Data-Informed Tools for Archaeological Reflexivity: Examining the substance of bone through a meta-analysis of academic texts

Sarah E. Jackson, Caleigh E. Richissin, Erin E. McCabe and James Jaehoon Lee
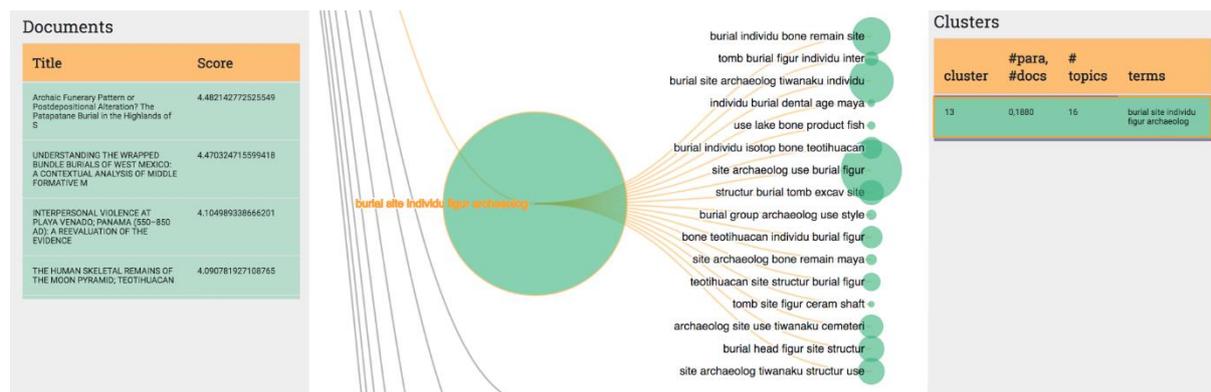
A topic-cluster with 16 subtopics, containing examples of broad and general archaeological language

Our study uses computational archaeology tools to investigate how researchers in our field present interpretations of the past in patterned ways. We do so in order to illuminate assumptions, naturalised categories, and patterned interpretative moves that may direct or impact the ways we interact with our evidence and write about our research. We approach this topic through a meta-analysis, using large-scale textual data from archaeological publications, focusing on the case study of bone. Are there patterned ways that archaeologists write about artefacts like bone that are visible when analysing larger datasets? If so, what underlying ideas shape these shared discursive moves? We present the results of three analyses: textual groundwork, conducted manually by field experts, and two machine-based interactive topic modelling visualisations (pyLDAvis and a hierarchical tree based on a Model of Models). Our results indicate that there are, indeed, patterns in our writing around how artefactual and archaeological materials are

discussed, many of which are overt and sensical. However, our analyses also identify patterned discourses that are less obvious, but still part of regularised discourses in written narratives surrounding bone. These include: the use of multiple conceptual positions within, rather than simply between, articles, and a lack of patterned centrality of indigenous ontologies in how our field writes about bone. This pilot approach identifies data-informed, applied tools that will aid reflexive practices in our field. These operate at a scale that impacts future scholarly interactions with both evidence and published interpretations by shifting observation and reflection from an individual or small group exercise to a larger and more systematic process.

1. Introduction

Archaeologists are trained to look carefully, and are increasingly sensitive to the ways that such looking is channelled, with resulting impacts on focus and interpretation. This may happen, for instance, through professional training (e.g. Goodwin 1994) or through local, cultural categories (e.g. Boivin and Owoc 2004; Henare *et al.* 2007; Meskell and Joyce 2003). An additional important element in how our analytical attention is directed relates to scale (Carr and Lampert 2016b). The scale we adopt in any research endeavour influences which perspectives are seen as appropriate, and also leads to the creation of boundaries around the analytical units that are perceived as meaningful at that scale (Carr and Lampert 2016a, 9–11). Adopting different scales of analysis may reveal blindspots that were not visible from a previous scalar perspective. Such repositioning can be especially effective in strengthening disciplinary reflexivity, which we invoke in the present work in a specific sense, as related to efforts to examine and critically understand the practices and structures of our field. Taking a larger perspective on archaeological analyses, as is now possible with computational archaeology and the use of large datasets, opens opportunities for better understanding how researchers in our field present interpretations of the past in patterned ways (Huggett 2013).

The computational archaeological methods we adopt build on important previous work in this area, including digital analyses, to better understand the structure and reasoning of our field (see, for instance, work by Aussenac-Gilles 2006; Gardin and Roux 2004; Morgan and Eve 2012). Like Gardin (1980), for instance, we are interested in observing and contextualising analytical practices by studying the structured elements of published research products, i.e., the units of language embedded in a given article, across a large textual corpus. However, the reflexivity afforded by the methods described in the present article does not have a typological aim of classifying research, nor does it propose a more explicit, formalised structure of archaeological analysis and its publications (Dallas 2016, 307, 311). Rather, our goal is to learn more about disciplinary ideas that structure how we write about our findings. To this end, in contrast with studies that have focused on earlier points in the research process (e.g. Marwick 2016; Vlachidis and Tudhope 2016), in our study we analyse journal articles in order to identify patterns in research writing that become visible from a broader scale.

We approach this topic through a meta-analysis, in which we use large-scale textual data from archaeological publications to identify patterns in language used. That is, our study conducts a meta-analysis of text (specifically, 599 journal articles) as data. Our

research is focused on bringing to light assumptions, naturalised categories, and patterned interpretative moves that may direct or impact the ways we interact with our evidence and talk about our research. As a starting point, our project asks: are there patterned ways that archaeologists write about artefacts that become visible when analysing larger datasets of writing? If so, what underlying ideas shape these shared discursive moves within written academic narratives? Our understanding of discourse (drawing on valences of the term that reference Natural Language Processing – discussed further below -- as much as Foucault; Lee and Beckelhimer 2020, 112-116) focuses on how we – archaeologists – write about archaeological material. This intersects with narrative processes such as 'the construction of subjectivities and experientiality through stories' (De Fina and Georgakopoulou 2015, 2), practices that can be recognised even within academic writing. Because such patterned ways of thinking and writing are often ambiguous and hard to recognise at a fine-grained scale, whether in our own work or the work of others, we approach this work through machine-learning analysis of large datasets, supplemented by human-directed textual groundwork (Lee *et al.* 2018).

Our research team began with intentionally narrow test cases in order to refine our methods and explore the utility of this approach; as detailed at the end of this article, we anticipate next steps in our study that will broaden our scope in terms of datasets used and resulting commentary. Given the areas of expertise of the archaeological members of our research team (Jackson and Richissin), we began with a focus on contexts from the ancient Americas, drawing on data from two prominent Latin American archaeology journals, *Latin American Antiquity* and *Ancient Mesoamerica*. Our preliminary work analysed multiple artefact categories (ceramic, lithics, obsidian, jade, and bone). In this article, we present the results of analysing the category of 'bone' holistically (identified analytically by the word 'bone', and thus including both human and animal), in order to demonstrate the possibilities of this approach. In what follows, we present the results of the three types of analyses that we carried out in order to investigate the patterns that emerge in how archaeologists write about archaeological materials. These three analyses include: 1. textual groundwork, conducted manually by field experts (Jackson and Richissin), followed by two machine-based interactive topic modelling visualisations (2. pyLDAvis and 3. a hierarchical tree based on a model of multiple models). These will be explained in greater depth below.

Based on our findings from these analyses, we argue that there are, indeed, patterns in our writing around how artefactual and archaeological materials are discussed, an idea that we support through our findings related to the specific substance of bone. Many of the patterns that we identify through our data analyses are overt and sensical – that is, they fit closely with particular methods, are unsurprising to specialists in the field, and at times echo explicitly articulated rationales. In this way, our case study confirms the power of this approach – that the machine methods used can identify themes that topic experts can confirm, even at this larger scale. However, our analyses make contributions by also identifying patterned discourses that indicate underlying assumptions, revealing less directly expressed ideas that appear as part of a regularised discourse in written narratives surrounding bone. Significantly, that means that this approach brings to light patterns that are visible only when we adopt a different analytical scale and way of organising our data.

While we believe that our specific observations about bone will be of interest to Latin American archaeologists, our primary contribution in this article is to demonstrate important interpretative, and reflexive, possibilities that emerge with a change in scale.

Archaeologists – and scientists generally – already know that when we re-engage with data in new ways, we are often able to learn something new (e.g. see recent work by Tshitoyan *et al.* 2019); our study demonstrates the productive possibilities for engaging with data – in this case, published texts – at a late stage in the research process. That is, completed research can be the starting point for new insights into some of the structured understandings that underpin research in our field. Bringing them to light allows them to be acknowledged and potentially queried, rather than remaining latent or invisible (akin to Miller's idea about 'the humility of things' that are not easily seen (1987, 85–108)). These are, perhaps, large claims, but through using the circumscribed, specific substance of bone, the approach we pilot identifies data-informed, applied tools to aid reflexive practices in our field. Significantly, these operate at a scale that shifts observation and reflection from an individual or small group exercise to a larger and more systematic process.

2. Contextualising Machine-Learning Analyses within Anthropological and Archaeological Frameworks

While we see the contribution of our study as primarily methodological, the research we describe intersects closely with multiple current theoretical discussions in our field. We contextualise our approach within larger anthropological and archaeological conversations in order to frame its impact in conceptual rather than mechanical terms.

