This PDF is a simplified version of the original article published in Internet Archaeology under the terms of the Creative Commons Attribution 3.0 (CC BY) Unported licence. Enlarged images, models, visualisations etc which support this publication can be found in the original version online. All links also go to the online original.

Please cite this as: Fadioui, A. 2026 Building Data Models for Archaeology: The case of the TETRARCHs Storytelling Data Model, Internet Archaeology 71. https://doi.org/10.11141/ia.71.1

Building Data Models for Archaeology: The case of the TETRARCHs Storytelling Data Model

Aida Fadioui

This article presents a methodology rooted in grounded theory which was developed through the crafting of a 'Storytelling Data Model' for the Transforming data rE-use in ARCHaeology project (TETRARCHs). This model aims to support the reuse of archaeological data for storytelling purposes across diverse audiences, be it archaeology professionals, creative practitioners, or partners from memory institutions, organisations, and platforms. In the context of TETRARCHs, storytelling is conceived as not just a means of supporting and reimagining data reuse, but simultaneously as a practice of interpretation and meaning-making. This article further discusses the representational implications of data modelling practices in archaeology and how the TETRARCHs Storytelling Data Model has sought to engage with and address these challenges.

1. Introduction

Archaeological data are notoriously diverse, incomplete, and in many cases difficult to organise. Putting things and thoughts into distinct categories is practically human nature; it certainly makes remembering and finding information a lot easier. In most fields, whether bioscience, geology or ceramic analysis, people rely on classifications to make sense of their data. Yet, classification is never neutral: how we choose to categorise phenomena reflects particular epistemologies, values, and aims (Bowker and Star 1999). In archaeology, every act of categorisation is also an act of representation, shaping what aspects of the past are made visible and what is left behind (Hacıgüzeller *et al*. 2021). Data models do not merely store information, they perform the work of interpretation by structuring knowledge, defining what counts as meaningful, and enabling certain interpretations at the expense of others.

Within the past decade, (born-)digital archaeological data production practices in the Global North (particularly Europe and North America) have thrived, driven by increasingly accessible hardware (often portable and internet-connected), user-friendly software, and growing digital literacy of the participants in these infrastructures (Batist and Roe 2023; 2024). In this context, it is also becoming clear that the implications of data modelling should not be an afterthought, but its careful planning and consideration of its agency should rather precede and structure data collection (Huvila and Ekman 2024; Hyvönen 2012; Kansa *et al.* 2014; Katsianis *et al*. 2023). If the objective is that all data collected is useful, or rather that it follows the FAIR (Findable Accessible Interoperable Reusable) data principles, having systems representing knowledge organisation or ontologies that support these goals is essential (Rabinowitz 2022). Scholarship on the relationship between archaeological data and the FAIR data principles (Wilkinson *et*

al. 2016) continues to grow and to advocate for gradual change driven by community-based efforts (Nicholson *et al*. 2023). Because they have historically been challenging for archaeologists, preservation and access remain priorities within the field, and significant work has been done regarding Findability and Accessibility, including encouraging the use of persistent identifiers and standardising cataloguing practices, especially the use of linked open data standards and ontologies which also push interoperability forward (Kansa *et al*. 2014; Neller *et al*. 2024; Wright 2021). Data reuse remains largely explored through technical and analytical terms (Faniel *et al*. 2018; Huggett 2018; Huvila *et al*. 2024; Huvila and Ekman 2024; Yakel *et al*. 2024). As Seaton and Lauzikas (2023) remark, lack of contextual data, issues regarding clear reuse licenses, inconsistently digitised datasets not following standards are all factors which discourage reuse. And while it is clear that reuse is a technical and infrastructural matter (Faniel *et al*. 2018; Huggett 2018; Huvila *et al*. 2024; Huvila and Ekman 2024; Yakel *et al*. 2024), it is also a cultural one, and ethically responsible and stronger community standards are also needed. In the context of the TETRARCHs (*Transforming Data Reuse in Archaeology*) project, storytelling is considered a means of supporting and reimagining archaeological data reuse (see Perry *et al*. 2025). Specifically, storytelling is understood as a means for data interpretation and meaning-making capable of prioritising embodied, affective, and sensory dimensions of past lives, crucial for narrative reconstructions (Bria and Vasquez 2022; Fennelly 2023; Habermas 2018; Praetzellis 2014). Existing data models typically used in archaeology (either local, project-specific, or overarching as in the case of CIDOC CRM, *see below*) are only beginning to accommodate the interpretative dimensions of data, still prioritising structure and material descriptions over human experiences, thus excluding the emotional, sensory, and narrative aspects that are central to how people engage with the past (Canning *et al*. 2022).

The TETRARCHs project is a three-year European initiative (2022-2025) which brings together an interdisciplinary group of archaeological specialists, data scientists, creative residents, and museum practitioners across more than a dozen European institutions and a half-dozen archaeological sites. The project aims to optimise archaeological data for reuse by bridging the gaps in knowledge regarding data reusability through the co-creation and assessment of experimental workflows (see Perry *et al*. 2025), and the development of tools and resources, and the promotion of best practices. TETRARCHs does so by collaborating mainly with three audiences, namely archaeological specialists, creative practitioners, and professionals from memory institutions. For clarity and concision, this article refers to creative practitioners, professionals from memory institutions, and other audiences engaging with archaeological data in storytelling contexts collectively as ‘non-specialist' audiences or communities, that is, *non-archaeologists*. This terminology is used solely for analytical distinction and does not imply a lack of expertise. On the contrary, these participants bring diverse professional and creative forms of knowledge, particularly storytelling and interpretation, which are central to the project's aim. Engaging non-specialist audiences is not just a matter of dissemination, it is also a matter of epistemic enrichment. Non-specialists bring interpretative and affective approaches that can challenge disciplinary assumptions of archaeological data.

