

ZOTERO2MAP: SVILUPPO DI UN NUOVO STRUMENTO ITC PER LA CONDIVISIONE E LA PUBBLICAZIONE DEI DATI BIBLIOGRAFICI PER LA RICERCA STORICA E ARCHEOLOGICA

1. INTRODUZIONE

Negli ultimi decenni la produzione scientifica si è ampliata in modo esponenziale, grazie alla moltiplicazione delle sedi nazionali e internazionali di confronto scientifico e il numero crescente degli spazi editoriali anche digitali dove pubblicare le ricerche. Anche se molte risorse non sono liberamente fruibili, l'avvento della cultura open access ha notevolmente contribuito alla diffusione delle pubblicazioni aperte, una mole di dati che ha generato nei ricercatori il problema del continuo aggiornamento della propria conoscenza su argomenti generali o specifici (CARVALE, PIERGROSSI 2012, 2015; CARVALE, PIERGROSSI, ROSSI 2021; CARVALE 2022). Infatti, la creazione di una o più collezioni di record bibliografici e il loro aggiornamento sono probabilmente alcune delle più importanti componenti di una ricerca scientifica. In questo ambito i ricercatori dispongono di software specializzati che aiutano a raccogliere in maniera semplificata, gestire, aggiornare, organizzare e usare in modo intuitivo i record bibliografici.

Tra questi il più importante è senza dubbio Zotero (<https://www.zotero.org/>; AHMED 2011; PUCKETT 2011), un diffuso strumento open source che oltre ad offrire un aiuto ai singoli ricercatori, integra anche strumenti collaborativi attraverso il paradigma del cloud computing, come la possibilità di sincronizzare il proprio database su un server online e dividerlo, totalmente o in parte con colleghi e collaboratori, oltre alla possibilità di creare collezioni pubbliche. Zotero, oltre ad essere integrato con i più diffusi web browser, consente di gestire allegati, genera automaticamente citazioni e note e permette di associare ad ogni oggetto della collezione bibliografica uno o più tag. Nonostante queste funzioni, rimane uno strumento di lavoro personale, che il singolo ricercatore può adattare alle proprie esigenze, rendendo spesso difficile una più proficua collaborazione.

Nel corso degli anni l'esigenza di connettere il software open source Zotero con una visualizzazione di tipo cartografico, e quindi di vedere le citazioni su una mappa fruibile online, è emersa tra gli utenti della community e ci sono stati vari tentativi di sviluppare degli applicativi simili, che tuttavia non hanno avuto una diffusione ampia e il cui sviluppo è stato abbandonato, come ad esempio quella della Entropy Free LLC (<https://github.com/zotero/zotero-maps>), il cui ultimo aggiornamento risale a 13 anni fa. Rispetto alle proposte che attualmente sono disponibili in rete, che si appoggiano sull'utilizzo



Fig. 1 – L'applicazione web Zotero2map.

di software come QGIS (TOBIAS 2021, <https://github.com/MicheleTobias/LiteratureMapper>), l'applicazione che viene proposta in questa sede non necessita di alcuna installazione di programmi esterni, poiché sarà direttamente la webApp a collegare le collection degli utenti Zotero con il sistema della mappa online.

Partendo da queste considerazioni il Laboratorio di Archeologia Digitale (LAD, <https://purl.org/lad>) della Sapienza Università di Roma ha sviluppato una applicazione web, denominata Zotero2map¹ (Fig. 1), che consente di estendere le funzionalità di Zotero, con lo scopo di favorire la condivisione dei dati bibliografici provenienti da vari colleghi/collaboratori che vogliono condividere le proprie collezioni bibliografiche, visualizzando i risultati su una mappa. Attraverso questa applicazione è possibile collegare Zotero, che è ormai uno standard *de facto* per la ricerca scientifica, con un webGIS che consente di visualizzare e interrogare i dati bibliografici anche su base geografica. Zotero2map dunque permette l'interscambio di informazioni relative a diversi contesti storici e archeologici, velocizzando il loro reperimento e garantendo un aggiornamento costante dei dati inseriti nel sistema.

