

---

# A CULTURE OF INERTIA: THE REALITIES OF CONTENT STANDARDS AND THE ARCHAEOLOGISTS WHO NEED THEM

---

by Corinne Softley

THE UNIVERSITY OF  
NEW SOUTH WALES



SYDNEY • AUSTRALIA

Submitted to the School of Humanities, the University of New South Wales, in partial  
fulfilment of the requirements of the Bachelor of Arts with Honours in History and  
Archaeology

# ACKNOWLEDGEMENTS

To my honours companion Georgia Burnett, thank you for your friendship and support throughout the year.

To the respondents of the “Vocabularies, Concepts and Standards” survey, thank you for taking part in my survey. Your participation made my research possible.

To the Federated Archaeological Information Management Systems (FAIMS) project, I am grateful for your help with disseminating and promoting the “Vocabularies, Concepts and Standards” survey. I also commend the Australian Archaeological Association (AAA) and OzArch for promoting the survey.

To my supervisor Dr. Shawn Ross, I extend a significant deal of gratitude for your support throughout the year. I also sincerely thank my co-supervisors Dr. Brian Ballsun-Stanton and Dr. Adela Sobotkova for their expertise and guidance.

To all my family and friends, thank you for your time, support and encouragement.

# TABLE OF CONTENTS

<b>ACKNOWLEDGEMENTS</b>	<b>2</b>
<b>TABLE OF FIGURES</b>	<b>5</b>
<b>ABSTRACT</b>	<b>6</b>
<b>INTRODUCTION</b>	<b>8</b>
<b>CHAPTER ONE: DIGITAL ARCHAEOLOGY AND DATA ACCESSIBILITY</b>	<b>16</b>
1.1. THE CRISIS OF DATA ACCESSIBILITY	16
1.2. THE IDIOSYNCRATIC NATURE OF ARCHAEOLOGY	18
1.3. DATA STANDARDS IN COMPUTATIONAL ARCHAEOLOGY	23
1.3.1. TECHNICAL STANDARDS	24
1.3.2. METADATA STANDARDS	25
1.3.3. CONTENT STANDARDS	27
1.3.4. THE LACK OF CONTENT STANDARDS	29
1.4. FRACTURED TERMINOLOGY	31
1.4.1. VOCABULARIES, CONCEPTS AND STANDARDS SURVEY	34
1.5 CONCLUSIONS	34
<b>CHAPTER TWO: THE “VOCABULARIES, CONCEPTS AND STANDARDS” SURVEY</b>	<b>36</b>
2.1. LIMITATIONS	38
2.2. SURVEY RESPONDENTS	39
2.3. RESULT PROCESSING AND ANALYSIS	42
2.4. PEARSON’S CHI-SQUARE TEST OF INDEPENDENCE	44
<b>CHAPTER THREE: SOCIAL FACTORS AND TERMINOLOGY STANDARDS</b>	<b>48</b>

3.1. PARTICIPANT UNDERSTANDINGS OF FRACTURED TERMINOLOGY	49
3.2. THE “NOT INVENTED HERE” SYNDROME	56
3.3. SELF-EXCLUSION	60
3.4. GOVERNMENTAL CONTROL OF TERMINOLOGY USE	65
3.5. NOBODY’S PROBLEM	67
3.6. HOW SERIOUS IS A LACK OF STANDARDS?	69
3.7. DISCUSSION	72
<b>CONCLUSION AND PROSPECTS FOR FUTURE RESEARCH</b>	<b>74</b>
<b>APPENDIX</b>	<b>83</b>
APPENDIX ONE – SURVEY FORM	83
APPENDIX TWO – SURVEY RESULTS	99
2.1. QUANTITATIVE RESPONSES	99
2.2. QUALITATIVE RESPONSES	112
APPENDIX THREE – PEARSON’S CHI-SQUARE TEST OF INDEPENDENCE	125
APPENDIX FOUR – “ARTEFACT SCATTER” COMPONENT FORM	138
<b>REFERENCES</b>	<b>142</b>

# TABLE OF FIGURES

<i>Figure 1. Question 32a</i>	<i>52</i>
<i>Figure 2. Question 32c</i>	<i>52</i>
<i>Figure 3. Question 32b</i>	<i>53</i>
<i>Figure 4. Question 32d</i>	<i>53</i>
<i>Figure 5. Question 27</i>	<i>55</i>
<i>Figure 6. Question 18</i>	<i>57</i>
<i>Figure 7. Question 30a</i>	<i>61</i>
<i>Figure 8. Question 30b</i>	<i>61</i>
<i>Figure 9. Question 30c</i>	<i>61</i>
<i>Figure 10. Question 30d</i>	<i>62</i>
<i>Figure 11. Question 16</i>	<i>64</i>
<i>Figure 12. Question 17</i>	<i>65</i>
<i>Figure 13. Question 34</i>	<i>71</i>

# ABSTRACT

As a discipline, archaeology suffers from inaccessible data online. Although digital repositories are progressive tools for modern archaeological practice, control over resources is hindered by the lack of data comparability. At the core of this deficit is a lack of semantic interoperability, aggravated by idiosyncratic terminology use. Archaeologists currently use loosely defined terminologies, creating a tangle of fractured definitions. Although attempts have been made to standardise terminology use, resistance and distrust within the archaeological community hinders any progress.