This article explores how archaeological data models can be designed to support storytelling as a form of meaningful data reuse in the context of the TETRARCHs project, and how such models might represent the sensory, emotional, and embodied dimensions of archaeological knowledge. In the first section, I explain the project's vision for data modelling and its relationship to storytelling and give a brief overview of the qualities of different data models used in archaeology and heritage. In the next section, I present the different stages of the methodology used to craft the TETRARCHs storytelling data model (in development),

informed by grounded theory (Charmaz 2014), detailing selected parts of the process. In the penultimate section, I draw on cultural studies and theories of representation, to situate data modelling as a performative, political and ethical act that represents the past. I conclude by reaffirming the importance of finding socio-technical solutions to computational issues, and challenging disciplinary assumptions.

2. Archaeological data in support of storytelling

Although the TETRARCHs project initially had only a broad vision for an alternative data model to support compelling storytelling, its specific objectives emerged gradually through the activities conducted during the project. Developing such a model was at first a peripheral consideration, but its importance was soon recognised as central to enabling richer and more ‘generous' forms of data retrieval (see Whitelaw 2015) that could support both specialist and non-specialists in their storytelling endeavours. Insights from the activities organised throughout the project revealed that storytellers required ways of recording affective and sensory dimensions of archaeological experience (see Krmpotich and Somerville 2016). Storytelling, as Tringham (2020) argues, can provide contextual information about archaeological features, but doing so requires more diverse forms of data, data about interpretation, emotion, and sensory perception. These information needs must therefore be at the centre of this design process. While archaeologists do engage in storytelling (see next paragraph), context sheets and databases are not designed with storytelling as a primary purpose or an anticipated mode of reuse. This section will focus on the qualities of existing data models, specifically whether they account for nuance and record interpretative, emotive, and sensory data. The challenge lies in determining what information recorded by existing models (e.g. CIDOC CRM) is already being reused by storytellers, and what other types of information are needed to support storytelling by different audiences. The recognition of storytelling as both a design need and an interpretive tool within TETRARCHs invites a broader reflection on why storytelling itself is central to how archaeological knowledge is produced and shared.

In his examination of scientist storytellers and the stories they tell, Terrell (1990, 4) noted that "facts may have a material basis outside people's heads, but on their own they don't have meaning". Storytelling is a sense-making and world-building activity through which facts are interpreted and given significance. Immersion in stories can shape individuals' beliefs, behaviours, and attitudes (Ratcliff and Sun 2020). Within TETRARCHs, storytelling is approached from a hermeneutical standpoint as a practice which is not only one of meaning-making, but also one of mediation and interpretation (see Brockmeier and Meretoja 2014; Meretoja 2018). Storytelling in this sense fosters interaction between people who find and care for heritage, and those who might find meaning and value in it. Previous research has shown that storytelling is already embedded within archaeological practice, whether through unconsciously crafted narratives or through the different modalities of storytelling deployed to communicate archaeological findings (Bria and Vasquez 2022; Hagedoorn and Sauer 2018; Helden and Witcher 2019; Mickel 2015; Pluciennik 1999; Terrell 1990; Thomas 2015). Archaeologists, as ‘experts' and as agents of this mode of knowledge creation, have a responsibility and a moral duty to work with communities of knowledge in world-making practices (Thomas 2015). This involves helping others investigate the potentialities of archaeology and supporting the crafting of more compelling and politically just narratives and representations.

This understanding of storytelling as a relational and interpretive practice extends beyond professional archaeology. The same sense-making and world-building capacities that shape archaeological narratives also underpin the diverse ways in which non-specialist

communities and creative practitioners engage with archaeological data. The reuse of archaeological data by non-specialist communities through storytelling offers both opportunities for wider engagement and risks of distortion. Non-specialist communities have long used archaeological information in diverse forms of storytelling (Bria and Vasquez 2022; Byrne 2012), regardless of whether such data was easily findable, complete, or properly contextualised. Laužikas *et al*. (2018) contend that archaeological knowledge is also generated within peripheral, non-professional communities, noting that jewellers, musicians, filmmakers, writers, game designers, tour guides, educators, and other creative practitioners frequently reuse archaeological data in their work. They further argue that when knowledge moves from the 'centre' to the 'periphery', its meaning undergoes transformations that might lead to misrepresentation or simplification. These dynamics echo Meretoja's (2018, 12) concern that narratives presented as "neutral, perspective-less statements", rather than as situated interpretations, flatten their meaning. Such flattening could contribute to the construction of harmful or exclusionary political discourse. Thomas (2015, 171) argues that archaeologists bear an ethical responsibility not to misrepresent their findings in a way that can be used to "perpetuate structural violence or socio-political injustice upon people of the present". This responsibility should be a driving force for them to communicate their narratives more reflexively. Within TETRARCHs, recognising that storytelling practices have great creative potential but can also entail risks is crucial to understanding how data reuse can support more inclusive, reflexive, and ethically grounded engagements with the past.