2.1 Looking, relating

Discussion of interpretative attention in archaeology has engaged with situated and contextually specific ways of looking, and thus knowing. We can think here of Charles Goodwin's work on looking and expertise (1994) (see also Grasseni 2007; Jay and Ramaswamy 2014; Mirzoeff 2011 for related work on visuality and interpretation), including specific discussion of the work of archaeologists. His research emphasises that particular practices make real or 'highlight' certain elements of the archaeological record (for instance, a post mold/posthole) (Goodwin 1994, 610–11), while – presumably – other elements of the archaeological record remain unmarked and unrecognised, for a variety of reasons. Goodwin's work is useful in sharpening awareness of our own normalised attentions. In this way, it intersects meaningfully with other work on multiple types of culturally specific understandings of the material world (e.g. Boivin and Owoc 2004; Jones 2012; Meskell 2004; Tilley 2004), some of which differ notably from the assumptions that modern, western archaeologists bring to our research. Through all these discussions, we are reminded of the multiple perspectives possible in relation to data and to the research process. Goodwin's work with archaeologists has also emphasised the role of talk and narrative in the process of making archaeological evidence visible; he discusses interactive narratives as a key element (Goodwin 2015). The patterns we identify in our study might be understood as the outcomes of such interactive narratives, as recorded in text.

Invoking the frame of 'looking' is not meant to singularly foreground the sense of sight and is instead a shorthand for 'directed attention'. This attention is important to our current project because we are considering the insights possible when our attention is shifted. While 'looking' or 'attention' imply uni-directional action upon the archaeological record and the data we record, we might productively think about 'relating' to data (of various types) and the research process. This opens up consideration of mutual influence and the process of accommodation by which we work with data and interpretations to reach conclusions. Jackson ([2014]) has elsewhere framed this process of relating to our data as dual processes of 'domestication' and 'liberation', representing different types of relationships with our data. These characterisations acknowledge that archaeological datasets (of a variety of types), especially as they grow increasingly large, may be subject to diametrically opposed processes that involve, on the one hand, the taming of our data to be more orderly, categorised, and named, while, on the other hand, engaging with the possibilities inherent when we make interpretative space for dynamism, multiplicity, and contradiction (e.g. see discussion of these topics in relation to big data in Cooper and Green [2016]). These different responses to our archaeological data – ones that are not mutually exclusive – indicate different expectations that we, as researchers, have for the outcomes of data analysis, and for how we present and frame our interpretations. This helps expand our critical thinking beyond just how we 'see' data, but also how we interact within it. Our project allows us to 'see' our data differently through the computational analyses we conduct, pointing towards categorisations and organisations that would not readily have been apparent to us without this method. We may not always be explicitly aware of these expectations, nor of our accompanying interpretative and linguistic or discursive moves; our study contributes to illuminating shared, but potentially unvoiced, relational stances to archaeological materials and interpretations.

2.2 Shifting scale

In our study, we refocus attention on archaeological discourses, using published texts as data, with particular interest in questions of scale. The scale of analysis we adopt relates not only to how our attention is directed, but also the perspectives we acknowledge as meaningful and/or appropriate. Researchers draw boundaries according to scale, similar to the constructive work of categorisation (as occurs through practices such as typification, organisation, and documentation – e.g. Berggren and Hodder [2003]; Bowker and Star [1999]; Cobb *et al.* [2012]). Indeed, Carr and Lempert argue that scale is not simply about perspective, but rather accomplishes related types of work, including orientation and valuation ([2016a], 9). Scales can feel naturalised, and institutionalised, in ways that make them hard to question or shift. In this way, scaling can obscure certain things – positioned perspectives leading to omissions – indicating that different scales are necessary for allowing different data and interpretations to emerge. Current discussions emphasise the need to acknowledge multiple ontologies (e.g. Alberti and Bray [2009]; Astor-Aguilera [2010]; Viveiros de Castro [1998]), in part because different questions and concerns will emerge as salient via different modes of engagement with the world; so should our analytical engagements show movement across different scales in order to identify productive areas of deeper questioning, as well as unexamined assumptions. Scale matters in important ways; analyses of larger datasets related to our disciplinary work offer opportunities for structured reflection into practices and patterns in archaeology.

In the present study, we shift scale by looking at a large set of published articles, drawn from two journals. While the corpus of 599 articles we use in our studies might not technically qualify as 'big data', our study is contextualised within increasing engagement with large datasets as technological advancements continue and data storage becomes increasingly low cost. Utilisation of large datasets has impacted multiple areas of our field, including the centralisation and accessibility of information (for instance, The Alexandria Archive Institute, focused on the Open Web); the possibility of integrating multiple datasets, and thus extending them through comparison and linkage (for instance, the cyberSW initiative at the University of Arizona; Blue and Communications 2017); and the scale of enquiries possible (e.g. many spatial initiatives such as LiDAR and GIS – for instance, Kozak and LaClair 2012). The data that we, as archaeologists, look at in aggregate vary in the type of information encoded, but one large subset is formed by primary source documents of various types: these include texts and images created by ancient users (such as those available through the Maya Vase Database (Kerr n.d.) or the Corpus of Maya Hieroglyphic Inscriptions Project at the Peabody Museum of Archaeology and Ethnology), or the increasingly less-grey literature of our own site reports, made progressively more available in both informal and structured sharing arrangements. Big data studies looking at archaeological data allow for comparative insights at a broader scale and may include analysis of archaeological knowledge practices (e.g. Cooper and Green 2016; Dallas 2016; Gattiglia 2015). Our study, which takes a meta-analytical approach to academic texts, looks not at data constitution or interpretative reasoning, but rather at the end product of the research process. By focusing on published journal articles, we are able to identify potential commonalities in how interpretations are presented in academic writing about particular artefactual materials.

2.3 Computing and ontologies

Our analyses are made possible through computational archaeology and the possibility of machine analysis of larger datasets. As alluded to above, others have undertaken related enquiries that use computer analyses to analyse archaeological data at larger scales (e.g. Gardin and Roux 2004; Huvila *et al.* 2017; Jørgensen 2015; Marriner 2009; Roux and Aussenac-Gilles 2013; Schmidt and Marwick 2020). We share with some of these previous enquiries not only the use of computer-driven analysis, but also a desire to make more explicit elements of archaeological reasoning, as well as better awareness of the state of our academic field. We take inspiration from Gardin's efforts to examine archaeological argumentation (1980), characterised by Dallas as 'a retrospective method of inquiry' (2016, 314) and framed as a digitally reflexive archaeology (2016, 316). Gardin (1980), and Dallas in his extension of Gardin's work (2016), think carefully about how digital methods can provide 'condensation' of archaeological analyses. By condensing research, Gardin aims to detect its essential analytical elements so that archaeology as a field might implement stricter schematisation, reducing the influence of rhetorical elements in published work (Dallas 2016, 314–15). Gardin's primary concern with scale, then, is ultimately a question of analysing sufficient text data to create and justify these potential schematisations. This approach has been extended in intriguing ways by Roux and colleagues, in their work on constructing knowledge bases from documents structured automatically with inference rules in the Arkeotek Project (e.g. Roux and Aussenac-Gilles 2013, 267, 270).

Like Gardin and Roux, we examine semantic relationships as a means of observing an ontology of analytical processes. However, our perspective seeks to include rhetorical

elements, more visibly patterned at the large scale, so that we can illustrate what archaeological discourses, latent or categorical, exist and so that we can begin to investigate their influence. To accomplish this semantic work, we derive meaning from computational word groupings (topic models). In this way, the relationship between the organisation and the meaning of language units, i.e. syntax and semantics, is blended more closely together. While our examinations focus on a different, later point in the interpretative process, both the earlier studies and our own indicate underlying structures discoverable via reorganisation of data (Dallas 2016, 317), a promising way to move towards data reuse and explicit disciplinary awareness.

As part of a diverse field of computational archaeology research, we see ourselves as interested in investigating ontologies, acknowledging the dual definitions (computational and anthropological) of that term. In a computational sense, our analyses are rooted in structured organisation of concepts that emerge from our textual corpus via machine learning. Computationally, we are defining ontologies by linguistic patterns and word use. Our second commitment to ontological enquiry is turned inwards, asking about the realities and ways of being we embody as anthropologists and archaeologists, and how we relate to the materials we encounter (Pike and Gahegan 2007). While we acknowledge that this may be a stretch of how ontology is typically conceptualised (and applied) anthropologically, we hope to be productively playful in suggesting that our discipline can examine not only structured practices or biases, but also inculcated and naturalised understandings of how best to interact with our materials and how to convey their meanings. In both of these ontological framings, we are specifically interested in a reflexive move made possible through the identification of larger-scale patterning. The reflexivity we explore does not focus on important elements of identity differences and representational challenges (see, for instance, Bardolph 2014; Conkey 2007 for examples of how these topics can be meaningfully integrated with data-informed inquiries), critical topics in our field but beyond the framing of this study. We are particularly interested in the written discourses that archaeologists use in their work because, in contrast to the tightly structured format of site paperwork or excavation reports, journal publications provide a place with some latitude in expression and may be a context in which authors provide textual indications of their relationship to and/or assumptions regarding archaeological materials.

In what follows, we discuss three textual analyses (one conducted by researchers and two by machine) that examine patterning of archaeological writing in Latin American archaeology publications, specifically regarding the substance of bone. Our contribution is to identify that these analyses, looking at archaeological publications at a large scale, illuminate both explicit and implicit structured archaeological understandings about certain materials (how they are perceived and how they are related to) that may not be visible without this particular scalar perspective.

3. Methods

3.1 Human analysis

3.1.1 Analysis 1: Textual groundwork

Our first step in this study was to carry out textual groundwork, in which we identified themes across articles apparent to human subject experts, drawing upon our perceptions and understandings as archaeologists. This initial analysis was intended to establish baseline expectations for the topic (that is, what are some of the ways that experts in our field talk about particular materials?), with grounding in published materials.

