di una ontologia del territorio dello Stato italiano che rispetto ai dati originari presenta alcune significative evoluzioni innovative.

3. Ontologia e Linked Open data ASI

Il modello ontologico definito è costruito attorno al concetto di regione spazio temporale, entità astratta attraverso cui vengono descritte le caratteristiche (denominazione, estensione territoriale, dipendenza amministrativa, ecc.) stabilmente attribuibili ad una determinata porzione di territorio (nei termini dell'ontologia unità territoriale) per l'arco

temporale definito come l'intervallo tra un evento che genera un cambiamento e il successivo. Di norma l'evento di cambiamento corrisponde ad un provvedimento normativo che determina la variazione (istituzione, cessione/acquisizione di territorio, mutamento di dipendenza gerarchica, estinzione, ecc.). È il concetto di regione spazio temporale che consente di ottenere, interrogando opportunamente l'ontologia, la rappresentazione esatta della configurazione amministrativa del territorio italiano in un determinato istante nel tempo. Da un punto di vista diacronico, invece, la storia di ogni specifica unità territoriale è costituita dalla successione di tutte le regioni spazio temporali ad essa riferibili.

Attualmente l'ontologia espone i dati relativi a una specifica sottoclasse di unità territoriale, l'unità amministrativa, istanziata da risorse RDF OWL corrispondenti a regioni, province e comuni dello Stato italiano, dall'unità ai giorni nostri; i dati relativi agli altri due livelli amministrativi, circondari e mandamenti, sono stati invece forniti solo per i primi anni successivi alla proclamazione del Regno d'Italia, con l'obiettivo di completarne progressivamente il quadro storico arrivando fino al 1927.

I dati dell'ontologia ASI sono raggiungibili attraverso due distinte piattaforme di pubblicazione e fruizione, una – l'endpoint SPARQL dati.san.beniculturali.it/sparql - destinata a forme di accesso e di riuso dei dati di tipo automatico; l'altra – una applicazione web raggiungibile all'indirizzo <http://dati.san.beniculturali.it/asi/local/> – è costituita da un sistema di ricerca e navigazione a faccette pensato invece per agevolare l'interazione coi dati RDF da parte di utenti umani.

I dati OWL di ASI sono a tutti gli effetti Linked open data, essendo stati inseriti link semantici alle corrispettive risorse delle principali ontologie geografiche disponibili (GeoNames, ISPRA, ISTAT, oltre ovviamente alla base di conoscenza di DBpedia), garantendo così forme di interoperabilità estesa col resto del dominio geografico. Sono stati previsti anche collegamenti semantici tra le risorse RDF rappresentanti le unità territoriali di ASI e le corrispettive risorse presenti nei dataset LOD del SAN (corrispondenti alla sede del soggetto conservatore e del soggetto produttore dei complessi archivistici descritti nel Sistema).

4. Risultati conseguiti e sviluppi futuri

Già in questa sua prima release l'Atlante Storico Istituzionale dell'Italia unita ha conseguito una serie di vantaggi derivati principalmente dalla rappresentazione in formato RDF dei dati. Un primo risultato di rilievo è rinvenibile nell'aver convogliato e reso immediatamente disponibili in un unico punto di accesso online aperto e di facile utilizzo informazioni complesse arricchite, ridefinite, e potenziate dai processi di interlinking semantico. Un ulteriore vantaggio risiede nella possibilità di riferirsi a ciascuna delle unità territoriali descritte dall'ontologia come risorsa RDF referenziata da un URI, richiamabile all'interno della descrizione di qualunque risorsa culturale (che si gioverà così automaticamente di tutto il potenziale informativo dell'istanza ASI. Infine, con ASI è stato messo a punto uno strumento fortemente caratterizzato dalla interdisciplinarietà: trattando dimensioni informative universali come lo Spazio, il Tempo, ed Eventi di carattere storico-istituzionale collocati su tali direttrici, l'intelaiatura informativa fornita da ASI risulta un fundamenta-

le supporto alla descrizione di qualunque tipologia di risorsa culturale.

L'Atlante Storico Istituzionale dell'Italia unita è stato concepito come un sistema aperto e in espansione: il modello si presta per sua natura ad essere progressivamente esteso con entità, concetti e dati relativi ad altre tipologie di circoscrizioni territoriali (giudiziarie, ecclesiastiche, militari, ecc.), per realizzare una sempre più esaustiva rappresentazione dell'evoluzione storico istituzionale italiana rapportata al territorio. L'ICAR intende dare avvio ad una nuova fase operativa del progetto indirizzata allo studio e all'analisi delle circoscrizioni territoriali di tipo giudiziario e della loro evoluzione nel tempo; in questa fase verrà anche esplicitato formalmente nel modello il legame tra luoghi, istituzioni che in tali luoghi hanno sede, e documentazione archivistica da queste prodotta. Questa estensione dell'ontologia permetterà di porre l'Atlante come base informativa centralizzata per la descrizione e la contestualizzazione del patrimonio archivistico nazionale.

Riferimenti bibliografici

Bonini, F., Blanco, L., Mori, S., Galluccio, F. (a cura di.) (2016), *Orizzonti di cittadinanza. Per una storia delle circoscrizioni amministrative dell'Italia unita*. Soveria Mannelli, Rubbettino editore, 2016. Cfr. in particolare: Carucci, P., Di Fazio, S., *Dalla Guida generale degli Archivi di Stato italiani all'Atlante storico istituzionale*, pp. 445-460.

Ontologia ASI, <http://dati.san.beniculturali.it/lode/asi/ASI.owl>.

Autrice

Stella Di Fazio - stella.difazio@beniculturali.it

Archivista, si è dedicata prevalentemente allo studio e all'analisi dei linguaggi formali e degli standard descrittivi archivistici. Si occupa da tempo di rappresentazione della conoscenza attraverso gli standard del Semantic Web, ed ha preso parte a numerosi progetti promossi da amministrazione archivistica, università ed enti di ricerca, che hanno portato alla definizione di ontologie informatiche di ambito archivistico e alla produzione di Linked Open Data. Dal febbraio 2018 è archivista di Stato presso l'ICAR.