These questions of representation and responsibility are not confined to the narratives archaeologists tell but are also embedded in the infrastructures that shape how archaeological knowledge is recorded and reused. Data models, as the technical embodiments of classificatory and interpretative choices, therefore, play a significant role in determining what kinds of stories can be told with archaeological data. Assessing their capacity to support storytelling requires a closer examination of some of the more commonly used models in the context of digital humanities and digital heritage in the Global North. The models assessed in this article serve as building blocks for the storytelling data model, because they are the most widely used said model will also be harmonised with their logic. The CIDOC Conceptual Reference Model is an event-centric model, providing a formal structure for cultural heritage documentation. While it can record complex narratives of provenance, temporal, or spatial relationships, it is not designed to capture sensory or emotive data, limiting its capacity to support storytelling. CRMarchaeo, a modular extension of CIDOC CRM for archaeological excavation process documentation, allows for great expressivity regarding stratigraphic relationships, excavation contexts, and the interpretation of chronological sequences, but similarly does not accommodate non-traditional or sensory knowledge. In contrast, the Odeuropa project, focusing on sensory heritage, created an olfactory extension of CIDOC CRM targeting smell-related information (Lisena *et al*. 2022). Its three-layered, event-based model combines classes and properties from CIDOC CRM and the Scientific Observation Model (CRMsci), and classes representing sensorial experiences. Odeuropa centres subjective and deeply affective modalities of interacting with smell heritage, capturing its perception and contextual embedding (van Erp *et al*. 2023). While the model is limited to olfaction, it shows exciting potential for data models to represent embodied, emotional, and sensory experiences in ways that directly support storytelling. The Europeana Data Model (EDM), an RDF-based model designed for the open data meta-aggregator Europeana, supports rich contextualisation and can hold multiple, possibly contradictory records. While primarily focused on resource description rather than sensory or emotional data, EDM accommodates data from third parties, and enables the integration of diverse metadata, enabling forms of contextual storytelling. One such ontology which has aimed to extend EDM is the Narrative Ontology (NOnt) (Bartalesi *et al*. 2019), a

CIDOC CRM extension developed to provide a formal representation of narratives in the context of digital libraries. The Narrative Ontology is event-based and a means to record narratives as an autonomous data type, offering a structured way to model the semantic relations in narratives and connect them to objects already stored in digital libraries (Meghini *et al*. 2021).

3. The TETRARCHs Storytelling Data Model: a bottom-up attempt at data modelling

The TETRARCHs storytelling data model was crafted by approaching data modelling from a bottom-up perspective. This methodology is not uncommon, the starting point of data modelling is often a particular collection or database that needs describing, as was the case with CIDOC CRM for instance (Bruseker *et al*. 2017). The TETRARCHs storytelling data model however is designed based on the needs of storytellers. To achieve this, two preliminary aims had to be realised: to understand who the storytellers reusing archaeological data are, and to assess what specific types of said archaeological data they were reusing. At this stage, the aims of the data model were defined: it was to enable the recording of sensory experiences, emotional data (e.g. feelings, moods, attitudes), and actions (including internal human processes such as cognition or perception). Because the model aims to support the crafting of stories, it must capture data which can help setting a scene. It must therefore integrate existing data classes from other models, for example those describing people (e.g. E39 Actor, E21 Person, E74 Group), find types, and condition states (e.g. E55 Type, E3 Condition State). As the model aims to enrich narratives and allow different perspectives to be integrated to the record, legitimising a diversity of truths, it must also strive to support polyphony, the narrative feature coined by Bakhtin as the notion of coexisting voices (see Bakhtin 1984). While the model primarily aims to capture sensory and emotional data related to archaeological evidence, rather than recording narratives, it is also intended to complement other models such as Narrative Ontology in the documentation of narratives as data types, both because sensations and emotional shifts drive narrative progression, and because it can catalogue key narrative elements such as actors, the incipit (describing setting or initial conditions), or peripeteia (actions).

The design of the storytelling data model was informed by an iterative, data-driven research process inspired by grounded theory (GT; see Birks and Mills 2015; Charmaz 2014; Chun Tie *et al*. 2019; Glaser and Strauss 2017). In GT, analysis proceeds inductively from empirical data, allowing theoretical categories to emerge from patterns observed in the data rather than being imposed in advance. Within TETRARCHs, this meant developing conceptual elements of the model through three types of storytelling activities involving the project's target audiences: archaeology specialists, creative practitioners, and professionals from memory institutions, organisations, and platforms. The first set of activities comprised three surveys, each incorporating a storytelling prompt adapted to distinct audiences. Participants conducted online searches to construct archaeological narratives and reported their search terms. While the first survey, sent out to archaeology specialists, was successful (32 respondents), a limitation in the methodology did arise during the survey phase. The two surveys targeting non-specialists received extremely low response rates: three from creative practitioners and five from professionals from memory institutions. Due to this limited participation, the resulting data could not be considered representative and was therefore excluded from further analysis. The reasons for low engagement remain unclear, whether stemming from the survey design, limited dissemination, or lack of relevance or interest among potential respondents. The second and most data-rich methodology involved annotation and reuse of archaeological data. During different workshops organised by TETRARCHs colleagues Sara Perry and Anna Simandiraki-Grimshaw in the UK, and Paola Derudas in Sweden, participants were asked to annotate by hand diverse materials such as

photographs, reports, 3D models, which were either printed out or accessible on a computer or tablet, and in some cases to reuse them in the creation of a story. Said participants included all three audiences previously described as the target audiences, and some also included other interested audiences such as children and university students (A table summarising all information regarding these activities will be accessible on Zenodo; see TETRARCHS project deliverable D3.1, Gap Analysis and Recommendations). The third source of data was focus group meetings, which were an integral part of the project (A report delving into the findings of these focus groups is available on Zenodo; deliverable D3.3, Modelling for Different Audiences Workshop). These focus groups involved all three target audiences and focused on their research processes when constructing stories. Across these three activities, data were coded and categorised using the NVivo software to identify recurring themes relevant to storytelling. Following the GT process, initial coding identified broad categories of information, focused coding then refined these into core categories and additional data was gathered to saturate said categories, meaning that further data collection through other iterations of the previously described activities no longer yielded new insights for these categories. Finally, advanced coding transformed core categories into conceptual terms representing the emerging classes (Figure 1).