L'applicativo svolge un ruolo di aggregatore di contenuti bibliografici da librerie di ricercatori che decidono di collaborare, consentendo oltre alle ricerche tradizionali, anche la visualizzazione su una mappa di ciò che è stato

¹ Lo sviluppo dell'applicazione è stato possibile grazie al finanziamento Programma Progetti per Avvio alla Ricerca della Sapienza Università di Roma stanziato nel 2022 e realizzata dal Laboratorio di Archeologia Digitale (LAD) da E. Iacopini e J. Bogdani.

pubblicato (articoli, monografie, etc.) riguardante una determinata area geografica o uno specifico tema di ricerca che investe diversi ambiti territoriali. Ovviamente il risultato sarà più soddisfacente nella misura in cui i ricercatori decideranno di condividere le loro banche dati. Il coinvolgimento del Laboratorio di Archeologia Digitale della Sapienza in diversi progetti tuttavia consentirà di testare l'applicativo con team di ricerca diversi sia nazionali che internazionali, come ad esempio quello relativo all'Epiro settentrionale.

Infatti, come caso di studio per lo sviluppo e il test di questo strumento software è stata scelta la regione antica dell'Epiro settentrionale, un'area che coincide all'incirca con l'attuale Albania meridionale e la parte nordoccidentale della Grecia. Si tratta di una zona non troppo estesa, che ha conosciuto negli ultimi decenni un'importante stratificazione di studi, condotti da gruppi di ricerca internazionali, spesso espressi in lingue diverse. Nella seconda metà del Novecento, la ricerca storico-archeologica è stata condotta soprattutto da team locali, albanesi per l'area entro i confini albanesi e greci per la parte greca; questa produzione scientifica presenta importanti problemi di accessibilità sia fisica che di contenuto. Negli ultimi trent'anni la ricerca si è aperta a collaborazioni internazionali, che hanno contribuito in maniera determinante all'arricchimento dei contenuti, espressi anche in italiano, francese, spagnolo e inglese. La moltiplicazione degli attori ha portato alla moltiplicazione dei contenuti, della loro qualità e delle sedi di pubblicazione. Questa moltiplicazione va intesa in maniera esponenziale, dal momento che l'aumento della produzione primaria, con nuovi dati e ricerche sul campo in lingue accessibili, rende possibile e amplia lo studio secondario, ovvero la sintesi e analisi anche da parte di chi non opera direttamente sul terreno.

Per questi motivi e per il diretto coinvolgimento del LAD in lavori sul campo in area epirota (<https://purl.org/lad/cuka-e-ajtoit>), la regione è il banco di prova ottimale per questo progetto di ricerca. Come primo step è stata creata una banca dati Zotero relativa all'area geografica dell'Epiro, consultabile ricercando su Zotero il gruppo pubblico "Epirus" (<https://www.zotero.org/groups/336647/epirus/library>), il quale attualmente conta più di 600 elementi, tra cui monografie, articoli da volumi, articoli su riviste o semplici report. Questa banca dati è stata sviluppata in seno alle attività didattiche del LAD, a cui hanno contribuito gli studenti del corso di Metodologia della Ricerca archeologica del prof. J. Bogdani.

Per ciascun record bibliografico sono stati assegnati uno o più tag, sulla base dell'indicazione geografica o per argomento. Ad esempio, per l'articolo *A colossal cult statue group from Dobër, Albania: Visual narratives of East and West in the countryside of Butrint* di Milena Melfi, pubblicato sull'«American Journal of Archaeology» (Fig. 2), sono stati assegnati sette "tag", di cui quattro relativi al posizionamento geografico (Bouthrotos, Cestrine, Dobra e Vagalati) e tre tematici (fortifications, temple e Trojan cicle).

