

ISTITUTO ITALIANO DI PREISTORIA E PROTOSTORIA

RIVISTA DI SCIENZE PREISTORICHE

numero speciale

Top of the Pop or Oldies but Goldies:
Approaches to Prehistoric Data from
the Central Mediterranean

$$f(SU) = \sum_{SU=1}^n \left(\frac{aDNA^{87sr/86sr}}{Pots} \right)^{14C}$$

C o u n t i n g R o u n D a t a

The volume collects the peer-reviewed revised versions of the papers presented at the *Vecchi o nuovi, purché condivisi e analizzabili: approcci ai dati della preistoria del Mediterraneo Centrale / Top of the Pop or Oldies but Goldies: Approaches to Prehistoric Data from the Central Mediterranean*, held in Rome on 23-24 October 2023.

Scientific and Organizing Committee

Andrea Di Renzoni, Maja Gori & Marc Vander Linden

Istituto Italiano di Preistoria e Protostoria

Oldies but Goldies

Top of the Pop or Oldies but Goldies: Approaches to Prehistoric Data from the Central Mediterranean

edited by Andrea Di Renzoni, Maja Gori, Marc Vander Linden

The volume collects the peer-reviewed revised versions of the papers presented at the *Vecchi o nuovi, purché condivisi e analizzabili: approcci ai dati della preistoria del Mediterraneo Centrale / Top of the Pop or Oldies but Goldies: Approaches to Prehistoric Data from the Central Mediterranean*, held in Rome on 23-24 October 2023.

Scientific and Organizing Committee

Andrea Di Renzoni, Maja Gori & Marc Vander Linden



ISBN 978-88-6045-108-8

© Istituto Italiano di Preistoria e Protostoria, 2025
c/o Museo Archeologico Nazionale di Firenze
Via della Pergola, 65, 50121 Firenze
www.iipp.it, email: iipp@iipp.it – segreteria@iipp.it

Con il presente volume l'Istituto Italiano di Preistoria e Protostoria amplia la serie dei "Numeri Speciali" della Rivista di Scienze Preistoriche, serie che è stata avviata nel 2020 con l'intento di pubblicare volumi dedicati all'approfondimento di tematiche specifiche, e che finora ha accolto solo testi relativi ad iniziative scientifiche promosse dall'Istituto.

A partire da questo volume, infatti, l'Istituto offre ai propri soci la possibilità di pubblicare come numero speciale della prestigiosa Rivista di Scienze Preistoriche anche testi generati da iniziative che non sono nate sotto l'egida dell'Istituto, ma che meritano di essere pubblicati per la loro qualità scientifica. L'intento è quello di arricchire ulteriormente i servizi che l'IIPP propone ai soci, valorizzando l'attività scientifica degli studiosi di Preistoria e Protostoria e mettendo loro a disposizione una sede editoriale rispondente agli attuali criteri di valutazione accademica.

I nuovi volumi potranno essere sia unicamente digitali Open Access, come il primo che presentiamo, sia tradizionalmente a stampa, sia con entrambe le modalità editoriali, come il secondo in uscita. In ogni caso è previsto che vengano approvati dal Comitato di Redazione di RSP, che non gravino sui bilanci dell'Istituto e che siano curati da guest editors. Essi osservano le norme redazionali della Rivista e seguono la rigorosa procedura di peer-review di tutti i contributi, in linea con i numeri speciali già usciti. Solo l'adozione di un colore differente per le scritte di copertina distingue le iniziative che fanno capo direttamente all'Istituto (scritte verdi su fondo grigio) da quelle proposte dai soci (scritte blu su fondo grigio).

Inaugura la nuova formula il volume *Top of the Pop or Oldies but Goldies: Approaches to Prehistoric Data from the Central Mediterranean*, curato da Andrea Di Renzoni, Maja Gori e Marc Vander Linden. Il volume, in inglese, è basato sul workshop svoltosi a Roma nel 2023 nel quadro di un accordo tra la Royal Society e il Consiglio Nazionale delle Ricerche, e raccoglie contributi sul tema dell'utilizzo e riutilizzo di diversi tipi di dati archeologici, sia relativi alla cultura materiale, sia provenienti dalle scienze dure. L'opera si propone come strumento di riflessione e orientamento per gli archeologi preistorici, anche in accordo con le politiche comunitarie che promuovono l'uso e il riuso dei dati scientifici secondo i principi "FAIR" (Findable, Accessible, Interoperable, Reusable), soprattutto se prodotti da ricerche finanziate con fondi pubblici.

In sostanziale contemporaneità con questo primo contributo viene pubblicato in questa serie anche il volume *HERDS – Animal Husbandry and its Economic Role in the Development of Central Mediterranean Protohistoric Societies*, al quale si rimanda. Il volume, anch'esso in inglese, è il risultato di una articolata ricerca finanziata dal MUR come Progetto di Rilevante Interesse Nazionale; è curato da Andrea Cardarelli, Giulia Recchia, Antonio Curci, Beatrice Demarchi, Claudia Minniti, Marco Pacciarelli e presenta i risultati di una ricerca che integrando metodologie già ampiamente consolidate, quali quelle dell'archeozoologia, con quelle isotopiche, ZooMS e DNA antico rappresenta un approccio innovativo allo studio dell'economia di allevamento nell'Italia protostorica.