Supported by a survey of practicing archaeologists, this thesis explores archaeological understandings and attitudes towards terminology standards. The survey results identify a culture of inertia with archaeology. Content standards change and limit the nature of archaeological practice to one defined semantic model. That is, a practice in which data collection, documentation and analysis is shaped by the individual project; standards apply a uniform structure to the creation of data. Given the realities identified by the survey, this thesis argues that content standards must embrace the diversity of data and archaeological practice. Rather than prescribing a rigid schema for capturing the meaning of data, terminology should be mapped using descriptive ontologies. Ontologies are networks of terminologies that represent objects, properties of objects and relations between objects. By adopting ontologies, the wider archaeological community may be willing to adopt content standards and create commensurate data.



# INTRODUCTION

*"The fact that researchers [engaged in synthetic research] can reach conflicting conclusions, not because their primary data differ but because they operationalize interpretive terms different, suggests even more subtle difficulties than those deriving from out dated or erroneous conclusions"<sup>1</sup>*

Over the last few decades, the disciplines of the humanities have undergone a revolution in the way resources are stored, accessed and shared. Ongoing and new alliances between computational specialists, librarians and humanities scholars have produced the "digital humanities"; technological innovations for the storage of humanities resources online for preservation and wide spread access.<sup>2</sup> Digital repositories have been particularly progressive for modern archaeological practice, as scholars can compare and preserve knowledge of the long-term human past more effectively. Despite the potentials, the World Wide Web – particularly the 'deep web', inaccessible through conventional search engines such as Google - is a "fire hose gushing with information" that cannot be fully harnessed or investigated.<sup>3</sup> For example, it is

---

<sup>1</sup> Keith. W. Kintigh, Francis P. McManamon, and Katherine Spielmann, "Enhancing Data Comparability and Enabling Synthesis with Tdar (the Digital Archaeological Record)" (paper presented at the 78th Annual Meeting of the Society for American Archaeology, Honolulu, Hawaii, April 3-7 2013), 2.

<sup>2</sup> David Mattison, "The Digital Humanities Revolution," *Searcher* 14, no. 5 (2006): 27.

<sup>3</sup> Paul Sturges and Anne Griffin, "The Archaeologist Undeceived: Selecting Quality Archaeological Information from the Internet," *Informing Science Journal* 6(2003): 221.



estimated that the information on the deep web is 400-550 times larger than the commonly defined World Wide Web.<sup>4</sup>

For archaeology, control over resources is hindered by the lack of interoperability. Interoperability is the ability to meaningfully and accurately exchange information across dissociated datasets. Semantic interoperability—the meaningful exchange of information based on data content—is central to the crisis of data access. Data content includes elements such as vocabularies.<sup>5</sup> Accurate understanding of data content within and between systems is vital for data synthesis and access. Currently, semantic interoperability is difficult to achieve for digital archaeology due to the nature of archaeological practice.

There are also two additional aspects of interoperability integral to the spread of information digitally: technical and syntactic interoperability. Technical interoperability is associated with the software/hardware components, systems and platforms that enable machine-to-machine communication to take place.<sup>6</sup> That is, technical interoperability concerns the practical aspects of linking computer systems and services.<sup>7</sup> Moving onto data itself, syntactical interoperability exchanges information

---

<sup>4</sup> Ceri Binding, Keith May, and Douglas Tudhope, "Semantic Interoperability in Archaeological Datasets: Data Mapping and Extraction Via the Cidco Crm," in *Research and Advanced Technology for Digital Libraries* (Berlin Heidelberg: Springer, 2008), 280.

<sup>5</sup> Sandra Heiler, "Semantic Interoperability," *ACM Computing Surveys* 27, no. 2 (1995): 271.

<sup>6</sup> Hans van der Veer and Anthony Wiles, "Achieving Technical Interoperability," (Sophia Antipolis Cedex, France: European Telecommunications Standards Institute, April 2008), 5.