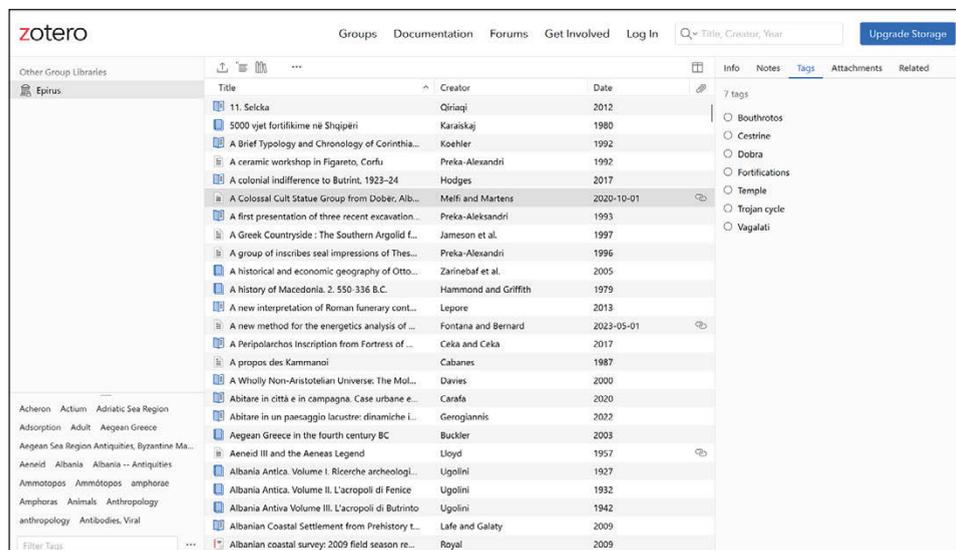


Fig. 2 – Libreria Zotero “Epirus”.

La possibilità di associare queste etichette agli elementi della collezione bibliografica è alla base della funzionalità dell'applicazione di incrociare questi dati con dei dizionari cross-language, una caratteristica, che consente a Zotero2map di avere una vocazione internazionale. Infatti, il sistema prevede di associare i termini relativi ad un medesimo argomento di ricerca o ubicazione geografica, ma trattati in lingue differenti, rendendo dunque la webApp fruibile da ricercatori di tutte le nazionalità. L'utilizzo di dizionari condivisi, multilingua, permetterà non solo di abbattere le barriere linguistiche, ma anche di definire una tassonomia articolata di relazioni tra i termini utilizzati. I singoli ricercatori, qualora intendano partecipare con i propri dati, non dovranno dunque adeguarsi a vocabolari prefissati, ma verranno invitati ad allineare, con strumenti semplici da usare, la propria tassonomia a quella condivisa, facendo uso di ontologie semplici ma potenti, come SKOS (<https://www.w3.org/2004/02/skos/>) o FOAF (<http://xmlns.com/foaf/spec/>).

2. SVILUPPO E CARATTERISTICHE

Zotero2map consente di visualizzare e ricercare dati bibliografici all'interno di una mappa online sviluppata con la libreria Maplibre GL JS (<https://maplibre.org/>), utilizzando parole chiave (tag), che definiscono i contenuti che sono archiviati all'interno delle librerie di Zotero. Dunque semplicemente cliccando sulla mappa è possibile risalire ai titoli dei contributi, dando la

```

{
  "type": "FeatureCollection",
  "name": "map",
  "features": [
    {
      "type": "Feature",
      "properties": {
        "name": "Chaonia",
        "altLabel": "Kaonia, Caonia",
        "broader": "Epirus"
      },
      "geometry": {
        "type": "Polygon",
        "coordinates": [
          [
            [19.572850963589893, 40.176334291240465],
            [19.68014489397072, 40.134306906071043],
            [19.88838585739381, 40.67070390342011],
            [19.912168018591871, 39.9217228894243],
            [19.994043256064288, 39.703600408414957],
            [20.042370792013909, 39.679635215754409],
            [20.081304662424344, 39.666673385832297],
            [20.102614745159883, 39.649640745995868],
            [20.10979108365076, 39.67411993637756],
            [20.27294309765808, 39.608152232624],
            [20.405122106741572, 39.8080616944640937],
            [20.431202407195883, 39.927894066696354],
            [20.340892092105548, 40.081636970322421],
            [20.091290593271822, 40.329207134616183],
            [19.892050406630773, 40.343161420197589],
            [19.6197905808646, 40.204077147014505],
            [19.572850963589893, 40.176334291240465]
          ]
        ]
      }
    }
  ]
}

