

Session 2

Open Data in Archaeology

Papers

2.1 MOD (Mappa Archaeological Open Data archive): new ideas for new minds

Francesca Anichini, Marco Ciurcina

Fabio Fabiani, Gabriele Gattiglia,

Francesco Ghizzani Marcà, Maria Letizia

Gualandi, Claudia Sciuto

MAPPAProject - University of Pisa

Valerio Noti

Terrelogiche

Il MOD (MAPPAProject archaeological Open Data archive), realizzato dall'Università di Pisa e finanziato dalla Regione Toscana (http://mappaproject.arch.unipi.it/?page_id=454), è la prima esperienza di archivio open data per l'archeologia italiana. Pubblicato nel giugno 2012 in una versione beta, oggi il MOD consta di numerosi archivi che raccolgono un campione significativo degli interventi effettuati a Pisa e in altri siti italiani.

1. La struttura informatica

Il MOD è un'applicazione web sviluppata su piattaforma tecnologica Open Source e consultabile attraverso browser compatibili con gli standard W3C. È composta di due segmenti distinti: una sezione amministrativa con accesso riservato, che consente l'inserimento e la modifica dei contenuti da parte degli operatori, ed una sezione pubblica accessibile dagli utenti (<http://www.mappaproject.org/mod>) e dotata di funzionalità di ricerca avanzata sugli archivi. L'applicazione è basata sui singoli interventi archeologici definiti da una serie di informazioni (titolo, autore, DOI, regione, località, anno, contatto primario, introduzione, overview) e dal collegamento con fonti dati eterogenee (documenti testuali, immagini, oggetti multimediali, files di database, dati geografici, ecc.) archiviate nel repository associato. Per gli interventi che rientrano nell'area oggetto d'indagine del progetto MAPPAProject, corrispondente alla città di Pisa, è disponibile la consultazione della componente geografica attraverso il collegamento diretto con il webGIS MAPPAProject (MAPPAGIS, <http://www.mappaproject.org/webgis>). In prospettiva di un progressivo e sostanziale

incremento dei dataset presenti, si prevede l'implementazione di un sistema di accounting attraverso il quale gli utenti potranno accedere alla piattaforma ed eseguire direttamente l'upload degli archivi, con uno step di verifica intermedia della correttezza formale dei documenti da pubblicare. Questa evoluzione garantirà un approccio maggiormente partecipativo, in linea con i portali Open Data già presenti in altri settori dell'informazione scientifica e tecnica, e un'accelerazione nell'implementazione del MOD.

2. Aspetti legali e modalità di pubblicazione

Oltre alla costruzione della struttura informatica, la realizzazione del MOD ha richiesto lo studio delle problematiche legali ed etiche connesse alla pubblicazione aperta dei dati archeologici, attraverso un percorso di analisi della legislazione vigente e di confronto tra questa e le prassi consolidate. Ad oggi, infatti, lo stato delle conoscenze, legate principalmente alle consuetudini adottate in ambito archeologico (universitario, ministeriale, professionale), risulta lacunoso e spesso discordante dalle norme sul diritto d'autore e sulla privacy, dal codice della proprietà industriale e dal codice dei Beni Culturali. Attraverso un serrato e fruttuoso confronto con la Soprintendenza per i Beni Archeologici della Toscana (di seguito SBAT) e con la Direzione Regionale per i Beni Culturali e Paesaggistici della Toscana, l'équipe di ricercatori e legali del progetto MAPPAProject ha cercato di fare luce sulla regolamentazione di alcuni punti specifici quali, fra gli altri, il riconoscimento della titolarità dei diritti sui dati, il ruolo del direttore scientifico nella pubblicazione open data della documentazione archeografica, le modalità di pubblicazione delle riproduzioni grafiche e fotografiche, le responsabilità dell'autore rispetto alla struttura/ente che pubblica online i dati (Università di Pisa), la tutela dei dati personali presenti all'interno dei documenti pubblicati, la ricognizione e la gestione di limiti imposti da accordi contrattuali tra archeologi e committenti del lavoro, le tipologie di licenze adottabili ecc...