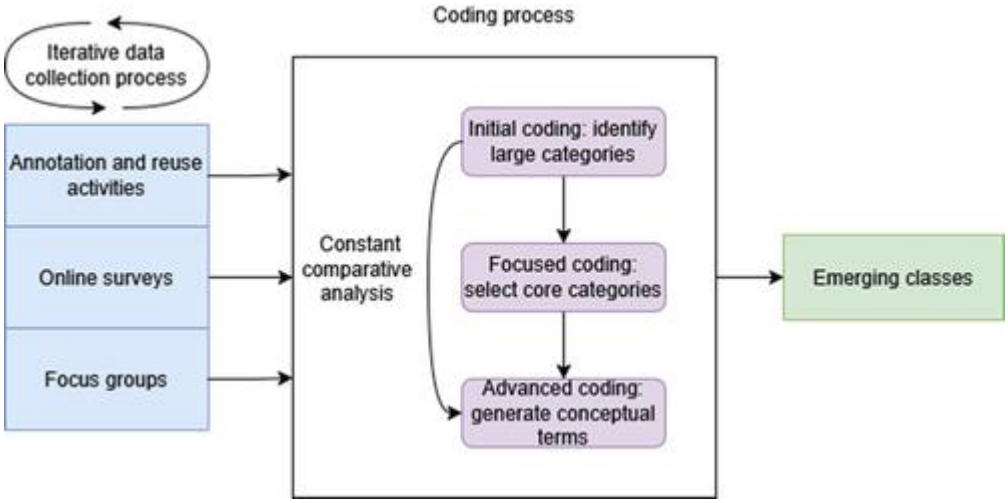

Figure 1: Iterative research design framework.

These emerging classes were then mapped to data classes of existing models such as CIDOC CRM when possible or became a new class of the TETRARCHs storytelling data model. Several core principles of GT guided this process, one being purposive sampling i.e. the selection of participants most likely to generate relevant insights, in our case the three target audiences described throughout this article which were identified in the TETRARCHs project description. Constant comparative analysis was another core principle, consisting of data, codes, and emerging categories, being iteratively compared to each other to systematically develop theory. Finally theoretical sensitivity, described by Chun Tie *et al.* (2019) as the ability to recognise a data segment relevant to one's theory, meaning how a researcher develops conceptual awareness through immersion in the data, theoretical knowledge, and the use of different analytical techniques. These principles ensured that the resulting model was both empirically grounded and responsive to the interpretative practices of its intended users. While the data model is still under active development, the classes represented in Figure 2 are stable classes which will be part of the final model.

Figure 2: Semantic representation of the storytelling data classes.

The GT analysis provided the conceptual foundation for the storytelling data model, identifying key entities and relationships relevant to storytelling. To ensure the emerging model could adequately represent them, a complementary method, the development of competency questions (CQs), was introduced. CQs are natural language sentences expressing "a pattern for a type of questions people expect an ontology to answer. The *answerability* of CQs hence becomes a functional requirement of the ontology." (Ren *et al*. 2014, 752). The ontology design is considered complete when each CQ can be expressed as a SPARQL query (a request written in the SPARQL language to retrieve data from an RDF-based knowledge graph), ensuring that all relevant elements and relations are represented. As such, CQs guide the design of the data model and will serve to evaluate it to iteratively improve it. TETRARCHs activity participants did not explicitly formulate these questions; they were instead derived from analytical memos made throughout the activities. For instance, during a focus group with professionals from memory institutions, participants discussed the creation of an exhibition on the theme of medieval children's toys, including interactive reproductions visitors could engage with to touch and play with. This scenario informed the first CQ listed in Table 1. Additional questions were also developed based on the needs of archaeological specialists, based on informal conversations held during conferences, excavations, and other professional contexts, and will be used to assess the model in the next phase of work.

Table 1: Competency questions for the TETRARCHs storytelling data model.

Competency question	**Interrogated classes**
What did Peterborough 12th century market square smell like?	Sensory Experience (smell), Landscape, Socio-Cultural Context
What did winter feel like in Punis Tharros?	Emotion, Landscape, Atmosphere
How did ancient Romans perceive disabled bodies?	Actor, Lived Experience, Cognitive Process<

The alignment of the data model with existing ontologies is ongoing to ensure seamless integration and interoperability. The implementation phase of this work, mainly involving its validation against a use case, will likely require additional technical expertise. Nonetheless, engagement with the data has already yielded insights extending beyond technical

considerations. The data from all TETRARCHs activities highlight a strong interest among storytellers in engaging with emotion, sensation, and embodied knowledge. Most of the narrative-relevant information which is of interest to storytellers already exists in archaeological practice, whether it is information regarding human activity and practice, emotion, affect, or sensory and embodied experiences. Such data is however often excluded from the formal archaeological record due to perceptions that subjective or affective interpretations lack professional legitimacy. The findings point to clear interest in diverse ways of knowing and meaning-making within archaeology that extend beyond conventionally defined objective empirical data, also including subjective, emotive, and sensory dimensions of data. This pattern was consistent across all three audiences. In addition, there is a strong interest in people; not only from the past, but also the contemporaries who might have interacted in a multitude of ways both with the data and its subject matter. Ultimately, people are interested in stories about people.

4. Data modelling: performance and representation

Representation, as conceived by sociologist and cultural theorist Stuart Hall, actively constructs reality rather than just reflecting it: it is doing, modelling, performing. Hall (1997b) argues that meaning can only be produced within shared frameworks of discourse, and that it is only through these cultural and linguistic frameworks that we can make sense of the material world. Meaning is therefore ever shifting, never fixed, and, crucially, never neutral. Every aspect of what a person's identity may be is unstable. Similarly, sociologist George Herbert Mead, one of the founders of symbolic interactionism, proposed that identity is a temporal state of the self, shaped by past experiences and mediated by memory (Ezzy 1998; Mead and Morris 1934). In Hall's lectures (see Hall 1997a), he describes race as a shifting category whose meaning is ever changing depending on historical and cultural context, as would culture, sexuality, ideology, and every other possible component that comes into play in the construction of identity. Hall further argues that identities are negotiated, constructed through creative cultural expression and resistance to stereotypes (Alexander 2009). This is reflected in the storytelling model through classes which can record an actor's (crm:E39 Actor) ‘Lived Experience', and ‘Socio-Cultural Context' (Figure 2). Both of these classes aim to allow the recording of a multitude of data, as one can for example be ‘from' a specific cultural context but living in a different one (for instance, accounting for an individual's social mobility or historical organisation of gender), or, as was famously the case for remains of the Viking warrior in Birka grave Bj 581 who were long misidentified as male, these classes can serve to record the different interpretations of social identity made by archaeologists through time.