```

Fig. 3 – Struttura GeoJSON della ontologia.

possibilità ai singoli utenti/ricercatori di essere sempre aggiornati sul dibattito scientifico e sulle pubblicazioni riguardanti una certa zona. Lo strumento ITC è stato sviluppato con tecnologie open source e software liberi ed è utilizzabile da una platea virtualmente universale, implementabile da altri centri di ricerca nazionali e internazionali. In dettaglio, il workflow di sviluppo della applicazione ha previsto quattro step: 1) la creazione della ontologia; 2) il reperimento dei dati bibliografici attraverso l'utilizzo dell'API di Zotero; 3) la programmazione JavaScript per l'incrocio dei dati di Zotero con l'ontologia; 4) la costruzione dell'interfaccia grafica e del motore di ricerca interno.

Per quanto riguarda il primo punto, i ricercatori sono chiamati a collaborare alla creazione di un vocabolario/ontologia condiviso, disponibile alla pagina del progetto. In pratica ai ricercatori viene chiesto di stabilire o confermare l'identità tra i concetti: ad esempio, di Caonia (italiano), Kaonia (albanese), *Xaonia* (greco) e Chaonia (inglese); oppure l'appartenenza di un concetto ad un altro: ad es., Caonia è parte dell' (è compreso nell') Epiro, oppure "fortificazioni" è un sottoinsieme di "architettura". La creazione di una mappatura molto chiara tra le varie parole chiave, superando le barriere linguistiche e di scala della ricerca, permette di collegare record bibliografici in maniera estremamente puntuale e concettualmente solida, fornendo uno strumento di ricerca di estremo interesse, capace di allargare in maniera importante gli orizzonti della ricerca. Questa ontologia è resa mediante un file GeoJSON (Fig. 3), che prevede: le properties name (nome del toponimo), altLabel (nome alternativo del medesimo toponimo), broader (appartenenza

ad un insieme) e la sezione riguardante la tipologia geometrica (punto, linea, poligono) con le rispettive coordinate.

Per quanto riguarda la connessione con le banche dati degli utenti, l'applicazione utilizza Zotero API client (<https://www.npmjs.com/package/zotero-api-client>) ovvero un client sviluppato in JavaScript che consente di leggere tutti gli elementi della libreria. Il recupero dei dati bibliografici tramite API (Application Programming Interface) che lo stesso Zotero offre consente di non duplicare il database bibliografico dei singoli ricercatori e di essere sempre aggiornati sui cambiamenti. Attualmente (marzo 2024), la versione predefinita e consigliata dell'API Zotero è la v.3; normalmente le richieste dati alla libreria si riferiscono ad un "gruppo id" o a un "user id" a cui segue l'indicazione di "/item" per estrarre tutti i dati di una libreria, ad esclusione degli elementi cestinati. Il programma a questo punto effettua un incrocio dei dati tra la bibliografia proveniente da Zotero, strutturati in formato JSON e il file dell'ontologia, seguendo questa logica:

- confronta la proprietà "tag" della bibliografia con le voci dell'ontologia;
- se un "tag" ha una corrispondenza con una o più voci dell'ontologia, il codice aggiungerà alla risorsa bibliografica una proprietà denominata "match" contenente un array con i valori dei tag che hanno corrispondenza, questo perché una risorsa bibliografica potrebbe essere associata ad esempio con più luoghi;
- il codice raggruppa tutti gli elementi della bibliografia per "matched tag" e crea un file GeoJSON a cui aggiunge la proprietà "biblio" che contiene un array dove sono indicati tutti i record bibliografici corrispondenti con l'indicazione del titolo, dei tag, della data di edizione e dell'autore.