Andrea Cardarelli, Maria Bernabò Brea, Maja Gori

With this volume, the Italian Institute of Prehistory and Protohistory expands the “Special Issues” series of the Rivista di Scienze Preistoriche, a series launched in 2020 with the aim of publishing volumes dedicated to in-depth studies of specific topics, and which until now has included only works related to scientific initiatives promoted by the Institute.

From this point forward, the Institute extends to its members the opportunity to publish, within this prestigious series, works originating from initiatives not conducted under its direct aegis, yet distinguished by their high scientific quality. This development reflects the Institute’s commitment to broadening the services offered to its members, supporting scholarly activity in Prehistory and Protohistory, and providing an editorial venue aligned with current standards of academic evaluation.

The new volumes may take the form of fully digital Open Access publications, as in the present case, traditional printed volumes, or a combination of both formats, as exemplified by the forthcoming second issue. In all instances, submissions are subject to approval by the Editorial Board of the Rivista di Scienze Preistoriche, must not entail financial costs for the Institute, and are to be curated by guest editors. All contributions adhere to the journal’s editorial standards and undergo a rigorous peer-review process, in continuity with previously published special issues. A visual distinction is maintained between volumes promoted directly by the Institute and those proposed by its members: the former are identified by green lettering on a grey cover; the latter by blue lettering on a grey cover.

The present volume, Top of the Pop or Oldies but Goldies: Approaches to Prehistoric Data from the Central Mediterranean, edited by Andrea Di Renzoni, Maja Gori, and Marc Vander Linden, inaugurates this new format. Written in English, it is based on a workshop held in Rome in 2023 within the framework of a bilateral agreement between the Royal Society and the Italian National Research Council. The contributions collected here address the use and reuse of diverse categories of archaeological data, ranging from material culture to datasets derived from the natural sciences. The volume aims to serve as a tool for reflection and guidance for prehistoric archaeologists, also in line with European policies that promote the use and reuse of scientific data according to the “FAIR” principles (Findable, Accessible, Interoperable, Reusable), particularly for data produced through publicly funded research.

Published concurrently within the same series is the volume HERDS – Animal Husbandry and its Economic Role in the Development of Central Mediterranean Protohistoric Societies, also in English, this work presents the results of a major research project funded by the Italian Ministry of University and Research (MUR) as a Project of Relevant National Interest (PRIN). Edited by Andrea Cardarelli, Giulia Recchia, Antonio Curci, Beatrice Demarchi, Claudia Minniti, and Marco Pacciarelli, the volume offers an innovative perspective on protohistoric animal husbandry in Italy by integrating established approaches, such as archaeozoology, with advanced analytical techniques including isotopic analysis, ZooMS, and ancient DNA.

Andrea Cardarelli, Maria Bernabò Brea, Maja Gori

INTRODUCTION

This collection of papers, focusing on different approaches to prehistoric data, stems from the international workshop *Vecchi o nuovi, purché condivisi e analizzabili: approcci ai dati della preistoria del Mediterraneo Centrale / Top of the Pop or Oldies but Goldies: Approaches to Prehistoric Data from the Central Mediterranean*. The workshop was held in Rome on 23-24 October 2023 and focused specifically on the use of different types of data (from legacy data to hard-science data) and was structured around the following themes: (1) data processing and sharing; (2) analysis and treatment of radiometric dating data; (3) collection and analysis of isotopic data; (4) DNA data in archaeology; and (5) quantitative treatment of data related to material culture. Each theme was addressed by different experts (two per theme), thereby providing an important opportunity for in-depth reflection on quantitative approaches to prehistoric archaeology. We are grateful to all the colleagues who participated in the event and contributed to the fruitful discussions, and in particular to those who authored the papers included in this issue. We also thank the external reviewers who kindly took on this task: Alfredo Cortell-Nicolau, Alejandra Galmés Alba, Gabriel García Atienzar, and Margaux Lucienne Depaermentier.

Top of the Pop or Oldies but Goldies was organised within the framework of a two-year bilateral agreement between the Royal Society and the Italian National Research Council entitled “*Modelling and Measuring Interaction Networks in the Adriatic during Late Prehistory*” (International Exchanges 2021). The CNR-ISPC research unit was composed of Andrea Di Renzoni (PI) and Maja Gori, founding members of the MAD-Lab (Modelling Archaeological Data Laboratory). The IMSET - Bournemouth University research unit was composed by Marc Vander Linden (PI), Fiona Coward and Philip Riris.

$$f(SU) = \sum_{SU=1}^n \left(\frac{aDNA^{87sr/86sr}}{Pots} \right)^{14C}$$

C o u n t i n g R o u n d a t a

23rd-24th OCT 2023

Fig. 1 – The conference logo. The formula plays with the language of mathematics to summarize the core idea of quantitative archaeology: each context is a complex function, built through the integration of different proxies (aDNA, isotopic ratios, radiometric dating, material culture, etc.), whose combination produces coherent and comprehensive interpretations..

