

Tra diritti e doveri. L'open data nell'archeologia italiana

Francesca Anichini, Marco Ciurcina, Fabio Fabiani,
Gabriele Gattiglia, Maria Letizia Gualandi

The MAPPA Open Data (MOD) archaeological repository (www.mappaproject.org/mod) stores and considers openly accessible all kind of archaeological data from raw data to linked open data produced during the research process, including metadata, and it's connected with the Journal of Open Archaeological Data for good quality data. Our infrastructure is hosted on GNU/Linux of the Centro Interdipartimentale di Servizi Informatici per l'Area Umanistica (CISIAU) of University of Pisa, is designed on an Open Source LAMP technological platform using an Apache HTTP Server, PHP 5.x scripting language and MySQL Open Source relational database.

Currently, as our main purpose was to persuade the archaeological community of the importance of open data, we use a basic policy for data management: we acquire raw data from archaeologists, and we validate the data from a legal point of view. Various types of records are published on the MOD which refer to archaeological interventions: preliminary reports, reports, context lists, context records, lists of finds, quantification records, period tables, stratigraphic diagrams, lists of activities, masonry contexts, masonry context records, rapid archiving records, excavation diaries, letters and communications, maps, sections, drawn records and photographs. Consequently, we checked the legitimacy of the activities carried out in the MOD with the records (uploading and dissemination to the public via the Internet) in the light of the following regulations:

- Law no. 633 of 22 April 1941, "Protection of copyright and rights related to its exercise" (LdA);
- Legislative Decree no. 42 of 22 January 2004, "Code of the cultural and landscape heritage, pursuant to article 10 of Law no. 137 of 6 July 2002" (CBC);
- Legislative Decree no. 196 of 30 June 2003, "Personal data protection code" (Privacy Code);
- Legislative Decree no. 30 of 10 February 2005, "Industrial property code, in accordance with article 15 of Law no. 273 of 12 December 2002" (CPI).

We published a detailed guide in which are explained the procedures that must be followed to prepare and provide the material to be published. In compliance with the laws, published documents are not expected to contain the personal data of natural persons who have not previously agreed to their publication, whether they are authors or third parties mentioned inside the records. Specific disclaimers have been prepared and can be downloaded to help authors correctly collect the authorisations needed to put their material online.

Once validated the data, we embed metadata to each dataset describing all the information regarding the dataset itself, we stored the data in our repository, providing protection. We use a metadata schema for each dataset describing all the information regarding the dataset itself: the structure and format of the digital data, the history of the archaeological investigation, the sources used, the method and the relationship with the physical data. The schema is composed partly from Dublin Core and partly from ISO 19115 metadata core for the geographical section. We don't describe the quality of data, because we firmly believe that the quality of research data must be responsibility of researchers in a sort of open peer review method.

Negli ultimi tempi parlare di open data è finalmente diventato usuale: tutti li citano, qualcuno li vuole, pochi li mettono in rete (almeno per ora). Oggi in Italia i dati archeologici non circolano liberamente e un'archeologia 2.0 italiana appare ancora una realtà molto lontana. Non ci riferiamo tanto ai dati inseriti in pubblicazioni ufficiali – che peraltro, viste le difficoltà economiche degli ultimi anni, sono sempre meno – ma ai *raw data* dell'archeologia, ovvero a tutti i dati archeografici prodotti nel corso di un'indagine: disegni, fotografie, schedature, diagrammi, report ecc., vale a dire a quella documentazione che, in molti casi, è tutto ciò che resta di un patrimonio che la stessa pratica archeologica concorre parzialmente a distruggere. Lo scavo archeologico, infatti, è un'attività non ripetibile: l'unico elemento di riproducibilità e ri-analisi è costituito dalla continua possibilità di attingere ai dati grezzi, attraverso i quali la comunità scientifica può ripercorre le tappe del processo interpretativo e formulare nuove ipotesi e ricostruzioni storiche, per non dire della possibilità di riuso delle informazioni su scale e con finalità diverse.