The performative nature of representation described by Hall is concretely expressed in classification practices: as they embed meaning into technical artefacts, discourse becomes infrastructure. As Bowker and Star (1999, 5) assert, standards and categories will inevitably valorise one point of view and silence another, making these classification choices inherently political acts, and potentially "dangerous" ones. Crafting a data model is a decision-making process in which most of the power resides in the hands of the data modeller. These decisions, as expressed by Canning *et al*. (2022), reflect value judgements: classes, the way they relate to each other and the meanings they hold are all ways in which power is exercised. Drawing on Butler's notion of *abjection* (see Butler 1993), what is not currently represented, by choice or because it requires additional effort to be represented as structured data, i.e. what is non-normative, transgressive, inconsistent, is equally important, because it highlights the limits of what is allowed to be represented, what counts as meaningful.

From this perspective, data themselves are not neutral records but situated representations derived from interactions recorded by specific means and tools, under particular circumstances and background conditions. Within these interactions, meaning, materiality, and agency are all at play and influence the data's content and character (Batist 2024; Bokulich and Parker 2021; Huvila *et al*. 2022; Wylie 2017; Yarrow 2008). The models used to collect data will always misrepresent aspects of the world, as there can be no perfect representation. Yet, as Bokulich and Parker contend in their paper on the pragmatic-representational view of data, representational adequacy is always dependent on purpose: possible misrepresentation does not preclude data being used "coherently and successfully to answer particular questions of interest" (Bokulich and Parker 2021, 31). If data are representational, this does not exclude their capacity to inform about various aspects of the world, only that these aspects are always enabled or constrained by the processes involved in deriving them. From a pragmatic-representational perspective, a model should be "adequate for the purpose of interest" (Bokulich and Parker 2021, 31); in the case of the TETRARCHs storytelling data model, that purpose is capturing emotive and sensorial data that improves affordances of storytelling. The use of competency questions to evaluate the data model, along with activities involving storytellers using the data collected with this model and the exploration of a use case (Fadioui and Giovannetti in prep) will serve as means to evaluate its "fitness-for-purpose" (Bokulich and Parker 2021, 31).

Archaeological storytelling discourses in this context provide the pragmatic horizon against which the data model is designed. They serve as fertile ground through which archaeological knowledge can be mediated by the interplay between representation, materiality, and embodied knowledge. As Swenson and Cipolla (2020, 316) argue, archaeologists transform material remains into "stories about said pasts in the present", where representation is not opposed to materiality but rather is itself a material process, enacted through bodies and discursive affordances. Representation is for these authors not separate from materiality as it is enacted by embodied engagements: on an excavation, representations of the past are produced by the embodied actions of sorting, labelling, cataloguing. Using trowels, 3D scanning, or hand drawing will offer different affordances which will shape these representations. These insights are critical for data modelling, as they define the kinds of relationships that must be afforded by the model. The TETRARCHs storytelling data model aims to put these theoretical insights into practice. By focusing on the collection of sensory experiences, emotions, and embodied practices, it attempts to expand the affordances provided by archaeological data to different storytelling audiences.

5. Concluding thoughts

This article has explored how archaeological data modelling can support storytelling as a meaningful form of data reuse, highlighting that the practice of storytelling is already embedded within archaeological practice. Through the development of the TETRARCHs storytelling data model, it has aimed to show that data relating to the sensory and affective dimensions of archaeological knowledge can be formalised while remaining methodologically rigorous. This article also aimed to highlight that data collection and modelling are themselves representational and performative acts, and that tackling this performativity exposes how technical design both reflects and constructs particular worldviews. This echoes Hall's (1997b) argument that representation constructs reality rather than simply reflecting it, and Bowker and Star's (1999) reminder that every classification is a moral and political choice. Data models, classifications, and ontologies, are not neutral containers, and recognising this is recognising that data modelling not just a technical activity but also an ethical and political practice, for "all information systems are necessarily suffused with ethical and political values" (Bowker and Star 1999, 321).

While large language models (LLMs) seem to be where a significant volume of computational research is heading, it begs the question of whether data modelling will still have any use for archaeology and heritage. I argue that it is of vital importance that prompt-based language models, which rely on statistical pattern recognition, are supported by fair and ethical semantic models. Where language models will lack understanding of cultural meaning or context, semantic data models which are fair and ethically informed can improve the outputs of said language models. Perhaps the advent of LLMs is an opportunity to craft more and more diverse and specialised semantic models, in hopes our data can be accurate, responsible, and compelling. Technical systems, as Flanders (2018, 290) writes, are "meaning systems and ideological systems, as far down as we are willing to look". From this perspective, the work of TETRARCHs reaffirms the necessity of cooperation between technical solutions and social theory.

As Radford and Joseph (2020) have argued, technical solutions can only take us so far with computational problems: social theory needs to be a part of our methodologies. Making our technical systems more humane requires integrating them into archaeological workflows, meeting archaeologists, creatives, and local communities where they are and within their own authority (Perry *et al*. 2025). Perhaps this begins simply, as the paper by Perry *et al*. suggest, through small interventions challenging archaeonormativity and normalising the recording of affective, sensory, and experiential data. The TETRARCHs project's contribution is a small step in that direction, an attempt to build alternative data models that could record a diversity of experiences, thereby making archaeological data reusable in more meaningful ways. Recording data that speaks of actions, emotions, sensations, or behaviours, allows for richer and more nuanced storytelling. This work attempts to challenge disciplinary assumptions and can hopefully be an invitation to recognise modelling as both a technical and a political act, and a form of world-building (Morgan 2025) in itself.