Il logo del convegno. La formula gioca con il linguaggio della matematica per riassumere l'idea di fondo dell'archeologia quantitativa: ogni contesto è una funzione complessa, costruita dall'integrazione di diversi proxy (aDNA, rapporti isotopici, datazioni radiometriche e cultura materiale etc.), la cui combinazione produce letture complessive e coerenti.



Andrea Di Renzoni ^(1, *), Maja Gori ⁽²⁾, Marc Vander Linden ⁽³⁾

Top of the pop or oldies but goldies: approaches to prehistoric data from Central Mediterranean

^(1*) National Research Council of Italy - Institute of Heritage Science (CNR-ISPC), Area della Ricerca di Roma 1 Strada Provinciale 35d, 9, 00010 Montelibretti (Rome). E-mail: andrea.direnzoni@cnr.it, ORCID 0000-0003-3857-0458

^(2*) University of Trento - Department of Humanities, Via Tommaso Gar, 14, 38122 Trento. E-mail: maja.gori@unitn.it, ORCID 0000-0002-8106-4801

^(3*) Department of Archaeology and Anthropology, Bournemouth University, Talbot Campus, Fern Barrow, Poole, Dorset BH12 5BB, UK. E-mail: mvanderlinden@bournemouth.ac.uk, ORCID: 0000-0002-0120-775

^(*) Corresponding author

Keywords:

archaeological data, data reuse, quantitative methods, data aggregation.

Parole chiave:

dati archeologici, riuso dei dati, metodi quantitativi, aggregazione di dati.

ABSTRACT: TOP OF THE POP OR OLDIES BUT GOLDIES: APPROACHES TO PREHISTORIC DATA FROM CENTRAL MEDITERRANEAN. This introductory paper addresses the epistemological and methodological challenges of reusing archaeological data, particularly within quantitative frameworks. The authors argue that archaeological data do not exist as neutral, pre-interpretative entities, but are produced through stratified analytical processes involving multiple interpretative layers. Key concepts discussed include informational levels, vertical and horizontal aggregation, and the distinction between data production, reconfiguration, and reuse. The paper emphasizes that aggregation is a central analytical practice that requires explicit documentation and archaeological expertise. Quantitative methods, far from neutralising interpretative issues, tend to amplify the consequences of data heterogeneity and aggregation choices. The contributions to this special issue explore these themes through case studies on radiocarbon dating, isotopic analysis, network approaches, and data sharing practices, converging on the conclusion that the increasing scale and complexity of archaeological research intensify the need for expert knowledge, transparency, and critical curation of legacy data.

RIASSUNTO: VECCHI O NUOVI, PURCHÉ CONDIVISI E ANALIZZABILI: APPROCCI AI DATI DELLA PREISTORIA DEL MEDITERRANEO CENTRALE. Questo contributo introduttivo affronta le sfide epistemologiche e metodologiche legate al riuso dei dati archeologici, in particolare nell'ambito di approcci quantitativi. Gli autori sostengono che i dati archeologici non esistono come entità neutre e pre-interpretative, ma sono il risultato di processi analitici stratificati che coinvolgono molteplici livelli interpretativi. Vengono discussi concetti chiave come i livelli informativi, l'aggregazione verticale e orizzontale, e la distinzione tra produzione, riconfigurazione e riuso dei dati. L'aggregazione è presentata come una pratica analitica centrale che richiede documentazione esplicita e competenza archeologica. I metodi quantitativi, lungi dal neutralizzare le questioni interpretative, tendono ad amplificare le conseguenze dell'eterogeneità dei dati e delle scelte di aggregazione. I contributi di questo numero speciale esplorano questi temi attraverso casi studio su datazioni radiocarboniche, analisi isotopiche, approcci di rete e pratiche di condivisione dei dati, convergendo sulla conclusione che la crescente scala e complessità della ricerca archeologica intensificano la necessità di conoscenza esperta, trasparenza e curatela critica dei dati legacy.

WHY TALK (AGAIN) ABOUT ARCHAEOLOGICAL DATA?

Archaeology is a strongly data-driven discipline, arguably even more so than other historical sciences, a characteristic that has become increasingly noticeable since the processualist turn (e.g. Binford 1962; Clarke 1968, 1973) and has expanded dramatically with the so-called “third science revolution” (Kristiansen 2014). Archaeological research is fundamentally grounded in the collection, recording, and analysis of material evidence, and much of its interpretative potential depends on the quality, structure, and accessibility of the data produced. At the same time, archaeological data are inherently heterogeneous, context-dependent, and profoundly shaped by the methodological and interpretative practices through which they are generated. These characteristics make archaeology a particularly challenging and productive field for critical reflection on the creation, reuse, and analysis of data.

In recent years, issues of data reuse and management have gained increasing attention across the sciences, and archaeology is no exception. Large quantities of legacy data, produced through decades or even centuries of excavation, survey, documentation, and analytical work, are now being revisited, digitised, and aggregated into new datasets. Archives, databases, and digital repositories have become central infrastructures for archaeological research, raising fundamental questions about how older data can be mobilised, reinterpreted, and integrated within new analytical and theoretical frameworks.