Oggi il MOD (MAPPA archaeological Open Data archive), creato all'interno del più ampio progetto di ricerca *MAPPAProject - Metodologie Applicate alla Predittività del Potenziale Archeologico*, realizzato dall'Università di Pisa e finanziato dalla Regione Toscana (http://mappaproject.arch.unipi.it/?page_id=454), si pone come la prima e ancora unica esperienza concreta di archivio open data per l'archeologia italiana (ad Agosto 2013 il MOD è entrato nella lista dei repositories raccomandati dal Journal of Open Archaeological Data (<http://openarchaeologydata.metajnl.com/about/editorialPolicies#custom-0>; ANICHINI *et al.* 2013). Oltre alla costruzione della struttura informatica, palesemente ispirata a quella ormai collaudata dell'*Archaeological Data Service* dell'Università di York, la realizzazione del MOD ha richiesto un grosso lavoro di studio di tutte le problematiche legali ed etiche connesse alla pubblicazione aperta dei dati archeologici (CIURCUNA 2013). In questo campo lo stato delle conoscenze, legate principalmente alle consuetudini adottate in ambito archeologico (universitario, ministeriale, professionale), si è infatti rivelato lacunoso e spesso discordante dalle norme sul diritto d'autore e sulla *privacy*, dal codice della proprietà industriale e dal codice dei Beni Culturali. L'*équipe* di ricercatori e legali del progetto MAPPA ha così deciso di intraprendere un percorso di analisi della legislazione vigente e di confronto tra questa e le prassi consolidate nel mondo archeologico, in vista delle sfide che la filosofia open data pone specificatamente nel settore dei Beni Culturali. È stato avviato un serrato e fruttuoso confronto con la Soprintendenza per i Beni Archeologici della Toscana e con la Direzione Regionale per i Beni Culturali e Paesaggistici della Toscana in merito alle modalità operative generali e su punti di confronto 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 immagini/fotografie, 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.

Da queste riflessioni è scaturita una revisione della versione beta del MOD, già *on line* da giugno 2012. Se non è stato necessario fare sostanziali cambiamenti nella struttura informatica e nemmeno nelle interfacce utente della versione già licenziata, è stato invece indispensabile rivedere i criteri d'inserimento dei documenti e la loro trattazione prima della pubblicazione *on line*. È stato inoltre essenziale riscrivere i documenti 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 (Soprintendenza per i Beni Archeologici della Toscana e terzi, diversi dagli autori). I confini e le modalità di utilizzo

del MOD sono stati sintetizzati in due documenti pubblicati all'interno della piattaforma: le "Condizioni di servizio" (<http://mappaproject.arch.unipi.it/mod/condizioni.php>) e il "Disclaimer" (<http://mappaproject.arch.unipi.it/mod/disclaimer.php>). I diritti e gli obblighi degli autori, invece, sono stati tradotti in una "Liberatoria" che ciascun autore deve sottoscrivere per ogni intervento che intende pubblicare.

Per definire le "Condizioni di servizio" è stato necessario specificare innanzitutto quali fossero i "contenuti": il MOD vuole essere il primo archivio di dati grezzi archeologici italiani, ma cosa intendiamo comunemente con dati grezzi? Visioni contrastanti, che probabilmente scaturiscono dall'assenza di definizioni condivise e standard precisi, inseriscono nei dati grezzi elementi diversi di quella che può essere genericamente considerata la documentazione in campo archeologico. Nel MOD si è scelto di definire "contenuti" dell'archivio tutti i "file contenenti relazioni preliminari, report, elenchi Unità Stratigrafiche (US), schede US, elenchi reperti, schede di quantificazione reperti, tabelle di periodizzazione, diagrammi stratigrafici, elenchi Attività, elenchi Unità Stratigrafiche Murarie (USM), schede USM, schede di archiviazione veloce (SAV), diari di scavo, lettere e comunicazioni, planimetrie, sezioni, rappresentazioni grafiche e fotografie ed altri testi, immagini, registrazioni audio e/o video, dati e/o informazioni organizzati in banca di dati (...)". Obiettivo del MOD è infatti dare massima diffusione a tutto il materiale primario prodotto nell'ambito di un intervento archeologico di qualunque entità, lo stesso materiale che rimane per lo più inedito e che è ciò che consente la verifica diretta delle analisi sintetiche o delle ricostruzioni storiche che spesso rappresentano gli unici elaborati che circolano nella comunità scientifica e non. Per non dire del fatto che vi sono interventi archeologici che, per vari motivi, non arrivano mai a essere pubblicati e per i quali i dati grezzi rimangono l'unica forma di documentazione (a giugno 2012 il MOD conteneva 13 archivi, oggi 119).