In order to accomplish this, we established the following workflow in order to locate articles for this part of the process. We used the article index Anthropology Plus, as a widely respected academic database that was also accessible through our university. Our initial search terms consisted of 'Maya' and '[artifact type]' in 'all text fields' search boxes (for the results presented below, this was 'bone'), with date range left open. The search results were exported into a spreadsheet, with the results subsequently narrowed using the following criteria. Articles were removed from the list if they were not about Maya archaeology; if they focused on Postclassic, Colonial, or modern time periods, were in Spanish, or showed up in the list more than once. For all of the artefact types we examined, we aimed to include 20 articles in this opening examination, and we agreed upon structured guidelines for how to choose or augment the list to reach 20 articles. These decisions were intended to narrow the article list to an approachable number for a subjective human gaze, while aligning with our own specific expertise in Classic Maya archaeology; English articles were privileged owing to the language analysis that forms the heart of this inquiry.

As noted, this initial textual groundwork was intended to establish a baseline expectation for the topic by identifying overt themes apparent to knowledgeable readers. It is inherently impressionistic and must be augmented by the more systematic approaches that we take in subsequent sections of this article.

3.2 Machine-based analyses

3.2.1 Corpus of articles for machine-based analyses

For the computational processes of this research project, we collected over 2300 articles of unstructured text from the journals *Ancient Mesoamerica and Latin American Antiquity* to compose our corpus. This range comes from their digitised publication history between 1990 and 2018. This scope of publication may suggest other considerations in the corpus's characteristics, such as changes over time, but at present our research interests aim to incorporate consolidated publication characteristics and represent a discipline-holistic view from a wider lens. Narrowing this further to focus on articles containing the term 'bone', as well as removing a small subset of Spanish language content, results in an operational corpus of 599 articles, or documents (Figure 1). The metadata for these articles is available as part of this article's accompanying data found at the cited DOI. To preprocess the dataset of articles we used several

functions from the python library, Natural Language Tool Kit (NLTK). We applied NLTK's stop words function to remove basic English stop words from being processed by the algorithm. From the same library, we then 'tokenise' our dataset to allow the algorithm to read each document term by term. Lastly, we stem these terms using NLTK's Porter Stemmer package so that morphological variants of a word can be interpreted together, e.g. bone/bones.
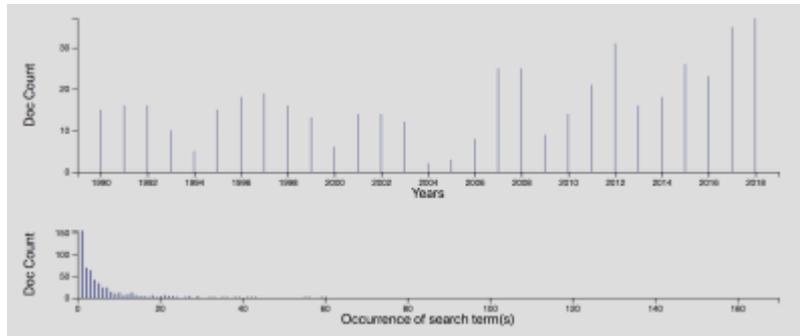

Figure 1: The top chart depicts the number of articles ('documents') over time that include the term 'bone'. The bottom chart shows the number of documents in relation to the number of times the search term 'bone' appears in those articles. 44.07% of these documents contain the search term 'bone' 5 or more times; 51.25% of documents mention bone 4+ times; 26.04% of documents mention bone 10+ times.

3.2.2 Topic modelling: the underlying machine-learning technique

Before describing the analysis of our two machine-learning based visualisations, we explain the algorithm that underlies them. Within the broader domain of Machine Learning, topic modelling is considered a task of Natural Language Processing (NLP), as it is concerned with computer processes that facilitate an understanding of large datasets of natural language. These algorithms provide a more accessible approach to a corpus of this size and assist in unearthing latent themes and structures in the collection of 'bone' documents. Specifically, we applied Latent Dirichlet Allocation (LDA), a type of statistical topic modelling that generates models based on the probability of latent topics in a document and the probability of words in those topics (Blei 2003; Mabey 2015). Here, the LDA sense of the word 'topic' refers to a collection of words with significant probability of co-occurring in a document. For example, one topic for the 'bone' corpus has listed high probabilities of the terms, *Bone/Isotope/Strontium/Enamel*. Whereas one topic might feature the term 'isotope' with high probability, its appearance in another topic might not be as strong or even have negligible probability. Similarly, one document might have a high presence of the topic containing *Bone/Isotope/Strontium/Enamel*, whereas that same topic features only faintly in another document. To accomplish the combination of terms into topics, the LDA topic modelling algorithm automatically detects the occurrence of every term in every document and can then compute the probability of terms co-occurring across all documents of the corpus. In brief, the power of topic modelling is that it allows us to understand how language works in the designated corpus at the level of the word and also at the level of the document.

Surveying a corpus of almost 600 full length academic articles to ascertain their collective topics is undoubtedly an immense task for any researcher. However, the primary strength of LDA topic modelling performed on these archaeology articles is not

simply a question of scale. Since this process relies on term occurrences in the corpus, resulting patterns are ascertained directly from the texts themselves, extracting quantifiable topic groups rather than applying a subjective, human grouping of vocabulary structures. Therefore, researchers can observe language structures beyond established or expected topics, to topics shaping more latent or implicit discourses, as the name, Latent Dirichlet Allocation, itself suggests (see similar applications in literary studies in work by Goldstone and Underwood 2014).

Relative to other forms of unstructured text, academic articles may already be understood as structured by somewhat explicit, disciplinary topics and are often even assigned keywords. However, algorithms charged with uncovering quantifiable, latent patterns offer a compelling means of reflecting on those structures. With digitised corpora now readily available, it is not uncommon in these reflective studies to create a corpora of texts from a prominent journal(s) in the field to serve as discipline proxy in order to provide a survey of a field (Martin 2019; Peirson *et al*. 2017) or specific, integral terminology (Cohen Priva and Austerweil 2015).

In addition to unearthing latent topics, the application of LDA topic modelling provides quantifiable results that can then guide researchers' interpretation of topics' relationships. Manually assessing relationships between topics significantly amplifies the magnitude but, more importantly, the subjective nature of a research question grounded in latent/implicit discourses. It is this second activity of interpreting the machine-based topic relationships within 'bone' articles where the algorithmic approach proves even more valuable. Distance between topics is reflective of those topics' term distribution similarity. These inter-topic relationships can be assessed in finer detail through asking: 1. What term(s) are shared between topics? 2. How does a term's probability vary between topics? 3. What different terms act as neighbours to a term across topics? In studying the topic modelling results of 'bone' articles with these questions in mind, we begin to shift towards more contextual interpretation, allowing us to construct an understanding of the corpus's discursive modes.

On our analysis platform, the LDA modelling algorithm passes through every word in the corpus of bone-related documents twenty times; in other words, we train our statistical model on the entire corpus twenty times to learn the corpus's structures thoroughly. Following extensive calibration, we have found that this number of passes, or 'reads', creates ample stability in which the algorithm is no longer acquiring new information.

Research informed by large-scale data relies on these computational methods; however, it also requires efficient visualisation to render human interpretability possible. The visualisations discussed below have an additional element of dynamic interactivity, allowing for more nuanced explorations. The visualisation of these LDA topic models facilitates exploration more conducive to research.

3.2.3 Analysis 2: pyLDAvis

One such LDA topic model visualisation approach on our analysis platform is an adaptation of the Python programming package, pyLDAvis (Mabey 2015). This is the second type of analysis that we undertook (and the first machine-based one). With our adaptation of this package, the topics (terms grouped together based on significantly probable co-occurrence in a document) are projected onto a two-dimensional plane, using Principle Component Analysis (PCA) projection to create a view of the relationship

between topics while also providing the ability to explore the terms in those topics with the terms' relevancy bar chart (Figure 2).
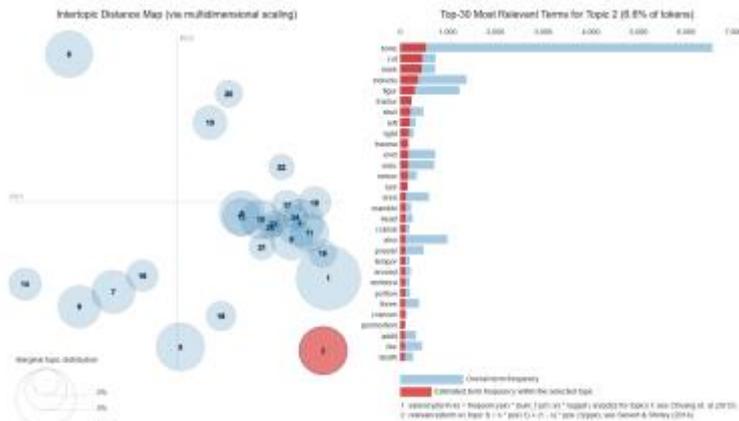

Figure 2: Complete view of the visualisation adapted from the Python programming package, pyLDAvis (Mabey 2015). Here, the highlighted Topic 2 of 25 is the second most prevalent topic with 6.6% of tokens. This topic contains terms pertaining to partibility such as skull, mandible, and portion. Terms like 'death' and 'trauma' indicate this topic's overlap with a more cultural consideration of bone.

In the two-dimensional plane section of this visualisation, each circle represents a topic. The size of each circle is determined by its prevalence in our corpus and the numbers reflect that prevalence in descending order, where Topic 1 is the most prevalent. This visualisation also takes into account the distance between topics as determined by the similarity of topics' distributions of terms. The terms' bar chart section of this visualisation allows for closer inspection of a topic. When a specific topic circle is selected from the two-dimensional plane, its corresponding list of most relevant terms is displayed in relation to the relevancy of those terms in the overall corpus. In this way, we can interpret the topic's meaning as well as the relationship of that topic's most relevant terms to the broader corpus (Sievert and Shirley 2014).