The title of this special issue, *Oldies but Goldies*, evokes this tension. It derives from the workshop “*Vecchi o nuovi, purché condivisi e analizzabili: approcci ai dati della preistoria del Mediterraneo Centrale*” (*Top of the pop or oldies but goldies: approaches to prehistoric data from the Central Mediterranean*), held in Rome on 23–24 October 2023. The workshop was organized within the framework of a 2-year bilateral agreement between the Royal Society and the Italian National Research Council entitled “*Modelling and measuring interaction networks in the Adriatic during Late Prehistory.*”, led by the present authors. The initiative emerged from the shared

need of the two research units to explore the potential for integrating datasets highly heterogeneous in their nature, structure, and modes of definition, and prompted sustained critical reflection on the epistemological and methodological challenges associated with their use.

The value of archaeological data does not depend on their age alone (for a discussion of the concept of legacy data, see Di Renzoni in press), but on the conditions under which they were produced, documented, and subsequently reused. “*Old*” data are not inherently obsolete, nor are they automatically suitable for contemporary forms of analysis. Rather, their analytical potential emerges only through critical engagement with the epistemic, methodological, and practical contexts of their production, aggregation, and interpretation. From this perspective, the central challenge is not simply to replace older datasets with newly generated ones, but to understand the conditions under which legacy data can be rendered analytically meaningful and productive once again. This issue is particularly pressing in the context of quantitative approaches. Statistical analyses, modelling techniques, and network-based methods promise new insights into past societies, but they also tend to amplify the consequences of data heterogeneity, uncertainty, and aggregation choices. Quantitative methods do not neutralise the interpretative nature of archaeological data; rather, they make explicit, and sometimes exacerbate, the assumptions embedded in their structure.

The contributions gathered in this volume address these challenges from a shared perspective. Focusing on prehistoric data, this collection of papers explores how legacy-based datasets can be critically reused through quantitative and comparative approaches. The central Mediterranean offers a particularly rich laboratory for this discussion, because it is characterised by a long-standing research traditions and diverse practices, as well as by the presence of a substantial body of published and unpublished data. By examining concrete case studies alongside broader methodological reflections, this volume aims to contribute to an ongoing debate on how archaeological “*Oldies*” can, under the right conditions, still become “*Goldies*”.

DO ARCHAEOLOGICAL DATA REALLY EXIST?

The apparently straightforward notion of “archaeological data” becomes problematic as soon as it is examined closely. Unlike many other scientific disciplines, archaeology does not deal with data that are directly observable or reproducible. What we usually call archaeological data are, rather, the result of a stratified analytical and interpretative process. In a recent work on legacy data, it has been argued that properly speaking archaeological data do not exist as such (Di Renzoni in press). What exists is a set of material traces that only become data through archaeological practice. In this perspective, the so-called “raw data” are a pre-discursive entity: they exist only before archaeology takes place. As for many other disciplines, the archaeological record, in its material wholeness, is transformed into data through its creation, in this case involving excavation, recording, classification, and interpretation. If we take as example a given archaeological context, to illustrate this process, we can argue that prior to excavation, this context exists as a complex material configuration. Through excavation, it is analytically decomposed into stratigraphic units, which are commonly treated as basic archaeological data. However, stratigraphic units should not be understood as raw observations. They are low-level interpretative entities, produced through an analytical process that depends on the observer’s expertise, theoretical background, and recording practices. The definition of a stratigraphic unit involves the use of multiple tools, standardised forms, drawings, photographs, written descriptions, but ultimately rests on an interpretative act performed by the archaeologist.

What becomes archaeological data, therefore, is already interpretation. We may define this as a low-level interpretation, in the sense that it does not build upon other explicitly articulated interpretative frameworks but emerges directly from the interaction between the observer and the observed material. This form of inference stands very close to abduction: it is immediate, situated, and only partially formalised.

The continuous addition of interpretative layers throughout archaeological practice blurs the distinction between data and knowledge. Archae-

ology constantly produces entities that lie somewhere between observation and interpretation, measurement and explanation. This raises a fundamental question: when archaeologists produce “data”, what are they actually producing? Several authors have highlighted the need to distinguish between data production and knowledge production as related but distinct activities (Baker, Matthew & Mayernik 2020). Within this framework, it is useful to introduce a third heuristic concept: information. While an exhaustive discussion of the definitions of data, information, and knowledge (see Zins 2007) lies beyond the scope of this introduction, it is important to clarify the position adopted here. Archaeological data, as we understand them, are best conceived in relation to information levels. Information emerges from the process of relating data to existing knowledge, personal, disciplinary, and contextual, thus enabling understanding.

From this perspective, data represent the lowest level in a hierarchy of interpretative layering. Yet even at this level, archaeological data are never “given” (see Huggett 2022 with reference therein) in a neutral sense. They occupy an intermediate position between something that exists independently of observation and something that is actively constructed by the data collector. Archaeological data are therefore neither purely objective nor entirely subjective: they are situated products of analytical practice.

