

Software libero nei processi di ricerca archeologica: l'applicazione allo studio della sigillata africana in Italia centro-settentrionale

Stefano Costa, Dipartimento di Archeologia e Storia delle Arti, Università degli Studi di Siena

steko@unisi.it 348 9133647

Con l'intenzione di interpretare lo spirito dei workshop ARCHEOFLOSS, questo intervento propone un caso di studio completo relativo all'utilizzo, sviluppo e applicazione di software libero nell'ambito della ricerca svolta per la tesi di laurea specialistica dell'autore.

La ricerca è stata finalizzata a delineare i tratti di complessità della distribuzione geografica e sociale della terra sigillata africana a partire dall'inizio del V secolo in Italia centro-settentrionale, basandosi sui dati editi disponibili.

Viene pertanto illustrato l'intero processo della ricerca, partendo dalla raccolta dei dati e contemporanea definizione del modello, soggetto ad una evoluzione continua per adattarsi alle esigenze di registrazione via via incontrate. Illustrando i motivi delle scelte tecniche, particolare attenzione viene dedicata allo sviluppo di una applicazione in linguaggio Python, basata sul *framework* GeoDjango. La possibilità di estendere l'applicazione tramite le numerose librerie di programmazione (tra cui la nota Matplotlib per la creazione avanzata di grafici) e il riutilizzo di codice precedentemente sviluppato dallo stesso autore hanno infatti permesso la rapida creazione di un prototipo funzionante, utile nella ricerca grazie alla facilità di automazione e di consultazione tramite una semplice interfaccia *web*, accessibile sul server locale.

Ogni informazione inserita è riconducibile alla propria fonte tramite tradizionali riferimenti bibliografici, inseriti in formato BibTeX e formattati tramite la libreria di funzioni Pybliographer.

La predisposizione di GeoDjango per la creazione dinamica di file in formati geografici standard quali KML, GML e GeoJSON ha inoltre facilitato l'interazione con la suite di software OSGeo per la visualizzazione avanzata di cartografia e la produzione di mappe geografiche di alta qualità destinate alla stampa.

Nella fase di analisi dei dati raccolti e della valutazione dei modelli interpretativi proposti si è potuta effettuare a più riprese la lettura di dati opportunamente preparati tramite il linguaggio di analisi R. In questa fase, è risultata di particolare efficacia la possibilità di automazione del processo di estrazione ed analisi dei dati, grazie all'utilizzo concatenato dei linguaggi Python (per l'estrazione) ed R (per l'analisi dei dati e la creazione di grafici riassuntivi di alta qualità): la stesura di opportuni *script*, eseguiti in sequenza, ha garantito la ripetibilità delle operazioni e la possibilità di parametrizzazione per analizzare lo stesso fenomeno al variare di determinate condizioni. In questa fase, è risultata determinante la scelta di un ambiente di lavoro avanzato quale GNU Emacs (versione 23) e della modalità ESS ("Emacs Speaks Statistics") per l'interazione con l'interprete R e la scrittura di codice in entrambi i linguaggi di programmazione.

Infine, la "materializzazione" dell'intero processo di raccolta, modellazione e analisi dei dati sotto forma di codice sorgente ha comportato il vantaggio, se vogliamo collaterale ma comunque di importanza strategica, di poter utilizzare un sistema di controllo versione che rende esplicito il processo evolutivo della ricerca, e in teoria permette di confrontarne i diversi momenti. Come per altri progetti software è stato utilizzato Mercurial, un completo sistema di controllo versione distribuito.

I risultati archeologici di questa ricerca sono in corso di pubblicazione in altra sede, ma verranno comunque illustrati brevemente per completezza.