Of the corpus's 25 topics, 14 are notably arranged very tightly; in other words, they have little inter-topic distance. However, these topics represent a range of prevalence relative to the complete corpus (as depicted by the area of each topic circle), with topics comprising anywhere from 11.7% to 2.1% of the corpus's total tokens. (A token is created for every word appearing in the corpus, following the stemming process and removal of stop-words such as prepositions and articles.) The relative prevalence of the topics in the corpus provide starting points for identifying patterns in writing about archaeological bone that form one of our central research questions. For example, shifting our gaze outside the tightly arranged group of 14 topics in Figure 2, we notice another group of topics formed by an overlap between the circles representing Topics 8, 7, and 16. The relative prevalence of this group's topic circles (denoted by the areas) suggested that these are not simply outlier topics. These topics – which are characterised by words like isotope, collagen, enamel, and carbon – share some common language among themselves, while also representing a distinctive methodological approach to bone that explains their distance from the central grouping of topics. Does a close vocabulary neighbour reflect a similarly close relationship in theory, or are there surprising relationships? On the other hand, why might topics

outside this neighbourhood utilise such different vocabulary in context with the same material?

3.3.4 Analysis 3: a multilevel model of models

LDA topic models likewise form the basis of the platform's Multilevel Model of Models visualisation, which is our third type of analysis (See Figure 5). Instead of running a single model from the corpus, as pyLDAvis does, this visualisation option runs six models, then groups the resulting 25 topics from each of the six models into topic-clusters. The clustering method used here is an affinity propagation on a vector space created from the top 25 words in each topic and their probability scores. This method takes the measure of similarity between pairs of data points as input. Affinity propagation then decides on the number of topic-clusters based on the data provided rather than using researcher input for this parameter. In generating and clustering these additional models, the resulting visualisation can offer more stability in assessing probability distributions as well as establish a hierarchy of topic to topic-clusters. Within this tree visualisation hierarchy, individual terms still serve as components of topics (as in the pyLDAvis) but topics, in turn, are branches of broader topic-clusters. In other words, topics are collections of terms with significant probability of co-occurring in a document and topic-clusters are groups of similar topics. We tested our 'bone' model's validity at various topic counts before determining that 25 topics per model allowed for the most meaningful results for our research direction. This evaluation relied both on our understanding of model structures, as well as humanistic interpretation provided by subject matter experts assessing topic relationships and utilising the platform's document viewer to further explore how reasonable those relationships might be. For instance, creating a multi-level model using 35 topics, we found the corpus being assigned into topic-clusters that were too broad, attempting to fit disparate, specific topics into one overarching grouping. This model resulted in 33 topic-clusters, one of which was an extremely broad topic-cluster of 72 topics (a topic-cluster that includes, for example, specific but disparate topics like 'ring tree chronolog sampl quemada' and 'codex hill place mixtec vienna'), and a long tail of 14 different single topic topic-clusters (e.g. 'vessel orang ceram holmul protoclass' and 'glyph emblem name titl ajaw'). On the other hand, at a division of 15 topics, the model provided insufficiently precise insight, with vocabulary endemic to broader archaeological research work (e.g. archaeology, site, maya) found pervasively throughout the model's topic-clusters, while only two smaller, more specific topic-clusters managed to surface at this scale. While further exploration in specific topic count selection for this and other language dataset's models could prove interesting, our analysis found that the 25 topic count model delineated topics and topic-clusters at the most successful level of specificity, given a grounding in subject matter expertise for both archaeology concepts and topic model validity.

This visualisation also allows for the interactive exploration of this 'tree' hierarchy. Alongside the 'tree' visualisation is an interactive, descriptive chart of the topic-clusters' metadata, including a count of topics within each topic-cluster and a sum of the documents that every topic within that topic-cluster appears in. The number of topics contained in a topic-cluster can serve as a measure of that topic-cluster's range, or the diversity of its vocabulary's application across this corpus. In knowing the sum of documents that a particular topic-cluster's topics appears in, we gain insight into that topic-cluster's weight within the corpus. Both of these metrics are useful in examining a particular topic-cluster, as well as in understanding a topic-cluster's place in relationship with the complete corpus. Interacting with either of these components of the visualisation will then populate a portion of the platform that allows for a view of a document's text,

allowing for a direct examination of specific articles corresponding to that topic, or topic-cluster, ordered by descending relevancy.

The 150 topics generated from this visualisation's combined six models (of 25 topics each) arranged themselves into 21 topic-clusters. The number of topics per topic-cluster ranged from 28, in the model's most broadly applied topic-cluster, to a trailing four topic-clusters that featured only one topic each. Notice the contrast, for instance, between the large Topic-Cluster 13, which contains 16 topics, representing general language related to bone and associated analytical discussions (including burials) and Topic-Cluster 4, which is much smaller with only two topics representing a narrow focus related to burials specifically. While there exists a recognisable top and bottom tier in terms of the topic-cluster's number of sub-topics, there is a fairly even spread, with the majority of topic-clusters that are established being in the 5-10 sub-topic range. As with the previous type of analysis, we see emergence of regularised ways of writing about bone, that then require interpretation to identify which patterns may be overt and/or unsurprising, versus those which reveal implicit or underlying understandings about bone. This portion of more even topic-clusters addresses the expectation of certain structured and expected elements of archaeological enquiry or academic writing modes, while directing attention to the deviations.

4. Bone Textual Groundwork: Emergent Themes in How Bone is Discussed

As discussed above, our first analysis employed a 'textual groundwork' approach in order to identify themes across articles that were apparent to us as experts in the field, in a human-scaled and human-directed analysis. Within the category of 'bone', we noted several thematic ways that bone is discussed in academic journal articles. Many of these thematic framings are not unexpected to expert readers, but nonetheless indicate how a particular material is discussed in the service of different topics or arguments. That is, when we look at patterns identifiable to knowledgeable readers, we see emergent differences in thematic treatment of the substance of bone. This analysis is inherently partial and qualitative but is nonetheless carried out in a way that allows for the articulation of some opening observations that are textually grounded, rather than simply the product of the researchers' own experiences or impressions. In this way, it provides an initial baseline for contextualising the two machine-based analyses that follow. In the subsections below, we offer some comments on how archaeologists write about bone as an artefactual category, including what types of written discursive moves we observed in how these materials are contextualised.

4.1 Bone is an extension of the body, used for analytical translation

Not surprisingly, bone is often analysed and discussed as an extension of the body. (As noted above, both human and animal bone are included in our corpus, but human bone was central to many of the themes we identified. Though we noted some overlaps in descriptive information in how human and animal bone are discussed, we overall saw clear categorisation of bone source or type, as well as anticipated differences in questions asked (Emery *et al.* 2000 or Novotny 2014), suggesting distinctions that depend on an understood categorical type.) Many articles report on post-excavation analyses and interpretations of human bone. A wide variety of scientific processes are employed to understand bone within these articles – focusing on topics like diet, age, health, trauma (articles in our textual groundwork corpus – as described above – that illustrate this include Gerry 1997 or White and Schwarcz 1989). All of these topics treat bone as a proxy to convey direct messages about ancient humans, including elements of their identities and experiences. Analytical translation occurs in many articles by using results of bone analysis to talk about larger cultural topics related to human personhood like identity and status (for example, Nystrom *et al.* 2005).

In some cases, attention to the deceased body is manifested via attention to ways bone was housed. In particular, we noted patterns of human bone discussed through the language of burials and tombs. This thematic grouping clearly focuses on context of deposition and discovery for bone. These framings, though dealing with the substance of human bone, were also preoccupied with some similar questions as noted above, such as gender, class, social identity (Wrobel 2007; Batta *et al.* 2013).

4.2 Discussion of bone often engages with fragmentation and parts

Such analyses rely variously on complete bones, pieces of bone, or distilled substances derived from bone, such as collagen (Gerry 1997). The inherent disarticulation, and often partial nature, of excavated bodies, as well as the nature of samples used for some analyses, means that small parts are often used to talk about a whole body, or even whole bodies. That is, these approaches include commentary on specific individuals, but also often larger populations (for example Gerry 1997 versus Wanner *et al.* 2007).

Description of, and attention to, fragmentation was related not only to preservation of extant osteological materials (i.e., osteological materials that are partial as a result of taphonomic processes), but also to form and function (i.e., bodies are analytically segmented). That is, different anatomical parts are categorised and discussed separately, or segments or pieces of bone are used to carry out analysis (e.g. Danforth *et al.* 2009). Many discussions of bone parts (and, by extension, fragmented bodies) are then used to cycle back to categories of imagined wholeness, in which an ancient individual's or demographic group's characteristics or practices were reconstructed, for instance.

4.3 Bone can be a substrate, a constructive material

In contrast to framings of bone that link back to bodies, we noted a distinctively different grouping of articles that focus on artefacts that are made out of bone (e.g. Emery 2009). In these manuscripts, discussion focused on topics related to craft production. Bone may appear as the substance from which tools are made, or as the base material of special objects requiring detailed attention (e.g. carved bones with texts). Some of the language used to describe bone focuses on modification to bone as a substance (e.g. bone as worked, carved, finished, polished). In some of these cases – for instance, objects that are marked with incised texts – the actual material of bone quickly fades from direct attention, as the analytical attention focuses on a different type of information (for instance, that derived from a hieroglyphic text).

In summary, this first analysis indicates some themes related to bone. These are particular to how bone is seen, understood, and framed. These framings in many cases relate to, or are informed by, research questions or methods; nonetheless, these are areas of focus that emerge in terms of the ways in which bone is written about, and different ways we talk about bone. For instance, bone may be addressed as referring to a body or bodies, as partial or complete, as a crafted substance, and/or one that engages in multiple types of translation to analyse and interpret it.

