

## UNA PROPOSTA DI STANDARD PER L'ARCHIVIAZIONE E LA CONDIVISIONE DI DATI STRATIGRAFICI

### 1. IL MATRIX DI HARRIS

La diffusione della pratica di condivisione dei dati archeologici ha coinvolto in modo diversificato le varie componenti della documentazione archeologica, interessando solo raramente in dettaglio l'archivio dello scavo stratigrafico, che al contrario rappresenta una base essenziale di conoscenza. Il cuore della documentazione dello scavo è la sequenza stratigrafica, documentata sotto forma di diagramma stratigrafico e ben nota nella sua forma analogica di ampio formato cartaceo. Sin dal 1990 sono stati disponibili strumenti digitali dedicati per la realizzazione e l'archiviazione del diagramma stratigrafico o matrix di Harris (BASP, Stratify, ArchEd), tutti basati sul riconoscimento del matrix come “grafo direzionato aciclico” (Directed Acyclic Graph o DAG), un genere particolare di grafo, con proprietà matematicamente note e di comune implementazione negli strumenti informatici specializzati (GASNER, NORTH 2000). Questa identificazione, corretta da un punto di vista strettamente teorico, rimane tuttavia problematica nell'effettiva applicazione, per via delle “eccezioni” non facilmente riducibili come il riconoscimento della contemporaneità di due unità stratigrafiche.

La descrizione del matrix di Harris come oggetto digitale non ha avuto vita facile, in un settore disciplinare molto ancorato a convenzioni grafiche di prevalente idiosincrasia (AMADORI *et al.* 2017): si adatterà quindi una convenzione opposta, dove il matrix di Harris è una base di dati che può avere molteplici rappresentazioni valide, ognuna legata a domande specifiche: quali sono le US più recenti di una data US? Quali sono le US più vantaggiose per effettuare una datazione radiometrica? Si propone quindi di abbandonare una concezione statica e descrittiva del matrix per considerarlo una base conoscitiva su cui elaborare la conoscenza cronologica dei contesti archeologici indagati, integrando le informazioni di carattere assoluto risultanti dallo studio dei reperti o dalle datazioni dendrocronologiche e radiometriche.

Questo è l'approccio che ha caratterizzato alcuni primi esperimenti sviluppati nel 2008 con il software Graphviz (<https://gitlab.com/iosa/harris-matrix>), poi adottato anche in altri strumenti dedicati alla pratica di scavo archeologico come pyArchInit (MANDOLESI, COCCA 2013), e in particolare dal software “hm” (DYE, BUCK 2015) che ha introdotto la principale innovazione: l'archiviazione dei dati in una struttura relazionale, composta da semplici tabelle CSV (Comma Separated Values), che vengono elaborate dal software per produrre le rappresentazioni del matrix come diagramma. La produzione della forma grafica finale rimane demandata sempre al software Graphviz,

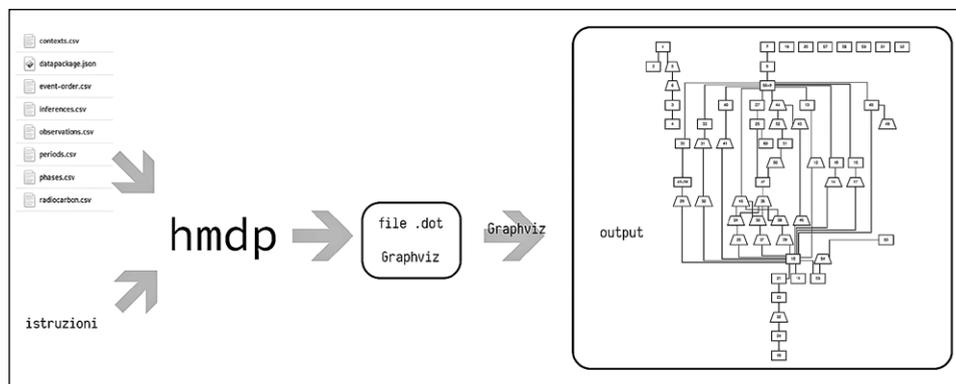


Fig. 1 – Harris Matrix Data Package: schema concettuale della procedura di elaborazione dei dati ed esempio di risultato grafico, basato sui dati di Ducci *et al.* 2012.

senza tuttavia utilizzare il formato dati Graphviz/DOT per l’archiviazione dei dati veri e propri.

La struttura relazionale è basata sulla pratica consolidata di altri software e prevede due tabelle obbligatorie: la prima con elenco e descrizione delle unità stratigrafiche (*contexts*), la seconda con elenco delle relazioni stratigrafiche osservate (*observations*). Significativamente, questa seconda tabella non contiene indicazioni né sulle relazioni del tipo “uguale a” né su quelle del tipo “contemporaneo a”, che vengono inserite in una tabella a parte (*inferences*). Infatti, questa categoria di rapporti stratigrafici, che è naturalmente parte della definizione intrinseca comunemente adottata nella pratica archeologica, rappresenta una difformità rispetto all’identificazione tra matrix di Harris e DAG, in cui non è contemplata la possibilità di un rapporto non direzionale tra due nodi. Generalmente la limitazione, che è ben nota anche nella letteratura archeologica, viene affrontata in due modi: da un punto di vista concettuale, accorpando in un unico elemento i due o più elementi ritenuti uguali o coevi; alternativamente, nelle soluzioni più strettamente grafiche, tramite tratti orizzontali. La soluzione più adatta al singolo caso può invece essere scelta, senza essere fissata in una forma definitiva, e il matrix, da oggetto finito che deve essere interpretato in modo impressionistico, diviene uno strumento che permette di costruire percorsi interpretativi rigorosi, non limitati alla descrizione delle relazioni stratigrafiche ma estesi a modelli cronologici complessi (DEAN 1978).