Rendiamo questa procedura più chiara con un esempio concreto (Fig. 4). L'elemento bibliografico della collezione "Epirus" di Zotero, *In the footsteps of Aeneas: Excavations at Butrint, Albania 1991-2* di Arafat Morgan del 1995, ha come tag "Butrint", "Chaonia", "Epirus", "Greek Archaeology" e "Kestrine". Quindi nella nostra ontologia avrà una corrispondenza con tre voci: Butrint, Chaonia e Kestrine (altlabel di Cestrine). Il programma, dunque, associa alla risorsa bibliografica questi tre valori all'interno della nuova proprietà "match" che non è nativamente presente nella risposta alla chiamata API di Zotero, ma viene aggiunta *ad hoc* all'interno del flusso del codice JavaScript. Una volta completata la procedura di valutazione di matching tra la bibliografia e la ontologia e dunque i record bibliografici hanno impostati i valori di match, il programma raggruppa i libri che hanno in comune uno stesso valore (ad es. Chaonia/Chaonia, Kestrine/Kestrine, etc.) e crea una nuova proprietà "biblio" in un file GeoJSON completamente nuovo, che da un lato riprende i valori dell'ontologia ma aggiunge un array bibliografico con tutti i record bibliografici che hanno in comune lo stesso valore di tag. Quindi proseguendo nel nostro esempio, nel file GeoJSON finale si troverà un

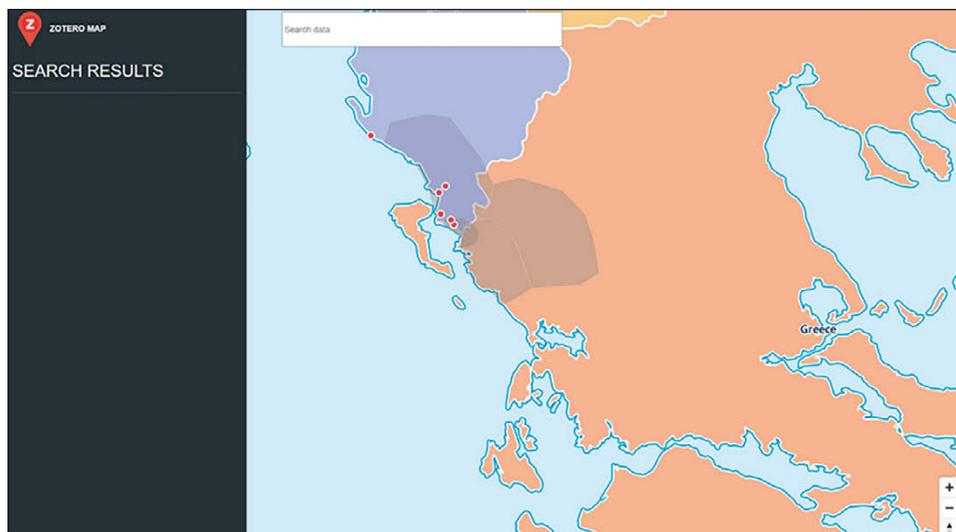


Fig. 4 – Interfaccia webGIS.

elemento con “properties name” Cestrine, Altlabell Kestrine, Broader Chaonia, Trespotia e biblio, il quale è un array di oggetti contenenti i valori della risorsa bibliografica con properties title *In the footsteps of Aeneas: Excavations at Butrint, Albania 1991-2*, tag un array contenente Butrinto, Chaonia, Epirus, Greek Archaeology e Kestrine e infine author Arafat, Morgan 1995.

Questa routine serve per collegare ad un singolo punto o area tutti i libri che riguardano quel luogo. Il file GeoJSON creato dal codice viene utilizzato come base per la realizzazione della mappa, per la quale è stata usata la libreria TypeScript MapLibre GL JS. Infatti, l’interfaccia utente, anch’essa sviluppata in JavaScript, consiste in una mappa (Fig. 5) che consente di visualizzare i luoghi (areali o puntuali) definiti nella ontologia, a cui è associato un array di elementi bibliografici che hanno in comune la medesima localizzazione geografica. L’utente cliccando su un’area o un punto riceverà sulla colonna di sinistra i risultati pertinenti alla sua azione.