Zooming out, we noticed certain limitations of this 'analog' process. For instance, our observations did not necessarily fit into neat categories, and we were keenly aware of how our own preconceptions shaped decisions at multiple points in the process. Our expertise as trained scholars in this field is important for reading and interpreting sources, but this process is also fraught, especially in terms of human limitations for identifying blindspots, implicit discourses, or less-overt patterns. These initial findings point towards the need for augmenting our own scholarly interpretations with machine-learning approaches that have the potential to capture larger patterns (including unexpected ones) and greater nuance. The machine-based analyses that we discuss next will contrast with the limitations of the corpus we used in our human-directed analytical process, and with the fairly explicit framings we were able to identify.

5. PyLDAvis Analysis: Sharper Divergence of Discourses

In contrast to the textual groundwork described in the previous section, the second analysis we conducted (one of two machine-based approaches that we used) looks at a much larger dataset, consisting of 599 documents that mention or discuss bone. In this sense, there is a greater opportunity to see patterns emerging across a large number of documents. This LDA based modelling is systematic in a different way, relying on machine analysis, which distinguishes it from the previous analysis, which relied on human observational acuity.

Preliminary observations that began to emerge in the textual groundwork section came more sharply into focus through the pyLDAvis analysis (Figure 2), with an apparent division between bone discussed as body or being, versus bone discussed through written discourses that frame it in more material terms, as an object. The congruence between these sections helps to confirm observations that appeared from textual groundwork, while amplifying and clarifying our initial commentary through a larger dataset and more systematic analysis. Additionally, the pyLDAvis analysis adds information on relationships between topics through visual representations of inter-topic distance, as well as topic weight (as represented by size), which allows for inroads in our understanding of patterns that we did not have with the textual groundwork.

5.1 Bone as body

A major theme that emerged from this analysis was the characterisation of bone as related to bodies. Many written discussions of bone in the archaeological literature focus on deceased individuals, excavation of burials, and osteological study of bodies.

Some of the specific ways that bone is discussed has to do with body parts (Figure 2; e.g. see top 30 most relevant terms for Topic 2), with a framing of anatomical structures: bone is discussed in ways that represent the body as partible, and made up of segments (for instance, different skeletal elements). Additionally, bone is characterised as fragmented. These discussions overlap with language that emphasises processes associated with bone and bodies: these include words related to cultural processes evident via bone (e.g. trauma, cause of death, etc.), as well as terms that point explicitly to analytical processes associated with bone samples (e.g. radiocarbon dating).

As part of noticing these analytical framings, we also observe terminology in these same examples that signpost archaeologists analysing the substance of bone in order to understand the lives and experiences of people in the past. Some of this language references specific bone-related substances (e.g. tooth enamel) or particular analytical queries (e.g. isotopic analysis). These terms then overlap with language indicating the ways that these analyses reveal elements of past lifeways (e.g. diet and associated social organisation), and, by extension, aspects of past lived identities (e.g. sex, age).
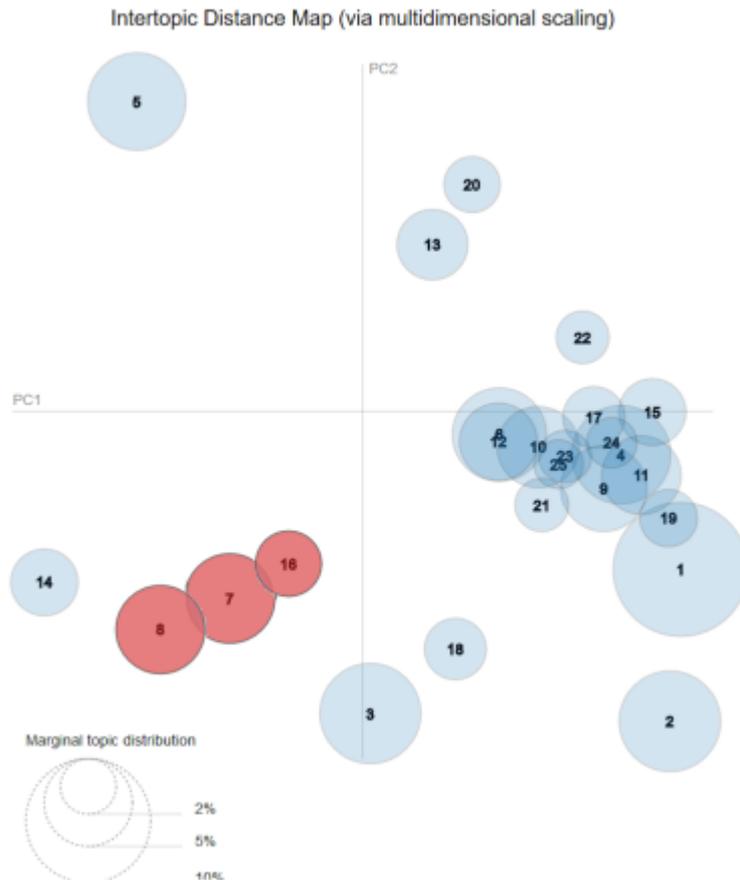
Figure 3: Inset view of Figure 2. Topics 7, 8, and 16 are grouped together though relatively distant from other topics in the corpus. Together they represent 13.1% of the corpus tokens (5.2%, 5.1%, and 2.8% respectively). Terms contributing to the assessment of these topics are isotope, collagen, enamel, and carbon.

Some other terms are also related to analytical processes that indicate increasing distance from the actual substance of bone (Figure 3, detail of Figure 2). Bone transitions from being an archaeological material to being a translated dataset – for instance, data on length or other measurements are extracted from bone and become the focus of discussion. In this way, we see the material of bone transformed throughout the archaeological investigative process and even interpretively fading from view. These points are interesting because they represent engagement with past bodies, but in a way that raises questions – materially, conceptually, relationally – about when bodies cease to be conceptualised as such. Note the relationship in this figure between three topics with close inter-topic distance in the visual model. Their common vocabulary suggests topics related to lab analyses of the biological substances of bone. While they are close neighbours, and large enough not to be considered 'fringe', they are nonetheless far from the tight central topic neighbourhood.

The ways that bone as body are written about also highlight another important category issue. Within our 'bone' pyLDAvis both human and faunal remains are captured. There appears to be both convergence and divergence in terms of how human and non-human bone are discussed. The pyLDAvis suggests that, for instance, some data gathering and analytical techniques are shared across human and animal bone. In contrast, terms that indicate detailed attention to excavation process and context of bone deposition seem to be almost exclusively focused on human bone. Additionally, clues about the

interpretative questions asked of human bone emerge more clearly, implying a divergence in the types of questions being asked across the data corpus. These observations imply some underlying cultural categories or assumptions on our part about appropriate treatment of, interaction with, and associated meanings vis a vis different types of bone.

5.2 Bone as object

We have already indicated that bone, even when understood as part of a past body, can be treated as an artefact. This brings us to discussions that frame bone primarily as an object rather than as a being or a past body.

We see evidence of discussion of artefactual objects made out of bone (e.g. a rasp or a whorl), coupled with related terms in the same topic, such as 'tool', 'product', and 'use'. This is a reminder that in some cases bone is considered primarily as a base material for, or related to, functional objects, with its role as a (formerly) living substance playing a passive or invisible role. Other terminology related to bone as an object indicated that these artefacts are often contextualised alongside other artefactual materials (like ceramics, obsidian, etc.), positioning them as another – and parallel – category of recovered material. Alongside these framings, we see an emphasis on spatial contextualisation, underlining their excavated nature. Intriguingly, we noted clues pointing towards a different framing of bone and making in this model, in which 'making' in relation to bone is not solely functional nor related to inert objects. Rather, we noted the inclusion of the word 'ancestor' in conjunction with other terms related to bone and making, potentially gesturing towards local understandings of how bone constructs non-material, spiritually meaningful entities. That this arises in our PyLDAvis indicates that it is being recognised by some archaeologists writing on this subject, though it does not carry significant weight in the model and is not central to the discourse around this topic.

A specific element of discussions of bone that places it within a material framing focuses on bone as contextualised by burial spaces (Figure 4, detail of Figure 2). Some of the related language involves the characterisation or categorisation of different types of burial contexts. Other common and related language references excavation processes carried out by archaeologists, emphasising the recovery and context of bone. Finally, bone is considered as part of a larger burial assemblage, one that goes beyond a body to include related artefacts and burial goods.
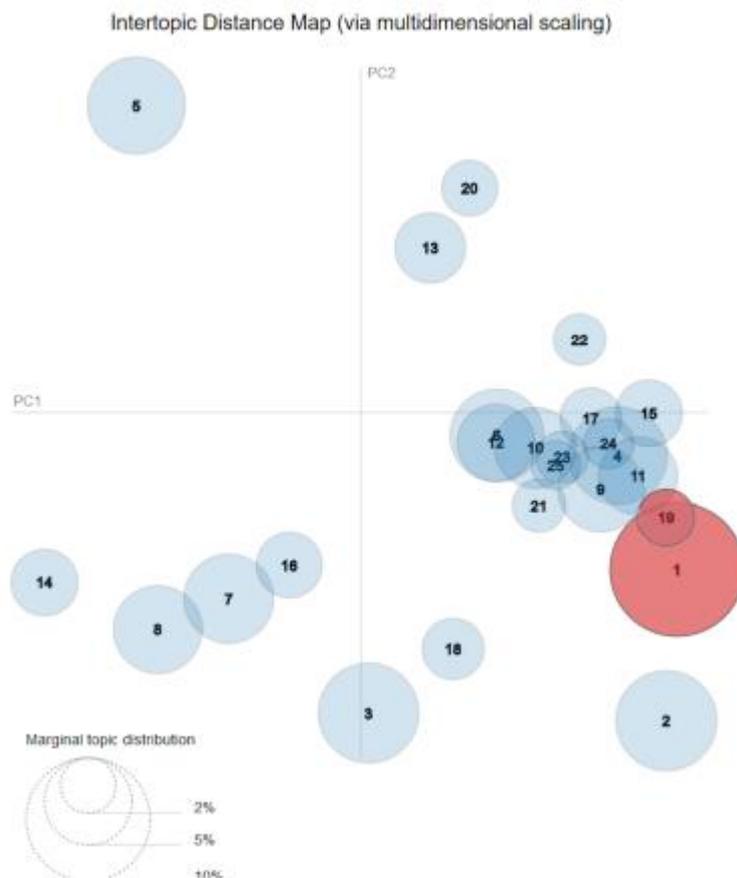
Figure 4: Inset view of Figure 2. Topic 1 is the most prominent topic representing 11.75% of corpus tokens. Topic 19 makes up a much smaller percentage of tokens in the corpus at only 2.1%. Topic 1 contains much broader archaeological language pertaining to burials while Topic 19 contains overlapping language, though this cluster's vocabulary is more culturally explicit, contributing to the difference in token percentage.