INFORMATIONAL LEVELS AND INTERPRETATION IN ARCHAEOLOGICAL DATA PRODUCTION

If archaeological data are the result of stratified interpretative processes, it becomes essential to clarify how these processes operate and how different informational levels emerge during archaeological practice. Data production in archaeology is not a single act, but a sequence of (continuous) analytical operations that progressively transform material traces into entities that can be recorded, compared, and analysed. A key concept in this respect is that of “informational level”. Archaeological entities are produced and defined at different levels of abstraction, each corresponding to a specific analytical operation and interpretative effort. These levels are not fixed or intrinsic properties of the archaeological record, but

the outcome of methodological choices made during data collection. In this sense, the archaeologist functions as the primary “measurement instrument”. Unlike instrumental measurements, where precision and error can be formally assessed, the “measurement” performed through excavation is intrinsically interpretative. The criteria used to distinguish one stratigraphic unit from another are theory-laden and context-dependent, and the precision of this measurement is often difficult to quantify. Nevertheless, these interpretative decisions form the foundation upon which most if not all subsequent analyses are built. A similar mechanism operates in the definition of other categories of archaeological data. A potsherd, for instance, becomes an archaeological datum not simply by being observed, but by being described, represented and linked to a hypothesised original vessel, and, in the end, being classified. This interpretative step adds a further informational layer, connecting the observed object to a broader analytical construct. In this case, the interpretative distance between observation and data is often smaller than in the definition of stratigraphic units, and the reliability of the measurement may be assessed more easily through comparison, replication, and typological consistency. Nonetheless, interpretation remains an essential component of the process.

These examples illustrate that archaeological data are produced through a succession of interpretative acts that operate at different informational levels. Each level reduces the complexity of the material record while simultaneously introducing new assumptions and constraints. Data production in archaeology is therefore not a process of extracting neutral facts, but of constructing analytically meaningful entities that are suitable for further use.

Recognising the existence of multiple informational levels is crucial for understanding both the strengths and the limitations of archaeological datasets. It allows us to move beyond simplistic oppositions between “raw” and “processed” data, and to acknowledge that even the most basic archaeological data are already shaped by methodological and theoretical choices. This awareness is particularly important when data are reused, aggregated, or subjected to quantitative analysis,

where differences in informational level can have significant analytical consequences.

TYPES OF ARCHAEOLOGICAL DATA AND DEGREES OF MEASUREMENT PRECISION

Archaeological datasets are characterised by a high degree of heterogeneity, not only in terms of content but also in the ways data are produced and measured. Different categories of archaeological data correspond to different measurement processes, degrees of precision, and levels of interpretative layering. Recognising this variability is essential for any analytical approach, and particularly for quantitative analyses that implicitly assume commensurability among data.

A first broad distinction can be drawn between data that are primarily produced through human observation and classification, and data generated through instrumental measurement. Stratigraphic units, typological classifications, and contextual attributions belong to the former category. As discussed above, in these cases the archaeologist acts as the measurement instrument. The criteria used to define and record such data are grounded in disciplinary conventions and individual expertise, and the evaluation of measurement precision and error is rarely straightforward. Uncertainty is often implicit, qualitative, and difficult to formalise.

By contrast, other categories of archaeological data are produced through the measurement of physical or chemical properties of the archaeological record. Radiocarbon dates, isotopic ratios, elemental compositions, and other archaeometric data are generated by technical instruments operating according to well-defined protocols. In these cases, measurement precision and error are explicitly quantified and documented, and the reliability of the measurement can be assessed in probabilistic terms.

However, the apparent epistemic clarity of instrumental data can be misleading. While the measurement process itself may be precise, the archaeological meaning of the resulting values depends on a series of interpretative steps that link the measured quantities to archaeological questions. A radiocarbon determination, for ex-

ample, becomes archaeologically meaningful only when the dated sample is securely associated with a context of provenance. The interpretative layers involved in defining that context, often based on stratigraphic reasoning and classification—are thus indirectly transferred to the numerical result.

This highlights a crucial point: precision in measurement does not coincide with accuracy in archaeological interpretation. Highly precise numerical data may rest on interpretative foundations that are comparatively fragile, while data produced through observational classification may, in some cases, be more robust with respect to specific analytical goals. The overall reliability of an archaeological datum is therefore not a property of the measurement alone, but of the entire chain of operations that produced it.

Different types of data also offer different analytical potentials depending on the level at which they are used. The same datum can be exploited at multiple informational levels, each associated with different degrees of uncertainty. For instance, a radiocarbon date may be suitable for modelling long-term demographic trends at a regional scale, while being insufficiently precise for resolving fine-grained stratigraphic sequences within a single site. Conversely, typological classifications may support detailed intra-site analyses but become increasingly unstable when aggregated across large spatial or temporal scales.

Understanding the relationship between data type, measurement precision, and informational level is therefore a prerequisite for meaningful analysis, a property magnified for quantitative ones. Treating heterogeneous data as if they were equivalent, or ignoring the interpretative layers embedded in their production, risks generating results that are formally correct but archaeologically unsound. Making these differences explicit is a necessary step toward responsible data reuse and aggregation.

AGGREGATION, CONTEXT, AND ANALYTICAL GOALS

The production and use of archaeological data cannot be understood without explicitly addressing the role of aggregation. Aggregation is not a

secondary or merely technical step that follows data collection; rather, it is a fundamental analytical operation through which data acquire meaning and become usable for specific research purposes. In archaeology, aggregation is unavoidable, and it operates along at least two distinct but interrelated dimensions: vertical and horizontal.