Nello specifico il codice di sviluppo della mappa prevede diversi passaggi tra cui:

- per ogni elemento del file GeoJSON crea un layer e controlla il tipo di geometria, se è un punto o un poligono e assegna le diverse caratteristiche grafiche (colore, pattern, etc.);
- aggiunge il layer sulla mappa;
- crea la sidebar ubicata sulla sinistra e per ogni elemento bibliografico riporta il titolo del volume, i tag a cui è associato e il nome dell’autore.

```
{
  "type": "Feature",
  "properties": {
    "name": "Cestrine",
    "altlabel": "Kestrine",
    "broader": "Chaonia, Thesprotia",
    "biblio": [
      {
        "key": "71JWJ2E4",
        "title": "In the footsteps of Aeneas: Excavations at Butrint, Albania 1991-2",
        "tag": [
          "Butrint",
          "Chaonia",
          "Epirus",
          "Greek Archaeology",
          "Kestrine"
        ],
        "author_date": "Arafat, Morgan. 1995"
      },
      {
        "key": "IF2EFS4K",
        "title": "Çuka e Aftoit: një kontribut i ri",
        "tag": [
          "Cestrine",
          "Chaonia",
          "Epirus",
          "Fortifications",
          "Hellenistic Archaeology",
          "Çuka e Aftoit"
        ],
        "author_date": "Bogdani. 2009"
      }
    ]
  }
},
```

Fig. 5 – Struttura GeoJSON del file prodotto dalla routine di corrispondenza tra elementi bibliografici e ontologia.

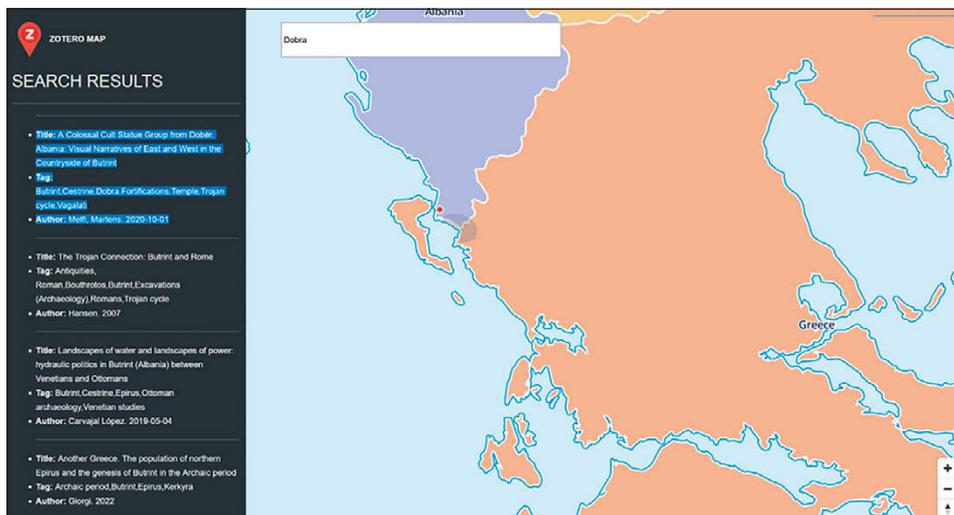


Fig. 6 – Esempio di restituzione dati.

Oltre all'azione diretta sulla mappa, l'utente può eseguire le proprie ricerche attraverso un form posizionato in alto a sinistra, nel quale può filtrare i dati per località, titolo del volume, autore o argomento. In questo caso l'utente vedrà direttamente sulla mappa il risultato delle proprie query (Fig. 6).

3. CONCLUSIONI

In conclusione, Zotero2map amplifica le funzioni di Zotero, fornendo importanti servizi di aggiornamento e conciliazione bibliografica per i singoli studiosi, senza introdurre un aggravio di lavoro e senza cambiare il flusso di lavoro che già mettono in atto. L'utente finale, dunque, semplicemente aprendo un browser web, potrà accedere alla visualizzazione della mappa e potrà effettuare delle ricerche sui termini di interesse. Le potenzialità innovative del sistema rispetto alle opzioni attualmente in uso stanno dunque nella facilità di utilizzo da parte degli utenti, che non devono installare alcunché sui propri dispositivi, nella realizzazione di un'interfaccia user friendly e cross-browsing, visualizzabile su diverse piattaforme, nella connessione automatica dei dati e nell'elaborazione multi-language dei "tag" e delle keyword, che di fatto rende il sistema fruibile a ricercatori di diverse nazionalità.