Per rendere chiare le scelte adottate alla luce di questa analisi, sono stati predisposti nuovi documenti di corredo alla pubblicazione nel MOD che illustrano agli utenti i criteri adottati, gli obblighi e le condizioni di chi acconsente alla pubblicazione dei dati (autore), di chi gestisce il servizio di pubblicazione on-line (Università di Pisa), di chi usufruisce dei dati pubblicati, di chi, infine, ha rilasciato le eventuali autorizzazioni alla pubblicazione. I confini e le modalità di utilizzo sono stati sintetizzati all'interno della piat-

taforma nei "Termini di utilizzo" che comprendono le "Condizioni di servizio" (<http://mappaproject.arch.unipi.it/mod/condizioni.php>) e il "Disclaimer" (<http://mappaproject.arch.unipi.it/mod/disclaimer.php>). L'autorizzazione da parte del MiBAC alla diffusione degli elaborati tramite il portale MOD non poneva problemi, stante la partecipazione al progetto delle Soprintendenze competenti; è stato così previsto di acquisire autorizzazione da parte dei singoli autori predisponendo un'adeguata "Liberatoria" che si chiede di sottoscrivere la pubblicazione di ogni archivio. Per definire le "Condizioni di servizio" è stato necessario specificare innanzitutto quali fossero i "contenuti". Obiettivi del MOD sono divulgare, conservare, dare visibilità e attribuire un riconoscimento effettivo a tutta la produzione archeografica legata ad un'indagine archeologica che solitamente rimane inedita. Parliamo quindi di quel materiale che rappresenta l'unica fonte di informazione di un determinato intervento: relazioni preliminari, report, elenchi e schede US e USM, elenchi reperti, schede di quantificazione, tabelle di periodizzazione, diagrammi stratigrafici, elenchi Attività, schede di archiviazione veloce, diari di scavo, lettere e comunicazioni, planimetrie, sezioni, prospetti, documentazione grafica e fotografica, registrazioni audio e/o video, dati organizzati in banche dati.

Parallelamente sono state precisate le modalità di pubblicazione nel MOD. L'idea che ha guidato la riflessione e l'analisi della legislazione sul diritto d'autore è stata comprendere innanzitutto a chi debba essere riconosciuta la paternità della documentazione prodotta durante un intervento archeologico. Alla luce delle norme della legge sul diritto d'autore, per aventi diritto si devono intendere gli estensori della documentazione i quali, in quanto autori, dispongono del diritto d'autore e degli altri diritti connessi previsti nella legge sul diritto d'autore che insistono sulla documentazione, fatto salvo il caso in cui nel contratto di affidamento del lavoro, od in altro modo, abbiano ceduto tale diritto. Le "condizioni" specificano pertanto come "ciascun contenuto è associato all'Utente che ha provveduto a pubblicare lo stesso contenuto e, se quell'Utente lo desidera, il Servizio rende nota la qualità di autore dello stesso Utente" (http://mappaproject.arch.unipi.it/mod/docs/Condizioni_Servizio.pdf).

La paternità dei dati è gestita mediante l'apposizione di DOI (Digital Object Identifier) distinti per la sezione "Relazione" (che comprende la letteratura grigia) e per la sezione "Dataset" (in cui è raccolta tutta la restante documentazione), i cui autori spesso possono non coincidere.

Agli autori che intendono pubblicare sul MOD si chiede di verificare che i documenti siano privi di "dati personali di persone fisiche (incluse fotografie, planimetrie e/o altri contenuti che si riferiscano ad abitazioni di proprietà o utilizzate da persone fisiche), a meno che queste non abbiano "dato il consenso al trattamento dei loro dati personali"; tale impegno è specificato nella Liberatoria a firma di ciascun autore. Laddove non sia stata richiesta o rilasciata la specifica autorizzazione alla pubblicazione dei dati

personali, questi sono coperti da omissis, mentre nelle fotografie sono resi irriconoscibili i volti delle persone e illeggibili eventuali elementi identificativi, quali targhe automobilistiche, insegne, ecc...

Le condizioni di servizio specificano inoltre agli utenti il tipo di licenza scelta dall'autore per liberare i propri dati, se CC BY o CC BY-SA. Un capitolo a parte è rappresentato dalla pubblicazione delle fotografie e riproduzione di Beni Culturali. Per quanto l'archeologo possa essere l'autore materiale della riproduzione, ai sensi degli artt. 106 e sgg. del D. Lgs. 22 gennaio 2004 n. 42, egli è tenuto a chiedere l'autorizzazione al MiBAC - e più precisamente alla Soprintendenza competente - poiché essa ritrae un bene culturale. Nel case study del progetto MAPPA (il territorio urbano di Pisa), l'autorizzazione alla pubblicazione delle fotografie è stata concessa dalla SBAT in virtù di un'apposita convenzione che consente l'utilizzo di immagini senza gli oneri di cui all'art. 108 del Codice dei beni Culturali (ANICHINI *et alii* 2012: 165-170). Con specifico accordo è stato inoltre definito che le fotografie pubblicate nel MOD abbiano dimensione massima di 1024 pixel per il lato lungo e quelle di proprietà esclusiva della Soprintendenza (non scattate cioè da soggetti terzi, che detengono diritti ai sensi della legge sul diritto d'autore) abbiano il marchio del MiBAC in sovraimpressione.