Vertical aggregation refers to the integration of data across different informational levels. Archaeological entities are routinely aggregated from lower to higher levels of abstraction: stratigraphic units are grouped into contexts, contexts into sites, and sites into regional or supra-regional analytical units. Each step of vertical aggregation involves a change of informational detail and the introduction of new interpretative assumptions. This process is not inherently problematic; on the contrary, it is what allows archaeological data to be mobilised beyond the scale at which they were originally produced. However, vertical aggregation always entails a trade-off between analytical generalisation and loss of contextual specificity.

Horizontal aggregation, by contrast, involves the combination of data produced at the same informational level but originating from different contexts, sites, or datasets. Examples include the aggregation of site-level inventories across a region, the merging of typological datasets derived from different publications, or the integration of excavation records produced by different research teams. Horizontal aggregation requires standardisation, harmonisation of categories, and the resolution of terminological and classificatory inconsistencies. These operations are intrinsically interpretative and depend on explicit or implicit assumptions about equivalence and comparability.

Both forms of aggregation are strongly shaped by the analytical goals of the research. There is no a priori correct level at which archaeological data should be defined or aggregated. Instead, the level is determined by the research question being addressed. When analysing the relative chronology of a settlement for example, stratigraphic units may represent the most suitable analytical entities for distributing artefact types and establishing temporal relationships. In contrast, when investigating the regional distribution of material culture to explore patterns of interaction or con-

nectivity, site-level aggregation is often necessary, collapsing multiple stratigraphic units into a single analytical context.

The choice of aggregation level also affects the degree of precision required from different components of the dataset. In a network analysis based on the distribution of artefact types across sites, high precision is demanded in the attribution of objects to their sites of provenance, as well as in the definition of those sites as distinct analytical entities¹. At the same time, the spatial precision of site location may be less critical, particularly when geographic coordinates are used primarily for visualisation rather than as variables in the analysis. Conversely, in spatially explicit models, even small uncertainties in geographic positioning may have significant analytical consequences and therefore should be taken into account (Conolly & Lake 2006: 77-79).²

These examples illustrate that aggregation is inseparable from analytical intent. Data are not simply aggregated because they are available, but because aggregation enables some forms of analysis while constraining others. Vertical and horizontal aggregation define the informational space within which synthesis operate, shaping both the possibilities and the limitations of interpretation.

Making aggregation processes explicit is therefore a crucial step toward responsible data use and reuse. Without a clear understanding of how and why data have been aggregated—whether in quantitative analyses or in qualitative classifications—synthetic interpretations risk operating on entities whose informational level, pre-

cision, and interpretative grounding are poorly defined. Recognising aggregation as a central analytical practice allows us to better assess the robustness of archaeological inferences and to situate results within their proper epistemic context.

AGGREGATION AND REUSE: EXPERTISE, STANDARDS, AND RISKS

While aggregation is a necessary analytical operation, it is also one of the most epistemically fragile stages in the archaeological research process. The reuse of existing data, particularly within quantitative frameworks, involves almost by definition re-aggregation, generally performed outside the original context of data production. It is at this stage that previously implicit assumptions, inconsistencies, and uncertainties may be amplified rather than resolved.

Aggregating archaeological data requires a set of competences that extend beyond computational skills. Whether aggregation involves harmonising typological classifications, selecting sites for inclusion in a regional dataset, or merging records produced by different research traditions, it presupposes detailed knowledge of the archaeological, chronological, and historiographic context in which the data were originally generated.

A critical aspect of these processes concerns the evaluation of uncertainty and error. In many cases, uncertainty is not explicitly documented in the original datasets, particularly when data derive from legacy publications or grey literature. As a result, decisions made during aggregation rely heavily on the perceived reliability of data producers, publication venues, and disciplinary consensus. Trust, authority, and reputation thus function as meta-factors shaping the acceptance and reuse of archaeological data, often substituting for formally defined measures of data quality.

This situation highlights a fundamental tension between data creation and knowledge production. Data creation is often described as a neutral, preparatory activity, while knowledge production is understood as question-driven and interpretative. In archaeological practice, how-

¹ In a recent exercise for the course "IT Applied to Archaeology", the classification of Middle Bronze Age ceramic productions from the southern Italian Peninsula (Cocchi Genick *et al.*, 1995) was revisited using a network approach. Due to unclear naming and definition of archaeological contexts, several ambiguities emerged during the construction of the adjacency matrix. These could only be resolved with the contribution of experts in southern Italian protohistoric archaeology.

² A basic example is provided by visibility analysis. A small change in the x and y coordinates of a site can lead to a different elevation value, potentially altering the results of the viewshed analysis and affecting any subsequent models that rely on it.

ever, this distinction is difficult to sustain. Excavation, classification, and recording are themselves analytically complex and question-oriented processes, typically conducted with the aim of producing traditional scholarly outputs.

The reuse and aggregation of such data require technical standardisation and critical curation. Curation, in this sense, refers to the explicit documentation of aggregation choices, assumptions, and transformations applied to existing data. Making aggregation processes transparent and measurable is a prerequisite for assessing the robustness of quantitative analyses and for enabling meaningful reuse by other researchers. Without such transparency, aggregated datasets risk becoming opaque analytical black boxes, difficult to evaluate and impossible to reproduce.