As began to emerge in the textual groundwork analysis, we observe here a divergence between bone discussed as evidence of beings and bodies, versus bone as a material that is made into objects. Thus, the results of this analysis support what we observed in our textual groundwork. This congruence between findings of human experts and machine-learning algorithms is significant in that it supports the plausibility of the pyLDAvis results. However, the findings in this section also suggest more complicated language patterns and associated discourses, and open possibilities for visualising vocabulary distributions and relationships within these discourses that were not possible through the human-directed analysis. With the findings from the pyLDAvis analysis, we see initial success in the identification of structured archaeological understandings of bone.

6. Model of Models – Tree: Returning to Documents, Complicating Discourses About Bone

The PyLDA visualisation discussed in the previous section helped to identify emergent themes, in particular dual framing of bone as body versus bone as object. The third type of analysis (and second type of machine-based analysis) we conducted allows us to examine these in a more subtle way. This happens in several ways. First of all, the 'tree visualisation' that we used within the Model of Models is more agile and more clearly shows hierarchical relationships of topic-clusters, topics, and terms (Figure 5). In this way, we can better see the relationships of topics within and between topic-clusters. Secondly, this visualisation allows for textual re-engagement and checking, connecting us back to the documents (articles) themselves (Figure 5). The result in this case was to make the categories we had tentatively identified through pyLDAvis more complex and – somewhat to our surprise – less clearly separated, underlining the importance of re-engagement with the textual sources themselves.
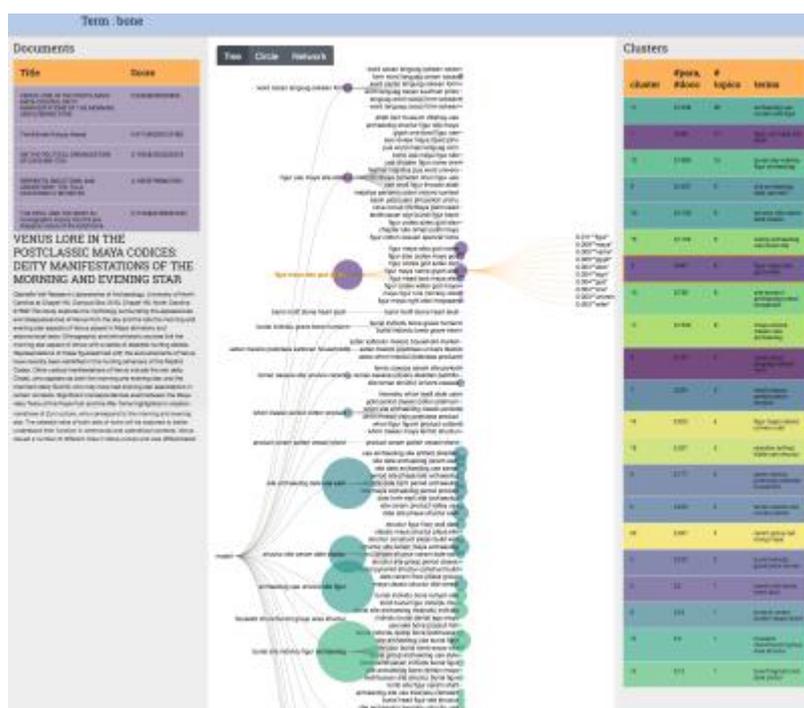
Figure 5: Complete view of the Multilevel Model of Models. This model visualisation displays hierarchical relationships between clusters and topics. This visualisation also allows for textual reengagement with documents as displayed in the document pane. This visualisation highlights clusters 0 and 2 as an example. These two clusters feature bone as a supporting role within discourses more central to other topics.

The topic-clusters most immediately of interest to us were the three that focused most directly on bone (Topic-Clusters 13, 4, and 16); we discuss each of these topic-clusters in turn in what follows. Given the language included in each topic-cluster's terms, we anticipated that these topic-clusters would allow us to better understand central framings

of bone within scholarly discourse. We approached these with our pyLDAvis observations in mind and wondered whether the overall framework of bone as being and bone as object would similarly be represented, or possibly further complicated. We paid particular attention in this analysis to structured archaeological understandings that might not have been evident in the previous analyses we undertook.

6.1 Broad and inclusive discussion of bone as body and object

Topic-Cluster 13 had 16 topics, a notably higher number than 18 of the other 20 topic-clusters, which means that it includes general archaeological language and is broadly inclusive of discussions of bone and related analytical/interpretative discussions and process (Figure 6, detail of Figure 5). The inclusion of all of these topics under the topic-cluster umbrella indicates meaningful relatedness of these discussions. That is, there is a connection between how bone is being discussed across all the topics in this topic-cluster, and likewise the contributing articles. It must be kept in mind that an individual article naturally contains a distribution of topics and will therefore appear across multiple topics in the model, even topics in different topic-clusters. For example, article_a may pertain to topic_a at a probability of 80%, topic_b at 15%, and topic_c at 5%. These topics may even be in distinct topic-clusters. In confirmation of what we had previously observed, we see bone being discussed as both body and object (e.g. Spence and Pereira 2007; Browne *et al*. 1993). Additionally, both human and non-human bone contribute to the documents underpinning this topic-cluster (deFrance *et al*. 2016). In this sense, this topic-cluster nicely confirms our previous observations from the pyLDAvis, but it does not treat them as entirely distinct framings. At a basic level, there is shared language in how archaeologists write about bone, that crosses the distinctions in linguistic orientation and framing that we identified in earlier stages of analysis.
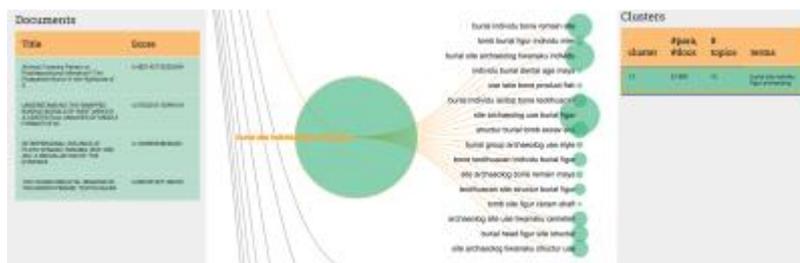

Figure 6: Inset view of Figure 5. Cluster 13 contains 16 subtopics making it one of the more broadly distributed cluster vocabularies, placing it in the top tier of prevalent clusters (third in terms of topic count). Examples of broad and general archaeological language within this cluster's vocabulary include structure, site, and archaeology.

6.2 Streamlined discourses focused on bone as body

Topic-Cluster 4 represents an intriguing contrast to Topic-Cluster 13, in that it includes just two topics, placing it in the trailing quarter of topic-clusters' topic counts, and indicating that it is much narrower in focus and more unified in source language (Figure 7, detail of Figure 5). This topic-cluster focuses exclusively on bone as being or body

and does not represent discussion of bone as object. Subtopics indicate focus on contextual placement of bone in burials (e.g. Rhodes *et al.* 2016), the partibility of bodies (e.g. Geller 2012) and, in one instance, related beliefs or conceptual engagement with bones (Millaire 2004). These topics are also those that we identified in pyLDAvis, confirming those observations. Such tightly wound language indicates the strong relatedness of the articles contributing to these topics and their parent topic-cluster and suggests particularly unified ways of discussing burials and bodies. This is interesting because bodies-as-past-human-beings appears to involve very streamlined and coherent modes of analytical and then linguistic engagement; this may point towards underlying training, approaches, but also beliefs about meaningful and appropriate modes of engagement with bone.

Figure 7: Inset view of Figure 5. Cluster 4 contains only two subtopics making it one of the most narrowly distributed cluster vocabularies. The two subtopics within cluster 4 contain overlapping vocabulary suggesting a more tightly wound discourses related to burials. Contributing terms include burial, individual, and grave.