Se si parte dalla considerazione che la pratica archeografica – sia essa frutto del lavoro di professionisti o di ricercatori afferenti a strutture pubbliche come le Università e le Soprintendenze – è sempre e comunque un'attività di ricerca, dal momento che produce dati unici e irripetibili, e che non vi è ricerca fin quando non vi è pubblicazione del dato, appare evidente come la condivisione dei dati grezzi e della letteratura grigia debba essere considerata come una pubblicazione a tutti gli effetti, salvaguardando sia le competenze che la capacità professionale e l'impegno, anche temporale, profuso da chi quel dato ha prodotto con il suo lavoro (ANICHINI, GATTIGLIA 2012).

Parallelamente sono state specificate le modalità di pubblicazione ("oggetto") nel MOD. L'idea (già delineata ampiamente in ANICHINI *et al.* c.s.) 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 intellettuale della documentazione prodotta durante un intervento archeologico. La versione beta del MOD riportava come autore della relazione dell'intervento il firmatario della stessa e, come autori del dataset contenente la restante documentazione, gli archeologi firmatari e il direttore scientifico dell'intervento (molte volte coincidente con il funzionario responsabile della Soprintendenza). La fase di studio ha portato a riconsiderare che, alla luce delle norme sul diritto d'autore, per aventi diritto si devono intendere gli estensori della documentazione indipendentemente da chi ha avuto la direzione scientifica dell'intervento e da chi (ditta esterna, ente locale o istituzione) lo ha finanziato, i quali, in quanto autori, dispongono del diritto d'autore sugli elaborati, 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".

La paternità intellettuale dei dati è gestita mediante l'apposizione di DOI (Digital Object Identifier) distinti per la sezione "Relazione" (che comprende la letteratura grigia) e la sezione

“Dataset” (in cui è raccolta tutta la restante documentazione), i cui autori possono non coincidere: poiché il lavoro archeologico è quasi sempre un lavoro d'équipe, è usuale imbattersi in relazioni a firma singola o multipla (il/i responsabile/i dell'intervento), mentre gli elaborati che costituiscono la restante documentazione sono realizzati da un numero maggiore di autori. In particolare si pongono questioni diverse per tipologie di prodotti diversi. Per la compilazione delle schede di US, ad esempio, rimane dubbio se sussista in ogni caso un diritto d'autore, come pure per gli schizzi e le piante di US, che si configurano più come schemi e spesso sono successivamente rielaborate da altri soggetti rispettivamente in planimetrie composite e, oggi, vettorializzate in ambiente digitale. All'autore, o agli autori, si chiede di verificare, prima della pubblicazione, la presenza interna ai documenti di “dati personali di persone fisiche (incluse fotografie, planimetrie e/o altri contenuti che si riferiscano ad abitazioni di proprietà di 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: CC BY o CC BY-SA. Firmando la liberatoria l'autore ha la possibilità di scegliere tra le licenze CC BY e CC BY-SA.