## 2. LO “HARRIS MATRIX DATA PACKAGE”

Lo standard per i dati stratigrafici proposto include ulteriori tabelle opzionali per i periodi (*periods*) e le fasi (*phases*) che completano il classico

quadro di interpretazione della sequenza di scavo. Si aggiungono tuttavia anche due tabelle dedicate ai dati di cronologia assoluta (*events*) collegati alle singole unità stratigrafiche, e all'ordine tra eventi collegati alla stessa unità stratigrafica (*event-order*). Se in una tradizionale rappresentazione come in un file di database relazionale (es. SQLite) la coerenza interna dei dati e la descrizione della struttura sono contenuti all'interno del file stesso, in questo caso le tabelle CSV vengono affiancate da un file data descriptor JSON, che segue lo standard del Tabular Data Package (<https://frictionlessdata.io/specs/tabular-data-package/>). La validazione formale dei dati sfrutta le librerie compatibili con lo standard generico Data Package, disponibili per vari linguaggi di programmazione (R, JavaScript, Python), integrando anche la validazione stratigrafica vera e propria (assenza di cicli chiusi, indicazione di unità stratigrafiche orfane e vedove).

L'insieme delle tabelle CSV e del file datapackage.json è un "Harris Matrix Data Package" che può essere facilmente salvato, compresso in un file ZIP, condiviso in repository open data o inviato tramite posta elettronica. La semplicità del formato dati permette di archiviare i dati stratigrafici provenienti anche da grandi interventi di scavo con dimensioni estremamente contenute.

Lo strumento specifico compatibile con questa proposta di standard, sviluppato in Python e rilasciato con licenza GNU GPLv3, è basato sulla libreria NetworkX (HAGBERG *et al.* 2008) ed è utilizzabile dalla linea di comando (<https://gitlab.com/steko/harris-matrix-data-package>) come eseguibile denominato "hmdp". Fornendo come argomento al programma il file data descriptor, viene caricato l'intero data package; in base al comando indicato il programma esegue diverse operazioni, tra cui la validazione formale, la validazione stratigrafica, la preparazione di una rappresentazione sotto forma di diagramma in formato DOT/Graphviz, che verrà poi eventualmente trasformato in immagine. Il software nella fase di sviluppo attuale può essere considerato un prototipo, la cui naturale evoluzione è rappresentata dalla creazione di un'applicazione web dedicata.

STEFANO COSTA\*

Soprintendenza Archeologia, Belle arti e Paesaggio per la Città  
Metropolitana di Genova e le province di Imperia, La Spezia e Savona  
[stefano.costa@beniculturali.it](mailto:stefano.costa@beniculturali.it)

## BIBLIOGRAFIA

AMADORI L., CAPPONI M., DE ROSSI G., FOSCHI A. 2017, *Esercizi di matrix*, Monte Compatri, Espera.

\* <https://orcid.org/0000-0003-1124-3174>.

- DEAN J.S. 1978, *Independent dating in archaeological analysis*, in M.B. SCHIFFER (ed.), *Advances in Archaeological Method and Theory*, San Diego, Academic Press, 223-255 (<https://doi.org/10.1016/B978-0-12-003101-6.50013-5>).
- DUCCI S., ANICHINI F., BERTELLI E., COSTANTINI A., DEL FREO A., GIORGIO M., GIANNOTTI S., MENCHINI M. 2012, *Via Galluppi – PiGAL 09 – Assistenza e scavo preventivo*, «MOD» (doi:10.4456/MAPPA.2012.26).
- DYE T.S., BUCK C.E. 2015, *Archaeological sequence diagrams and Bayesian chronological models*, «Journal of Archaeological Science», 63, Suppl. C, 84-93 (<https://doi.org/10.1016/j.jas.2015.08.008>).
- GANSNER E.R., NORTH S.C. 2000, *An open graph visualization system and its applications to software engineering*, «Software – Practice and Experience», 30, 11, 1203-1233.
- HAGBERG A.A., SCHULT D.A., SWART P.J. 2008, *Exploring network structure, dynamics, and function using NetworkX*, in G. VAROQUAUX, T. VAUGHT, J. MILLMAN (eds.), *Proceedings of the 7<sup>th</sup> Python in Science Conference – SciPy (Pasadena, CA 2008)*, Pasadena, 11-15.
- HERZOG I., SCOLLAR I. 1991, *A new graph theoretic oriented program for Harris Matrix analysis*, in S.P.Q. RAHTZ, K. LOCKYEAR (eds.), *CAA90. Computer Applications and Quantitative Methods in Archaeology*, Oxford, Tempus Reparatum, 53-59.
- MANDOLESI L., COCCA E. 2013, *PyArchInit: gli sviluppi dopo ArcheoFoss 2009*, in M. SERLORENZI (ed.), *Atti del VII Workshop Open Source, Free Software e Open Format nei processi di ricerca archeologica (Roma 2012)*, «Archeologia e Calcolatori», Suppl. 4, 128-138.

## ABSTRACT

Archaeological stratigraphy-datasets are widely represented and visualised by the Harris Matrix. However, even digital Harris matrices retain most of the limitations their paper originals experienced and they are still in need of a standardised format for automatic recording. This paper presents an example, based on the well-known CSV open format, with a Data Package descriptor. The Harris Matrix Data Package standard is the result of two independent software implementations, one in Common Lisp by T.S. Dye and one in Python. The proposed format is not limited to describing the stratigraphy for the sole purpose of generating a diagram to look at, but it can be used to model stratigraphy as a Directed Acyclic Graph and to guide the interpretative process and integrate the results with established procedures for Bayesian modelling.