Recognising aggregation and data curation as central activities has important implications for archaeological research practice. These activities should not be regarded as ancillary or purely technical, but as intellectually demanding contributions that warrant explicit recognition. Elevating the status of data curation and aggregation is essential if archaeological datasets, particularly legacy data, are to be reused responsibly and effectively within new analytical paradigms.

From this perspective, it is useful to distinguish between different modes of data transformation that are often conflated under the general label of “data reuse”. Archaeological data may be created or acquired through direct observation and recording, as in the definition of a context or a stratigraphic unit, where close attention is paid to contextual information and metadata.

Data may also be reconfigured, when existing records are transformed into structured datasets with limited reference to their original metadata. In these cases, datasets emerge primarily through the formalisation of descriptive attributes, often resulting in analytically convenient but contextually impoverished representations (Bevan *et al.* 2017).

Finally, data may be aggregated, when multiple datasets or classificatory systems are combined to address new analytical questions, for example through the construction of networks based on different typological traditions. Unlike simple

reconfiguration, aggregation entails explicit assumptions about comparability and equivalence across data sources.

The dynamics described above can be further clarified by considering the identity of the different actors involved. Archaeological enquiry typically involves at least three types of agents: data producers, data aggregators or curators, and data reusers. At the risk of over-simplifying, data producers, most notably excavators, generate highly contextual, largely qualitative data that provide the framework within which quantitative measurements can acquire archaeological meaning. At the opposite end, data reusers are often specialists applying formal or quantitative models to highly layered and standardised datasets.

Acting between these two poles are data aggregators and curators, whose role is to translate heterogeneous and context-dependent data into forms that are reusable across analytical frameworks. This intermediary position is epistemically crucial but institutionally fragile. Data aggregation is often highly demanding in terms of time, expertise, and familiarity with specific archaeological contexts, yet it offers limited academic recognition compared to both data production and data reuse. This expertise does not merely concern technical skills, but the ability to recognise and interpret specific informational layers, as well as to make explicit the scientific orientations that generated them. Even within the same category of data, different researchers may produce distinct layers as a result of divergent theoretical perspectives and recording practices.

The material nature of archaeological data introduces a further, often underestimated, element into this model. Archaeological data originate from unique material entities: while a potsherd or a stratigraphic unit may belong to a broader class, the informational layers emerging from its observation are irreproducible. A vessel found in a specific context of a given site cannot be re-observed elsewhere, nor can the same stratigraphy be documented twice.

Moreover, certain links between data can only be created at the moment of data production. The possibility of associating a radiocarbon date with a specific context for instance largely depends on

sampling decisions taken during excavation, and therefore on the awareness, expertise, and research orientation of the data producer. These choices condition not only what data exist, but also which forms of aggregation and reuse will be possible in the future.

This irreproducibility has important consequences in terms of academic power. Since archaeological observations cannot be replicated and primary data cannot be duplicated, privileged access to specific datasets becomes a structural advantage, fostering research opportunities and, ultimately, academic careers. In this sense, archaeological data function more as stable objects than as strategic assets, whose value lies in their potential for aggregation, reuse, and mobilisation across contexts.

As noted by Sabina Leonelli (2019), data curators have long remained the “Cinderellas of academia”. A similar marginalisation applies to those engaged in data aggregation and the construction of new informational layers. Recognising the centrality of these activities is essential for understanding both the potentials and the limitations of data reuse in archaeology, particularly within quantitative research paradigms.

DIFFERENT APPROACHES, A SHARED ATTITUDE: THE CONTRIBUTIONS TO THIS VOLUME

The papers collected in this issue address the relationship between quantitative approaches and archaeological data from complementary perspectives, focusing on different moments of the data lifecycle. Taken together, they provide a coherent reflection on some of the most critical dimensions of contemporary archaeological practice: the conditions and limits of data sharing, the epistemic implications of working with inherently quantitative datasets such as radiocarbon dates and isotopic measurements, and the reuse and transformation of material culture data within formalised analytical frameworks, including Social Network Analysis.

The contribution by Gattiglia and Vander Linden discusses data sharing in archaeology, starting from the apparent paradox of a discipline that is both strongly data-driven and yet often re-

luctant to make data openly accessible. By reviewing existing initiatives and practices, the authors show how data sharing remains uneven and often constrained by informal systems of control and privileged access. Crucially, they frame sharing not as a purely technical matter, but as a socially situated and ethically grounded practice, one that requires trust, recognition, and responsibility. As archaeological datasets grow in size and heterogeneity, effective sharing increasingly depends on the availability of hybrid expertise, capable of bridging archaeological knowledge, technical skills, and infrastructural awareness.

Closely related issues emerge from Cavazzuti’s discussion of isotopic data, which foregrounds the interpretative density of quantitative measurements. Categories such as “local” and “non-local” cannot be mechanically derived from numerical values alone, but require contextual knowledge, archaeological reasoning, and a critical understanding of analytical procedures. The paper explicitly warns against delegating interpretative responsibility to statistical tools, especially when addressing questions that are fundamentally archaeological. In this respect, quantitative data demand not less, but more expertise. Interpretation and modelling are shown to be mutually dependent: analytical frameworks shape how data are produced, documented, and reused, while the structure and quality of data constrain the kinds of questions that can meaningfully be asked.