Acknowledgements

I am sincerely grateful to all of the collaborators involved in the TETRARCHs project, in particular Sara Perry for her guidance and feedback on a draft on this article, and for organising with Anna Simandiraki-Grimshaw as well as Paola Derudas, the different annotation and reuse data collection activities this article is based on. This research could not have happened without their support. I would also like to thank my supervisors Piraye Hacıgüzeller, Hélène Verreyke, and Christophe Verbruggen, for their guidance and expertise. All responsibility for the content of the paper is my own.

Disclosure statement

The author reports there are no competing interests to declare.

Funding

This work was done as part of the TETRARCHs project, supported by the Research Foundation – Flanders (FWO) in Belgium, the Arts and Humanities Research Council (UKRI) in the UK, the Research Council of Lithuania, the Ministry of Education, Science and Sport in Slovenia, and the FORTE Swedish Research Council for Health, Working Life and Welfare, under the CHANSE ERA-NET Co-fund program, which has received funding from the European Union's Horizon 2020 Research and Innovation Programme, under Grant Agreement no. 101004509.

Alexander, C. 2009 'Stuart Hall and 'race'', *Cultural Studies* **23**(4), 457–482. https://doi.org/10.1080/09502380902950914

Bartalesi, V., Meghini, C., Metilli, D. and Benedetti, F. 2019 'Introducing narratives in Europeana: A case study', *International Journal of Applied Mathematics and Computer Science* **29**. https://doi.org/10.2478/amcs-2019-0001

Bakhtin, M. 1984 *Problems of Dostoevsky's Poetics*, University of Minnesota Press. https://doi.org/10.5749/j.ctt22727z1

Batist, Z. 2024 'On the Value of Informal Communication in Archaeological Data Work', *Open Archaeology* **10**(1). https://doi.org/10.1515/opar-2024-0014

Batist, Z. and Roe, J. 2023 'Open-archaeo: A Resource for Documenting Archaeological Software Development Practices', *Journal of Open Archaeology Data* **11**. https://doi.org/10.5334/joad.111

Batist, Z. and Roe, J. 2024 'Open Archaeology, Open Source? Collaborative practices in an emerging community of archaeological software engineers', *Internet Archaeology* **67**. https://doi.org/10.11141/ia.67.13

Birks, M. and Mills, J. 2015 *Grounded theory: A practical guide*, 2nd edition, SAGE.

Bokulich, A. and Parker, W. 2021 'Data models, representation and adequacy-for-purpose', *European Journal for Philosophy of Science* **11**(1), 31. https://doi.org/10.1007/s13194-020-00345-2

Bowker, G. C. and Star, S. L. 1999 *Sorting things out: Classification and its consequences*, MIT press. https://doi.org/10.7551/mitpress/6352.001.0001

Bria, R. E. and Vasquez, E. C. 2022 'Digital Archaeology and Storytelling as a Toolkit for Community-Engaged Archaeology' in K. Garstki (ed) *Critical Archaeology in the Digital Age: Proceedings of the 12th IEMA Visiting Scholar's Conference* (Vol. 2, pp. 49–66). Cotsen Institute of Archaeology Press at UCLA. https://doi.org/10.2307/j.ctv2fcctzd

Brockmeier, J. and Meretoja, H. 2014 'Understanding Narrative Hermeneutics', *Storyworlds: A Journal of Narrative Studies* **6**. https://doi.org/10.5250/storyworlds.6.2.0001

Bruseker, G., Carboni, N. and Guillem, A. 2017 'Cultural Heritage Data Management: The Role of Formal Ontology and CIDOC CRM' in M. L. Vincent, V. M. López-Menchero Bendicho, M. Ioannides and T. E. Levy (eds) *Heritage and Archaeology in the DigitalAge*, Springer International Publishing. 93–131. https://doi.org/10.1007/978-3-319-65370-9_6

Butler, J. 1993 *Bodies that matter: On the discursive limits of 'sex'*, Taylor and Francis.

Byrne, S. 2012 'Community Archaeology as Knowledge Management: Reflections from Uneapa Island, Papua New Guinea', *Public Archaeology* **11**(1), 26–52. https://doi.org/10.1179/175355312X13311392295513

Canning, E., Brown, S., Roger, S. and Martin, K. 2022 'The Power to Structure: Making Meaning from Metadata Through Ontologies', *KULA: Knowledge Creation, Dissemination and Preservation Studies* **6**(3), Article 3. https://doi.org/10.18357/kula.169

Charmaz, K. 2014 *Constructing Grounded Theory*, SAGE Publications Ltd.

Chun Tie, Y., Birks, M. and Francis, K. 2019 'Grounded theory research: A design framework for novice researchers', *SAGE Open Medicine* **7**, 2050312118822927. https://doi.org/10.1177/2050312118822927

Ezzy, D. 1998 'Theorizing Narrative Identity: Symbolic Interactionism and Hermeneutics', *The Sociological Quarterly* **39**(2), 239–252. https://doi.org/10.1111/j.1533-8525.1998.tb00502.x

Fadioui, A. and Giovannetti, C. in prep 'Weaving voices into the record: the Sley data model for the integration of senses and sensitivities'

Faniel, I. M., Austin, A., Kansa, E., Kansa, S. W., France, P., Jacobs, J., Boytner, R. and Yakel, E. 2018 'Beyond the Archive: Bridging Data Creation and Reuse in Archaeology', *Advances in Archaeological Practice* **6**(2), 105–116. https://doi.org/10.1017/aap.2018.2

Fennelly, K. 2023 'Archaeology, Emotional Storytelling and Performance', *Public Archaeology* **21**(1-4), 108–122. https://doi.org/10.1080/14655187.2022.2149683