6.3 Crossing categories: bone as body….or is it object?

The other overtly bone-centric topic-cluster is Topic-Cluster 16, which contains nine topics, placing it between the expansive Topic-Cluster 13 and the narrowly focused Topic-Cluster 4 (Figure 8, detail of Figure 5). At first glance, this topic-cluster seems very unified, focusing on isotopes and diet, discussing both human and animal bones as part of these analyses. Looking at the source documents confirms and supports this perspective (e.g. Andrushko *et al.* 2009; Williams *et al.* 2017). We noticed similar attention to analytical approaches used to reconstruct elements of past human lifeways in the pyLDAvis visualisation. We were intrigued that a topic-cluster characterised by what we would expect to be standardised or aligned discussions of a particular analytic approach was more diffuse than the topic-cluster (Topic-Cluster 4) that addresses bone as body. Digging more deeply into Topic-Cluster 16, counter to our initial expectations, we made some surprising observations that might help explain the broader reach of this topic-cluster.
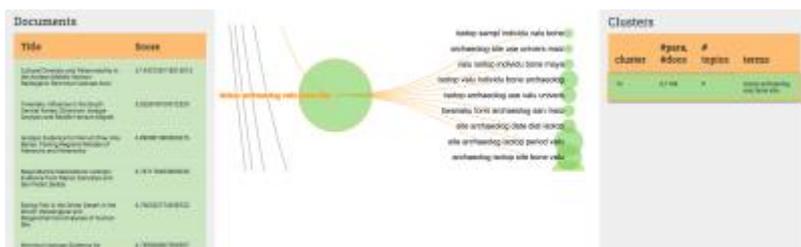
Figure 8: Inset view of Figure 5. Cluster 16 contains nine subtopics placing it in the middle range of cluster vocabularies within the corpus. Contributing terms to this cluster's vocabulary include general language such as site and archaeology, as well as more focused language such as isotope and diet.

Rather than focusing exclusively on isotopic analyses and related dietary issues – which we categorised within a bone as body framing in the pyLDAvis discussion – we discovered in the documents multiple discussions that clearly engaged with bone as an object. For example, we see documents that discuss crafted artefacts made from bone, such as bone flutes, painted bones, and other portable objects (Emery 2008). In some cases, these appear in the same article as discussion of human bone – particularly in the case of burial contexts – and in others not. In order to understand this unexpected appearance of examples of bone as object within an apparent bone as body framing, we manually checked whether discussions of isotopic analyses co-appeared in the articles in which bone artefacts were being discussed; they did not. Another notable linkage may meaningfully position these bone artefacts in conversation with discussions of isotopes and diet: both of these foci deal with partibility and subsequent transformations of bone (i.e. it can be segmented and then modified such that it has different characteristics and functions). In this way these disparate examples of bone (e.g. flutes, isotope samples) share some key material and relational characteristics that cross body and object categories and suggest some similarities in how archaeologists frame these entities in their writing.

In addition to artefacts made out of bone, bone-as-object also appeared within this ostensibly body-orientated topic-cluster in several articles that discussed bone being used as ceramic temper (Fargher 2007; Levine *et al.* 2015; Neff *et al.* 2006; Sharer *et al.* 2006). These examples were extremely interesting to us because they take ideas of partibility and transformation of bone to an extreme. In grinding up bone to act as temper in the process of ceramic production, bone is reduced and changed as a substance (and adopts a transformed function) in a way that is perhaps analogous to samples prepared for isotopic analysis. This highlights in a more dramatic way the transformation of bone material, again intriguing in terms of its appearance under an isotope/diet topic-cluster, which in turn fits into the bone as body framework. This represents a next step beyond partibility, with bone understood to be meaningful, identifiable, and analytically communicative, even if it is no longer visually apparent as bone.

In this way, very different manifestations of bone, research questions, and associated analytical processes co-occur, suggesting that the division between bone as body and bone as object at times pales in significance compared to shared understandings of bone as changeable, while still maintaining meaning and identity as an archaeologically meaningful and interpretable substance. Through this third type of analysis, we mostly clearly identified structured and patterned framings for our test case of 'bone', finding both results that conformed to anticipated disciplinary expectations, but also ones that were more surprising, opening the door to insights into associated beliefs or expectations.

7. Discussion and Conclusions

7.1 Fluid, cyclical framings

We are intrigued by the possibilities for deeper awareness on two fronts that emerge from our analyses. The first is a sense of the multiple conceptual positions contained within, rather than simply between, articles. That is, rather than solely revealing differences in framings that might correspond to differing theoretical or methodological approaches between researchers (and between their scholarly writings), we also note evidence of shifts in writers' framings of bone happening at an intra-article level. This is important because it identifies productive spaces of connection and overlap between scholarship that might seem methodologically or theoretically distinct at first glance. It also illuminates productive places of conceptual friction (in Tsing's (2004) sense of conflict or interaction that leads to changed states) occurring within single works, places that may be rich for deeper investigation.

At the opening of this article, we indicated our interest in identifying assumptions, naturalised categories, and patterned interpretative moves used by archaeologists in writing about particular types of artefactual materials. We wondered if the strengths of large datasets and machine-learning approaches could help bring to light these patterns, illuminating ideas that underlay these patterns, and allowing for disciplinary insights that are not possible without a shift to large-scale analysis. Our results indicate that for our case study of the substance of bone there are regular, recognisable, and sometimes clustered patterns that characterise how experts write about bone. Notably, we were able to move beyond straightforward characterisations of methodological approaches to bone analyses to uncover indications of scholarly assumptions and beliefs about bone, not all of which are overtly acknowledged or expressed. We see these emergent observations not as a critique of disciplinary weaknesses, but rather as a new opportunity. We can use now-available large data corpuses for digital meta-analysis, yielding deepened awareness of our own material ontologies and the ways in which they may continue to appear in our interpretations, despite explicit efforts to decentre them. The analytical possibilities of larger datasets, coupled with machine-learning approaches, can be combined with a disciplinary curiosity to better understand foundational premises in our field. While the present study was conducted on an intentionally limited dataset and topic, this work is readily scalable in a methodological sense, as will be discussed below. At a more conceptual level, we link back to our opening discussions about the impact of scale, and of ways of seeing and organising data, underscoring the ways that steps towards disciplinary self-awareness should draw upon anthropological understandings of visibility, boundaries, and classification or typification.

Our first step of textual groundwork – in which the archaeologists in the team identified themes from a sample set of articles – suggested some initial foci, which, using the pyLDAvis results, began to resolve into contrasting thematic categorisations of bone as body or being, versus bone as object. However, our analysis with multi-level modelling revealed that the situation is more complicated, and that a separation of discursive 'camps' obscures the more complex work that archaeologists are undertaking. Specifically, through our analyses, we are able to see an intriguing reorientation that many archaeologists engage in, in which they fluidly transition between treating bone as being and bone as object. That is, within scholarly texts (and presumably the

accompanying analyses and engagements), bone undergoes cyclical transitions in which it is characterised as different types of substances (body versus object). This is significant because it happens on a short time scale – within a lab, for instance, but also as represented within an article text – and indicates that our treatment and characterisation of this material is notably dynamic. The upshot of these observations is that our material engagement with the archaeological substance of bone is not fixed. That is, our relationship with this material involves analytical positioning that is renegotiated at multiple points in the interpretative process, per the intra-article heterogeneity that we observed. This is significant because it has the potential to shift values and meanings associated with materials. These observations indicate optimism about the nimbleness of archaeological repositioning around materials in a way that is extremely contextually sensitive, and able to incorporate or move between multiple frames of reference. At the shifted scale of observation that we adopt in this article, and leveraging the changes (in boundaries, values, and visuality) that result from a different scale of analysis, we are able to recognise patterns of fluidity in archaeological writing that indicate our field's strengths in connecting between multiple analytical, ontological, and theoretical perspectives.

7.2 Analytically buried: indigenous beliefs about bone

The second point is about raising our awareness of active contextualisation of our research materials; our analyses indicate a lack of patterned centrality of indigenous ontologies in the ways that our field writes about bone. The field of archaeology is already reflective regarding our language and writing practices, and processes of narrative making, in ways that shift awareness of our own ways of doing archaeological work (e.g. Buccellati 2017; Chapman and Wylie 2016; Joyce and Preucel 2002; Lucas 2019); our field has also begun the critical process of questioning how we can move forward on decolonising archaeological practices and interpretations (efforts that must take many forms, per Atalay (2006), but that often involve processes of de-centring (Atalay 2006, 295–97)). We are at a moment when our field is striving to be more inclusive and to more deeply examine underlying assumptions that may structure it. It is transformative to have tools that let us see – using shifted lenses and different scales – the ways in which certain topics or perspectives may not yet be fully integrated into our scholarly practices of writing. This observation moves beyond a recognised need for better intellectual partnership between multiple stakeholders and provides guidance on specific opportunities for where and how this might be productively accomplished.

Specifically, despite the fluidity and multiplicity of perspectives noted above, we also observed an area to which this multiplicity did not extend in notably patterned ways. While in this article we are focused on understanding our own beliefs about artefacts, and bone in particular, the analyses we undertook indicated that language related to *ancient* beliefs about bone (e.g. topics related to indigenous ontologies or ideologies) was not central, in patterned ways, to the discourses used to discuss bone. (Note that we identified a small indication in our pyLDAvis analysis of attention to construction of ancestors, but it was not central to the discourse represented.)