Capuzzo and Crema’s contribution on radiocarbon dating further reinforces this point by illustrating how accuracy and precision are the outcome of a long and cumulative chain of practices. Radiocarbon dates are not isolated measurements, but the result of sampling strategies, fieldwork decisions, laboratory protocols, and quantitative analyses. Their reliability, and thus reusability, depends on the transparency and robustness of each step in this chain. The paper highlights both the opportunities and the risks associated with large datasets: while increased sample sizes allow uncertainty to be explored rather than suppressed, they also amplify the consequences of poorly documented or weakly contextualised data. Legacy data, in this sense, are in-

separable from the research histories and interpretative frameworks that produced them.

These themes find a particularly clear synthesis in the contribution by Iacono, Michienzi and Arena, which addresses the relationship between archaeological data and quantitative modelling through the lens of network analysis. Their review highlights the considerable potential of network-based approaches for the comparative analysis of large and heterogeneous datasets, while also stressing that meaningful comparison requires a careful formalisation process aimed at transforming diverse archaeological data into commensurable analytical units. Network methods are shown to be especially effective in revealing patterns of interaction that are not immediately visible at smaller spatial or chronological scales, allowing researchers to “zoom out” and explore relational structures across broad regions and long time spans.

At the same time, the paper makes clear that such transformations are never neutral. Different types of networks—based on locational data, artefactual evidence, or hybrid combinations of the two—entail different assumptions, proxies, and interpretative consequences. More broadly, the contribution underlines a relational understanding of archaeological phenomena: what matters is not the intrinsic coherence of cultural categories, but the network of relationships that connects places, objects, and practices across different scales. Abstracting archaeological data into network components can therefore generate new insights, but only if the process of abstraction itself is critically examined and theoretically informed.

Taken together, the papers in this issue converge on a shared conclusion. The increasing scale, complexity, and interdisciplinarity of archaeological research do not reduce the need for expert knowledge; they intensify it. Quantitative methods amplify the epistemic consequences of decisions made during data collection, formalisation, aggregation, and reuse. Addressing these challenges requires not only methodological innovation, but also sustained investment in skills, documentation practices, and shared infrastructures that make archaeological data intelligible,

traceable, and reusable over time. In this sense, comparison, modelling, and synthesis are not endpoints, but ongoing processes that depend on the continuous negotiation between data, theory, and scholarly communities.

REFERENCES

- BAKER, MATTHEW & MAYERNIK (2020) - Disentangling knowledge production and data production, *Ecosphere* 11(7):e03191. [10.1002/ecs2.3191](https://doi.org/10.1002/ecs2.3191)
- BEVAN A., COLLEDGE S., FULLER D., FYFE R., SHENNAN S. & STEVENS, C. (2017) - Holocene fluctuations in human population demonstrate repeated links to food production and climate, *Proceedings of the National Academy of Sciences of the United States of America* 114(49), E10524-E10531. <https://doi.org/10.1073/pnas.1709190114>
- BINFORD R. L. (1962) - Archaeology as Anthropology, *American Antiquity*, 28(2): 217-225. <https://doi.org/10.2307/278380>
- CLARKE D (1973) - Archaeology: the loss of innocence, *Antiquity* 47(185): 6-18. [10.1017/S0003598X0003461X](https://doi.org/10.1017/S0003598X0003461X)
- CLARKE D. (1968) - *Analytical Archaeology*, London: Routledge.
- COCCHI GENICK D., eds. (1995), *Aspetti Culturali Della Media Età del Bronzo nell'Italia centro-meridionale*. Firenze: Octavo.
- Conolly J. & Lake M. (2006) - *Geographical Information Systems in Archaeology*, Cambridge: University Press, Cambridge.
- DI RENZONI A. (in press) - Concluding remarks. On the Shoulders of Giants, in Pizzuti E., Palazzini F., Bossolino I., Casarotto A., Monti D. (eds.), *Out of Neglect, Into the Spotlight. Strategies and Good Practices for Rediscovering Archaeological Sites with Troubled Histories*, Proceedings in Archaeology and Heritage, Springer.
- HUGGETT J. (2022) - Data legacies, epistemic anxieties, and digital imaginaries in archaeology *Digital*, 2(2): 267-295. <https://doi.org/10.3390/digital2020016>
- KRISTIANSEN K. (2014) - Towards a New Paradigm? The Third Science Revolution and its Possible Consequences in Archaeology, *Current Swedish Archaeology* 22(1): 11-34. [10.37718/CSA.2014.01](https://doi.org/10.37718/CSA.2014.01).
- LEONELLI S. (2019) - Data — from objects to assets, *Nature* 574: 317-320. <https://doi.org/10.1038/d41586-019-03062-w>
- ZINS C. (2007) - Conceptual approaches for defining data, information, and knowledge, *Journal of the American Society for Information Science and Technology* 58: 479-493. <https://doi.org/10.1002/asi.20508>