Flanders, J. 2018 'Building Otherwise' in E. Losh and J. Wernimont (eds) *Bodies of Information*, University of Minnesota Press. 289–304. https://doi.org/10.5749/j.ctv9hj9r9.19

Glaser, B. and Strauss, A. 2017 *Discovery of Grounded Theory: Strategies for Qualitative Research*, Routledge. https://doi.org/10.4324/9780203793206

Habermas, T. (ed) 2018 'How narratives evoke emotions' in T. Habermas *Emotion and Narrative: Perspectives in Autobiographical Storytelling*, Studies in Emotion and Social Interaction, Cambridge University Press. 93–176. https://doi.org/10.1017/9781139424615

Hacıgüzeller, P., Taylor, J. S. and Perry, S. 2021 'On the Emerging Supremacy of Structured Digital Data in Archaeology: A Preliminary Assessment of Information, Knowledge and Wisdom Left Behind', *Open Archaeology* **7**(1), 1709–1730. https://doi.org/10.1515/opar-2020-0220

Hagedoorn, B. and Sauer, S. 2018 'The Researcher as Storyteller: Using Digital Tools for Search and Storytelling with Audio-Visual Materials', *View* **7**(14/2018), Article 14. http://dx.doi.org/10.18146/2213-0969.2018.jethc159

Hall, S. 1997a *Race, The Floating Signifier* [Lecture transcript]. https://www.mediaed.org/transcripts/Stuart-Hall-Race-the-Floating-Signifier-Transcript.pdf [Last accessed: 2 Feb 2026]

Hall, S. 1997b *Representation: Cultural Representations and Signifying Practices,* SAGE.

Helden, D. van and Witcher, R. 2019 'Historical fiction and archaeological interpretation: Introduction' in D. van Helden and R. Witcher (eds) *Researching the Archaeological Past through Imagined Narratives*, Routledge.

Huggett, J. 2018 'Reuse Remix Recycle: Repurposing Archaeological Digital Data', *Advances in Archaeological Practice* **6**(2), 93–104. https://doi.org/10.1017/aap.2018.1

Huvila, I., Andersson, L. and Sköld, O. 2024 'Patterns in paradata preferences among the makers and reusers of archaeological data', *Data and Information Management* **8**(4), 100077. https://doi.org/10.1016/j.dim.2024.100077

Huvila, I., Börjesson, L. and Sköld, O. 2022 'Archaeological information-making activities according to field reports', *Library and Information Science Research* **44**(3), 101171. https://doi.org/10.1016/j.lisr.2022.101171

Huvila, I. and Ekman, S. 2024 'Documentation of data making, processing and use facilitates future reuse of research data: The CAPTURE project' in I. Volodina, G. Bouma, M. Forsberg, D. Kokkinakis, D. Alfter, M. Fridlund, C. Horn, L. Ahrenberg and A. Blåder (eds) *Proceedings of the Huminfra Conference (HiC 2024), 10-11 January, 2024, Gothenburg, Sweden,* Swedish Research Council. 26-30. https://doi.org/10.3384/ecp205004

Hyvönen, E. 2012 *Publishing and Using Cultural Heritage Linked Data on the Semantic Web*, Springer International Publishing. https://doi.org/10.1007/978-3-031-79438-4

Kansa, E., Kansa, S. and Arbuckle, B. 2014 'Publishing and Pushing: Mixing Models for Communicating Research Data in Archaeology', *International Journal of Digital Curation* **9**. https://doi.org/10.2218/ijdc.v9i1.301

Katsianis, M., Uleberg, E., Opitz, R., Derudas, P., Ore, C.-E., Hiebel, G., Hivert, F., Marlet, O., Nenova, D., and Bruseker, G. 2023 'Semantic Modelling of Archaeological Excavation Data. A review of the current state of the art and a roadmap of activities', *Internet Archaeology* **64**. https://doi.org/10.11141/ia.64.12

Krmpotich, C. and Somerville, A. 2016 'Affective Presence: The Metonymical Catalogue', *Museum Anthropology* **39**(2), 178–191. https://doi.org/10.1111/muan.12123

Laužikas, R., Dallas, C., Thomas, S., Kelpšienė, I., Huvila, I., Luengo, P., Nobre, H., Toumpouri, M. and Vaitkevičius, V. 2018 'Archaeological Knowledge Production and Global Communities: Boundaries and Structure of the Field', *Open Archaeology* **4**(1), 350–364. https://doi.org/10.1515/opar-2018-0022

Lisena, P., Schwabe, D., van Erp, M., Troncy, R., Tullett, W., Leemans, I., Marx, L. and Ehrich, S. C. 2022 'Capturing the Semantics of Smell: The Odeuropa Data Model for Olfactory Heritage Information' in P. Groth, M.-E. Vidal, F. Suchanek, P. Szekley, P. Kapanipathi, C. Pesquita, H. Skaf-Molli and M. Tamper (eds) *The Semantic Web,* Springer International Publishing. 387–405. https://doi.org/10.1007/978-3-031-06981-9_23

Mead, G. H. and Morris, C. W. 1934 *Mind, self and society: From the standpoint of a social behaviorist* (Facsim. ed.), University of Chicago press.

Meghini, C., Bartalesi, V. and Metilli, D. 2021 'Representing narratives in digital libraries: The narrative ontology', *Semantic Web* **12**(2), 241–264. https://doi.org/10.3233/SW-200421

Meretoja, H. 2018 *The ethics of storytelling: Narrative hermeneutics, history and the possible*, Oxford university press.