Un capitolo a parte è rappresentato dalla pubblicazione delle fotografie e riproduzione di Beni Culturali. Nel *case study* del progetto MAPPA (il territorio urbano di Pisa), l'autorizzazione alla pubblicazione delle foto (ancora ai sensi degli artt. 106 e sgg. del D. Lgs. 22 gennaio 2004 n. 42) è stata concessa a monte dalla Soprintendenza per i Beni archeologici della Toscana, 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 al.* 2012, 165-170). Tutte le fotografie sono pubblicate nel MOD nella dimensione massima di 1024 pixel per il lato lungo; inoltre sulle fotografie di proprietà esclusiva della Soprintendenza (non scattate cioè da soggetti terzi, che detengono diritti ai sensi della legge sul diritto d'autore), viene apposto in sovraimpressione il marchio del Mibact. Con l'entrata in vigore della Legge n. 106 del 29/07/2014 (“Art Bonus”) la prassi adottata in Mappa (e già prima nel progetto “Wiki Loves Monument”), è stata probabilmente normata. Questa è, infatti, la lettura più ragionevole del controverso nuovo comma 3-bis dell'art. 108: le fotografie dei beni possono essere pubblicate con licenza CC-BY-SA in sito che non ha fini di lucro, con idoneo avvertimento che fa salve le previsioni degli artt. 106 e seg. del Codice dei beni culturali e del paesaggio.

È doveroso precisare che il MOD è uno strumento ancora ampiamente migliorabile, ad esempio lavorando sul formato dati e sulla loro metadazione (in corso di ultimazione). Come già ricordato in altre sedi però, il primo obiettivo che ci si è posti nella creazione del MOD è stato l'avvio di una riflessione – non più procrastinabile – sul tema della condivisione dei dati archeologici, allo scopo di sensibilizzare tutta la comunità archeologica sui temi dell'open data (e dell'open access) e di far comprendere le grandi opportunità scientifiche e professionali che sono intrinseche in un nuovo modo di approcciarsi ai dati e alla loro disseminazione. Per raggiungere tale obiettivo, è stato scelto deliberatamente di non vincolare l'inserimento dei dati nel MOD a quelle che sono le buone regole (e le “buone stelle”) che definiscono un dato veramente open, ma di “accontentarsi” di far arrivare alla portata di tutti, aprendo faldoni e cassette, quanto più materiale possibile, nella consapevolezza che un lavoro di ottimizzazione dei formati e dei metadati potrà essere fatto in uno *step* successivo. Sondando un vasto campione di utenti, infatti, è apparso chiaro come, nonostante la sempre maggior diffusione della filosofia open data, siano ancora presenti 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 della dignità di pubblicazione scientifica alla pubblicazione dei *raw data*. È sembrato quindi prioritario iniziare, lanciare il sasso nello stagno attorno al quale si muovono tutti gli attori del mondo archeologico.

Con il MOD vogliamo proporre un modello operativo – e già attivo – che può diventare sede, oltre che di pubblicazione, anche di riflessione comune, nella convinzione che i fatti guidino le parole. Apriamo i dati, pubblichiamoli, condividiamoli in formati sempre più aperti e, contemporaneamente, confrontiamoci su modalità e standard per renderli ancora più utili per il lavoro di noi tutti.

Autori

- Francesca Anichini: Università di Pisa.
- Marco Ciurcina: Studio Legale Marco Ciurcina
- Fabio Fabiani: Università di Pisa.
- Gabriele Gattiglia: Università di Pisa.
- Maria Letizia Gualandi: Università di Pisa.

Bibliografia

ANICHINI F., CIURCINA M., NOTI V. 2013, Il MOD: *l'archivio Open Data dell'archeologia italiana*, in ANICHINI F., DUBBINI N., FABIANI F., GATTIGLIA G., GUALANDI M.L., *MAPP.A. Metodologie Applicate alla Predittività del Potenziale Archeologico, vol.II*, 121-132 doi: 10.4458/0917-09

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, *MAPP.A. Metodologie Applicate alla Predittività del Potenziale Archeologico, vol.I*, 165-170.

ANICHINI F., GATTIGLIA G. 2012, #MappaOpenData. *From web to society. Archaeological open data testing*, in GUALANDI M.L., *MapPapers* 2, 2012, 53. doi: 10.4456/MAPPA.2012.05

CIURCINA M. 2013, *Parere legale sul portale Mappa Open Data*, in "MapPapers", 4, 2013, 87-106 doi:10.4456/MAPPA.2013.76