3. Uno sguardo ad un futuro già presente

È doveroso precisare che il MOD è uno strumento ancora ampiamente migliorabile. Per quanto nell'immediato futuro sarà indispensabile lavorare sul formato dei dati e sulla loro metadattazione, l'avvio del MOD non è stato volutamente vincolato a standard precisi, cercando invece di coinvolgere il maggior numero di archeologi ad aprire i loro dati e sensibilizzare, anche in forma critica, tutta la comunità archeologica sui temi dell'open data e dell'open access. Riteniamo infatti necessario che si faccia sempre più ampia la riflessione sulle grandi opportunità scientifiche, professionali e di tutela che sono intrinseche a questo nuovo modo di approcciarsi ai dati e alla loro disseminazione. Per quanto a distanza di un anno dalla prima messa in rete del MOD le politiche e le posizioni sembrano essere maggiormente aperte alla filosofia open, rimangono ancora forti la diffidenza verso lo sharing, i dubbi sul riconoscimento della paternità intellettuale a chi ha prodotto materialmente il dato, lo scetticismo nei confronti dell'attribuzione di dignità scientifica alla pubblicazione dei raw data archeologici.

Il MOD si propone quindi come concreto modello operativo che può spostare la discussione nella concretezza di una piattaforma già in uso. Se in Italia era "importante partire" (ANICHINI, GATTIGLIA 2012: 52), adesso è indispensabile non arrestarsi; è importante confrontarsi, discutere, ma, contemporaneamente, caricare, condividere, scaricare, aprire i dati.

Bibliografia

ANICHINI F., GATTIGLIA G., GUALANDI M.L., NOTI V. C.S., *MOD (Mappa Open Data), Conservare, disseminare, collaborare: un archivio open data per l'archeologia italiana*, in SERLORENZI M. (a cura di), *Open Source, Free Software e Open Format nei processi di ricerca archeologica, Atti del VII Workshop (Roma 11-13 giugno 2012)*, «Archeologia e Calcolatori», Supplemento 4.

ANICHINI F., FABIANI F., GATTIGLIA G., GUALANDI M.L. .2012, *MAPPA. Metodologie Applicate alla Predittività del Potenziale Archeologico*, vol.I, pp. 165-170.

ANICHINI F., GATTIGLIA G. 2012, *#MappaOpenData. From web to society. Archaeological open data testing*, in GUALANDI M.L., *MapPapers 2/2012*, pp.51-54.

CIURCINA M., GROSSI P., *Open data: alcune considerazioni sulla pubblica amministrazione e sui beni culturali e paesaggistici in Italia*, in M. SERLORENZI (a cura di), *Atti della seconda giornata di studio sul SITAR - Roma, Palazzo Massimo alle Terme, 9 novembre 2011*, Roma, 2013

2.2 The Semantic Web and the Digital Archaeological Workflow: a Case Study from Sweden

Marcus Smith

Swedish National Heritage Board

Two related projects from the Information Development Unit at the [Swedish National Heritage Board](#) will be presented: SOCH, an existing system which aggregates and semantically links diverse cultural heritage datasets from a variety of sources; and DAP, an ongoing programme of work to digitise the Swedish archaeological process.

SOCH

[Swedish Open Cultural Heritage](#) is a web platform which aggregates metadata from a variety of Swedish cultural heritage institutions, and makes it accessible as RDF linked open data via an API. There are almost 40 different institutions in Sweden delivering data to SOCH, including the National Monuments Register, the Historic Buildings Register, and a number of national, regional, and local museums. SOCH currently holds information on 4.7 million objects with coverage of artefacts and monuments ranging from the palaeolithic up until the present day.

As an aggregator, SOCH enriches the metadata it harvests with semantic links before releasing it as open data. This linking facilitates cross-searching of the different datasets, and makes the aggregated data more accessible than it would be on its own.

An 'object' in SOCH is an entry in a database: it might well be a description of a physical artefact in a museum catalogue, or a drawing or photograph, but it might just as well be an ancient monument, a historic building, a digitised archaeological report or other documentation, a person, event or thematic grouping. At the time of writing, SOCH contains semantically linked information about, among other things:

- 2.1 million artefacts
- 880 thousand photographs
- 830 thousand monuments
- 440 thousand documents
- 110 thousand historic buildings
- 40 thousand personages
- 2000 historical events
- 1500 historic maps

Because SOCH links together objects from different datasets and collections in a meaningful way and through a common interface, it makes finding information on a particular topic, place or thing simpler. Searches need not be carried out individually using