Si tratta di uno strumento software, che in questa fase iniziale di implementazione è dedicato all'area epirota, ma che potrà essere personalizzato e implementato per qualsiasi ambito della ricerca scientifica. L'applicazione è costruita interamente con tecnologie open source ed è rilasciata a sua volta con licenza open source (AGPL 3.0, <https://www.gnu.org/licenses/agpl-3.0.en.html>), come gran parte dei progetti a cura del LAD. È liberamente scaricabile dalla pagina GitHub del Laboratorio di Archeologia Digitale al seguente indirizzo: <https://github.com/lab-archeologia-digitale/zotero2map>.

ELEONORA IACOPINI

Sapienza Università di Roma
eleonora.iacopini@uniroma1.it

BIBLIOGRAFIA

- AHMED K.K.M., AL DHUBAIB B.E. 2011, *Zotero: A bibliographic assistant to researcher*, «Journal of Pharmacology & Pharmacotherapeutics», 2, 4, 303-305 (<https://doi.org/10.4103/0976-500X.85940>).
- CARVALE A. 2022, *Le banche dati archeologiche. Articolazione e formalizzazione delle conoscenze*, Firenze, Edizioni All'Insegna del Giglio.
- CARVALE A., PIERGROSSI A. 2012, *Archeologia in rete. Le riviste open access: risorse e prospettive*, «Archeologia e Calcolatori», 23, 187-207 (https://www.archcalc.cnr.it/indice/PDF23/11_Caravale_Piergrossi.pdf).
- CARVALE A., PIERGROSSI A. 2015, *Archaeological open access journals: The case of «Archeologia e Calcolatori»*, in F. GILIGNY, F. DJINJIAN, L. COSTA, P. MOSCATI, S. ROBERT (eds.), *Concepts, Methods and Tools. Proceedings of the 42nd Annual Conference on Computer Applications and Quantitative Methods in Archaeology (Paris 2014)*, Oxford, Archaeopress, 257-263.
- CARVALE A., PIERGROSSI A., ROSSI I. 2021, *Open Data, Open Knowledge, Open Science: The New OpenLab of the Institute for Heritage Science (CNR)*, in J. BOGDANI, R. MONTALBANO, P. ROSATI (eds.), *ArcheoFOSS XIV 2020, Open Software, Hardware, Processes, Data and Formats in Archaeological Research, Proceedings of the 14th International Conference*, Oxford, Archaeopress (<https://www.archaeopress.com/Archaeopress/download/9781803271248>).

PUCKETT J. 2011, *Zotero: A Guide for Librarians, Researchers, and Educators*, Chicago, Association of College and Research Libraries.

TOBIAS M., MANDEL A. 2021, *Literature mapper: A QGIS plugin for georeferencing citations in Zotero*, «Air, Soil and Water Research», 14 (<https://doi.org/10.1177/11786221211009209>).

ABSTRACT

Zotero2map is a web application for sharing bibliographic data from various collaborators who want to share their bibliographic collections, using the Zotero software with the specificity of displaying and querying data also on a geographical map. With this tool the LAD: Laboratory of Digital Archeology at Sapienza (<http://purl.org/lad>) aims to facilitate the exchange of information relating to different historical and archaeological contexts, to speed up their retrieval and ensure constant updating of the data entered into the system. The webApp allows you to view and search for bibliographic data within an online map, using keywords (tags) that define the contents that are archived within the Zotero Libraries, which individual researchers already use for own scientific production. By simply clicking on the map it will therefore be possible to go back to the titles of the contributions, giving the possibility to individual users/researchers to always be updated on the scientific debate and on the publications concerning a certain area. The international vocation of the project presented is given not only by usability and open distribution, but also by the development of the cross-language functionality of keywords, a system that will allow the association of terms related to the same research topic but treated in different languages. The use of shared, multilingual dictionaries will allow not only to break down language barriers, but also to define an articulated taxonomy of relationships between the terms used. Individual researchers therefore do not have to adapt to pre-established vocabularies, but are invited to align their taxonomy with the shared one, making use of simple but powerful ontologies, such as SKOS or FOAF. Currently the software has been used on the specific case of Caonia, Northern Epirus.