This does not mean, of course, that archaeologists are not writing about ancient ideologies about bone and the dead. However, we return to the idea of scale, and the different ways of seeing that emerge in looking at large-scale data, to contextualise our

observations. The three topic-clusters in our multi-level models that most clearly addressed bone (Topic-Clusters 4, 13, and 16) for the most part did not include topics or terms that pointed towards investigations of emic meanings or beliefs associated with bone. Again, this is looking at large-scale textual patterns (certainly, there are individual studies that discuss these topics), and this suggests that our own beliefs about bone tend to dominate discourse about methods, analysis, treatment, and interpretation of bone artefacts and materials; one wonders if this links back to ideas about expectations of our data, and inclinations to 'domesticate' them in certain ways that make sense without our own bone-related ontologies.
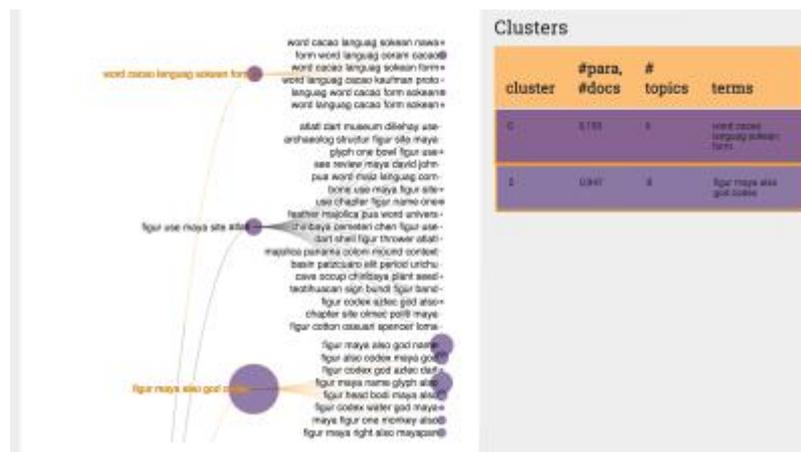
Figure 9: Inset view of Figure 5. Highlighted Clusters 0 and 2 (viewed here with Cluster 1, not highlighted) contain terminology suggesting indigenous ideology and perspectives where bone plays a supporting role, rather than the primary focus of these topics.

Individual articles discussing topics related to indigenous ideology or beliefs appeared buried in Topic-Clusters 0 and 2, for instance, but 'bone' as a term did not rise to the level of term, topic, or topic-cluster in those instances. Intriguingly, the two topic-clusters contributing to these observations, numbers 0 and 2 (see Figure 5 and Figure 9), do not immediately appear to deal with bone. For instance, at the highest (i.e., topic-cluster) level, Topic-Cluster 2, which has a relatively average number of sub-topics, deals with ideology-related topics ('figur maya also god codex'). Looking at the topics within Topic-Cluster 2, we see 'head' and 'body' (bodi) within one of them. These are instances in which bone as a word can be seen as playing a supporting role within a discourse that may be focused on broader or multiple topics. For instance, bone might be appearing in our analysis as part of a discourse about bodies, or gods, or rulers. This may be a productive avenue for future analysis in terms of understanding how patterned written frameworks relating to specific materials cross over into other realms of discussion, underlining the importance of realigning categories in order to avoid obscuring meaningful linkages (per discussions of focused attention at the beginning of this article). This also suggests that further interpretative opportunities will be possible when this type of analysis extends beyond the test case of a single material type.

A shifted awareness is made possible by the scale of the dataset and the machine-learning approach: despite scholarly interest in including indigenous meanings and perspectives in our work (perhaps especially important with regards to osteological material), local ontologies about the artefact category we investigated appeared as systematically peripheral, in patterned and structured ways. Looking at terms that are included in Topic-Clusters 4, 13, and 16 (that is, those topic-clusters that most clearly

address bone), we see a notable absence to references to indigenous or non-western perspectives on bone. Rather, these perspectives seem to be siloed into their own topics. For instance, Topic-Cluster 2, 'figure maya also god codex' is the only topic-cluster that features the term 'god' as an indicator of belief-orientated discussion at both the topic-cluster level and within multiple topics. Outside of this topic-cluster, 'god' appears only one other time, as part of a topic ('figure codex aztec god also') within Topic-Cluster 1; these appearances draw attention to indigenous written sources (that is, indigenous texts) as a distinct source for understanding topics related to bone from a local perspective, again, siloed into a particular topic and topic-cluster (Figure 5). In looking for other terms that could possibly indicate attention to local understandings of bone, we observe Topic-Cluster 0 ('word cacao languag sokean form') could be pointing towards indigenous linguistic elements related to bodies and human remains (e.g. see the appearance of cacao, sometimes understood as an analogue for human blood [Meskell and Joyce 2003, 140; Novotny 2014, 55], within the terms). Looking at the document viewer as a check on these observations, Topic-Cluster 2 and the specified topic from Topic-Cluster 1 do, indeed, indicate articles and discussions that are more orientated towards beliefs and local meanings (and, relatedly, towards text and iconography as evidentiary sources). The articles viewed in the document viewer do not support Topic-Cluster 0 as a place of discussion of local conceptions of bodily substances.

Overall, this indicates to us that indigenous perspectives related to bone should be included in our writing with greater regularity, in connection with topics such as our discussion of excavation and recovery, methods, and documentation of burial contexts. Obviously, not all studies or articles lend themselves to including a discussion of religion, ideology, or ontologies (and on the flip side, nor do all articles omit these perspectives). But, the dominance of a modern, western, scientific framing is clear in our results, which, in our minds, indicates the need to change the dynamics of the voices represented at the table in scholarly products. In this case, it seems that local ontologies and indigenous perspectives should not be treated primarily as separate areas of study (with an accompanying implied hierarchy of value), in the ways that lead them to appear as separate and non-integral in our machine visualisations; rather, they should be systematically considered, inserted, and referenced in our consideration of excavated materials (bone, in this case, but also others). While researchers will continue to have methodological specialisations that emphasise particular evidentiary sources (e.g. osteological analyses versus hieroglyphic ones), awareness of related scholarship is key, and we believe should be integrated into interpretations. Our remarks on the fluid and flexible framings in archaeology articles indicate that we have the disciplinary capacity to accomplish this. The tools we use in this article establish a baseline and show how we can then subsequently monitor change to track our progress as a field.

7.3 Future directions

Our case study in this article has focused on bone: we chose this as an initial test case to explore the utility of this meta-analysis of journal texts as primary source data, using two Latin American archaeology research publications, an admittedly narrow focus. With the studies we presented in this article, we determined that this approach allows for identification of patterned language use and disciplinary discourses that are frequently used to frame this material but that may not be overtly expressed or recognised. Some elements of the present study could be investigated in even greater detail, such as parsing differences between how human and animal bone are framed (e.g. is human

bone more standardised, with differences in animal species examined leading to a greater diversity of textual treatments?), or differences in geographic/cultural context across the Americas (e.g. can we attribute some differences in framings to differences in training and areas of study?). Next, we hope to extend this work more broadly to include other artefactual materials beyond bone and also to move beyond Latin American-orientated research, but also to explore other intriguing category issues. For instance, we are curious about whether particular disciplinary discourses that emerge in conjunction with different artefact types might actually cross material categories, perhaps revealing meaningful commonalities in how we write about (and interact with, analyse, and conceptualise) very diverse artefact classes, indicating how we relate to different substances and our assumptions about them.

Additionally, we are curious about whether different emergent material framings that we are able to see using LDA and other digital scholarship techniques might allow us to shed light on shifting trends in the discipline of archaeology over time (something that Marriner 2009 and Schmidt and Marwick 2020 are able to look at in their datasets, for instance, using computational approaches, though not LDA), the impact of particular centres of training or experience (representing research traditions and/or communities of practice), as well as possible differences based on some identity markers (such as gender, nationality, etc.). How can we more subtly describe differences within subareas of our field, recent directions or changes in our field, and understand some of the elements that may contribute to them? Some of these larger enquiries, including ones with a diachronic perspective, will be most revealing with a larger dataset, indicating the need for inclusion of a greater number of journals. In short, the work presented here is a starting place for a larger endeavour. Our invocation of 'reflexivity' in the present article reflects only a limited perspective on how the field of archaeology can better understand its own discourses, assumptions, and foundations – further, and more specific, work is needed. Our goal with the present work is to open the door for further studies using digital tools in order to query our own field in a spirit of curiosity and increased awareness. While these are not new impulses, the analytic possibilities demonstrated here suggest new and productive avenues for large-scale yet sensitive understandings of elements of our field.

At the beginning of this article we wondered if we could use large datasets and machine-learning analysis to identify patterned ways that archaeologists write about artefacts (bone, in this case), and whether any underlying ideas that shape these patterns might emerge through our analysis. We found that the LDA models were able to identify such patterns. Placing machine learning in dialogue with field experts' interpretations allowed for identification of differentiated written discourses, yet somewhat surprisingly, the result was not a tidy list of different methodological or theoretical 'camps'. Rather, we observed fluidity and flexibility in how the archaeological material of bone was written about and understood. Depending on their background, training, or scholarly orientation, this may be surprising to some archaeologists, and not to others. Rather than being an empty commentary about increased 'nuance' (with the accompanying critiques about the inutility suggested by Healy of the ever-more 'fine-grained view' (2017, 126) and the move to 'embrace complexity rather than to cut through it' (2017, 119), in discussing sociological theory), we suggest that there is analytical and observational impact in the ability to link large-scale datasets with broad questions about our field. We hope that readers will move beyond our work with their own queries that derive from their own positionality in the field. Our findings, driven by our own desire to better understand underlying and often invisible interpretative assumptions in our own scholarship, indicate an opportunity for greater reflexive awareness of our own beliefs about artefacts, as well

as the opportunity to create space for the inclusion of other culturally grounded ways of relating to this material.

Acknowledgements

Code availability: The data processing and computational components of this research rely on free software libraries available through Python including (1) the Natural Language Toolkit [NLTK], (2) Gensim, (3) Scikit-learn, and (4) pyLDAvis (Pedregosa *et al*. 2011; Řehůřek and Sojka 2010; Sievert and Shirley 2014; Wagner 2010). From the NLTK library, we first remove stop words and stem all the words in our dataset with the Stopwords and Porter Stemmer packages respectively. To generate our multi-level model, we build six topic models from this processed dataset using the Latent Dirichlet allocation package from the Gensim library. The results of those six models are then clustered into the topic-clusters, using the Affinity Propagation module from Scikit-learn's Cluster package. Example code for Gensim's LDA is widely available and clustering process code is available on Scikit-learn's website.