Mickel, A. 2015 "Archaeology's Epic Battles with Storytelling and Stereotypes', *Norwegian Archaeological Review* **48**, 1–4. https://doi.org/10.1080/00293652.2015.1051581

Morgan, C. 2025 'Archaeology as Worldbuilding', *Cambridge Archaeological Journal* 2025:1-10. https://doi.org/10.1017/S0959774325100164

Neller, A., Heckman, J., Bollwerk, E., Myers, K. N. and Wells, J. 2024 'Making Archaeological Collections More Findable and Accessible through Increased Coordination', *Advances in Archaeological Practice* **12**(1), 34–42. https://doi.org/10.1017/aap.2023.31

Nicholson, C., Kansa, S., Gupta, N. and Fernandez, R. 2023 'Will It Ever Be FAIR?: Making Archaeological Data Findable, Accessible, Interoperable and Reusable', *Advances in Archaeological Practice* **11**, 63–75. https://doi.org/10.1017/aap.2022.40

Perry, S., Simandiraki-Grimshaw, A., Morgan, C., Taylor, J. S., Fadioui, A., Foket, L., Hacıgüzeller, P., Sampatakou, D. V., Derudas, P., Wright, H. and Clough, A. 2025 'Towards New Futures for Archaeological Data Production: Challenging Archaeonormativity through Storytelling', *Journal of Field Archaeology* 50(8), 796-815. https://doi.org/10.1080/00934690.2025.2504235

Pluciennik, M. 1999 'Archaeological Narratives and Other Ways of Telling' *Current Anthropology* **40**(5), 653–678. https://doi.org/10.1086/300085

Praetzellis, A. 2014 'Narrative and Storytelling for Archaeological Education' in C. Smith (ed) *Encyclopedia of Global Archaeology*, Springer. 5135–5138. https://doi.org/10.1007/978-1-4419-0465-2_2095

Rabinowitz, A. 2022 'Metadata for the Masses: Facilitating Discoverability and Access in Digital Archaeological Archives' in L. Goldstein and E. Watrall (eds) *Digital Heritage and Archaeology in Practice: Presentation, Teaching and Engagement,* 1st edition, University Press of Florida. 61–84. https://doi.org/10.2307/j.ctv2pfq2jj.8

Radford, J. and Joseph, K. 2020 'Theory In, Theory Out: The Uses of Social Theory in Machine Learning for Social Science', *Frontiers in Big Data* **3**, 18. https://doi.org/10.3389/fdata.2020.00018

Ratcliff, C. and Sun, Y. 2020 'Overcoming Resistance Through Narratives: Findings from a Meta-Analytic Review', *Human Communication Research* **46**. https://doi.org/10.1093/hcr/hqz017

Ren, Y., Parvizi, A., Mellish, C., Pan, J. Z., van Deemter, K. and Stevens, R. 2014 'Towards Competency Question-Driven Ontology Authoring', in V. Presutti, C. d'Amato, F. Gandon, M. d'Aquin, S. Staab and A. Tordai (eds) *The Semantic Web: Trends and Challenge*, Springer International Publishing. 752–767. https://doi.org/10.1007/978-3-319-07443-6_50

Seaton, K.-L. and Lauzikas, R. 2023 'Digital Archiving Best Practice and Future Challenges', *Internet Archaeology* **63**. https://doi.org/10.11141/ia.63.8

Swenson, E. and Cipolla, C. N. 2020 'Representation and materiality in archaeology: A semiotic reconciliation', *World Archaeology* **52**(3), 313–329. https://doi.org/10.1080/00438243.2021.1925582

Terrell, J. 1990 'Storytelling and Prehistory', *Archaeological Method and Theory* **2**, 1–29. https://www.jstor.org/stable/20170203 [Last accessed: 2 Feb 2026]

Thomas, J. T. 2015 'The Archaeologist as Writer', in R. M. Van Dyke and R. Bernbeck (eds) *Subjects and Narratives in Archaeology*, University Press of Colorado. 169–187. https://doi.org/10.5876/9781607323815.c008

Tringham, R. 2020 'Closely Observed Layers: Storytelling and the Heart' in K. Supernant, J. E. Baxter, N. Lyons, and S. Atalay (eds) *Archaeologies of the Heart*, Springer International Publishing. 239–252. https://doi.org/10.1007/978-3-030-36350-5_15

van Erp, M., Tullett, W., Christlein, V., Ehrhart, T., Hürriyetoğlu, A., Leemans, I., Lisena, P., Menini, S., Schwabe, D., Tonelli, S., Troncy, R. and Zinnen, M. 2023 'More than the Name of

the Rose: How to Make Computers Read, See and Organize Smells', *The American Historical Review* **128**(1), 335–369. https://doi.org/10.1093/ahr/rhad141

Whitelaw, M. 2015 'Generous Interfaces for Digital Cultural Collections', *Digital Humanities Quarterly* **9**(1). https://dhq.digitalhumanities.org/vol/9/1/000205/000205.html [Last accessed: 2 Feb 2026]

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. 2016 'The FAIR Guiding Principles for scientific data management and stewardship', *Scientific Data* **3**(1), 160018. https://doi.org/10.1038/sdata.2016.18

Wright, H. 2021 'Archaeology and the FAIR Principles. A case study from the Archaeology Data Service ', [Slideshow presentation] Archaeology Data Service. https://archaeologydataservice.ac.uk/app/uploads/2022/08/SSHOC_ArchaeologyFAIR.pdf [Last accessed: 2 Feb 2026]

Wylie, A. 2017 'How Archaeological Evidence Bites Back: Strategies for Putting Old Data to Work in New Ways', *Science, Technology and Human Values* **42**(2), 203–225.

Yakel, E., Faniel, I. M. and Robert Jr, L. P. 2024 'An empirical examination of data reuser trust in a digital repository', *Journal of the Association for Information Science and Technology* **75**(8), 898–915. https://doi.org/10.1002/asi.24933

Yarrow, T. 2008 'In Context: Meaning, Materiality and Agency in the Process of Archaeological Recording' in C. Knappett and L. Malafouris (eds) *Material Agency: Towards a Non-Anthropocentric Approach*, Springer US. 121–137. https://doi.org/10.1007/978-0-387-74711-8_7