<sup>7</sup> Herbert Kubicek, Ralf Cimander, and Hans Jochen Scholl, "Layers of Interoperability," in *Organisational Interoperability in E-Government: Lessons from 77 European Good-Practice Cases* (Berlin: Springer, 2011), 86.

based on data formats and syntax elements.<sup>8</sup> Syntax refers to the grammar and formal rules for defining sets of data.<sup>9</sup>

In archaeology, accurate and sophisticated interoperability has not yet been achieved. If interoperability is achieved, computer systems are able to compare and retrieve the most relevant data for researchers.<sup>10</sup> Archaeology suffers from data inaccessibility. At the core of the crisis rests a lack of semantic interoperability, aggravated by idiosyncratic terminology use.

Although archaeology is collaborative and multi-disciplinary, requiring large datasets to be shared across teams of researchers, the discipline is idiosyncratic in nature. Archaeology is dominated by a belief that every project is, and should be, unique; this attitude is described by Julian Richards as “the not invented here syndrome”.<sup>11</sup> The attitude stems from the diversity of data encountered on archaeological projects, and an association with the small sciences. Small science disciplines straddle both the humanities and sciences.<sup>12</sup> With a focus on case-specific research, testing, scientific methodologies and the re-use of data, small sciences demand diverse documentation methods.<sup>13</sup> As a result, archaeologists generally understand the collection and analysis of data as a practice shaped by the individual project;

---

<sup>8</sup> Veer and Wiles, "Achieving Technical Interoperability," 5.

<sup>9</sup> Kubicek, Cimander, and Scholl, "Layers of Interoperability," 89.

<sup>10</sup> Veer and Wiles, "Achieving Technical Interoperability," 5.

<sup>11</sup> Julian D. Richards, "From Anarchy to Good Practice: The Evolution of Standards in Archaeological Computing," *Archeologia e Calcolatori* 20(2009): 27.

<sup>12</sup> Eric Kansa and Ahrash Bissell, "Web Syndication Approaches for Sharing Primary Data in "Small Science" Domains," *Data Science Journal* 9(2010): 42.

<sup>13</sup> Ibid.

archaeological data does not adhere to any specific content standards.<sup>14</sup> However, content standards are fundamental to computational archaeology, and more specifically, to semantic interoperability.

Content standards dictate how data is represented. If implemented at the beginning of the data life cycle, content standards define parameters for how data is articulated.<sup>15</sup> Archives and repositories might use these parameters to accurately compare resources and identify relevant data. If the content parameters are standardised across a discipline or database, the time spent mapping a dataset to an archival standard is reduced. Therefore, content standardisation is essential to the future of computational archaeology. Without content standards, sophisticated sharing and comparison of data beyond local levels of research cannot happen.<sup>16</sup> Synthetic digital archives and repositories have the potential to advance long-term, large-scale understandings of human history through the sophisticated integration of data.<sup>17</sup> Content standards that reconcile fractured terminology would be particularly beneficial to the discipline.

Archaeologists currently use loosely defined vocabularies between domain-specific knowledge representations. Individualistic description arises out of the

---

<sup>14</sup> Shawn Ross et al., "Creating Eresearch Tools for Archaeologists: The Federated Archaeological Information Management Systems Project," *Australian Archaeology* 77(2013): 110.

<sup>15</sup> Andrew Baines and Kenneth Brophy, "What's Another Word for Thesaurus? Data Standards and Classifying the Past," in *Digital Archaeology: Bridging Method and Theory*, ed. Thomas L. Evans and Patrick Daly (London: Routledge, 2006), 237.

<sup>16</sup> Richards, "From Anarchy to Good Practice," 29.

<sup>17</sup> Eric Kansa, "Openness and Archaeology's Information Ecosystem," *World Archaeology* 33, no. 4 (2012): 498.

synonymous (two words with the same meaning), polysemous (a single word with multiple meanings) and conceptual nature of archaeological terms. Semantic connections between data are limited as a result of fractured terms.<sup>18</sup> If computer systems do not connect accurate meaning within data content, they cannot operate synthetically.<sup>19</sup> Thus, fractured terminology obstructs any reuse, repurposing and reproducibility of archaeological data. According to Keith Kintigh, access issues have come to limit the work of many projects, by preventing their understanding of phenomena that operate on a large scale.<sup>20</sup> Individual projects alone are, in principle, rarely able to answer questions of more than local interest.<sup>21</sup> For a semantically interoperable database, precise meaning must be acquired through the integration of the synonymous, polysemous and conceptual entities of archaeological data. Terminology standards are the first step towards semantic interoperability.

Terminology standards are basic, and more intricate, sets of terms that are common across a variety of domain-specific concepts and vocabularies.<sup>22</sup> If terminology is standardised and used during the creation of data, digital repositories can operate

---

<sup>18</sup> Jeff Heflin and James Hendler, *Semantic Interoperability on the Web*(Maryland University College Park Department of Computer Science, 2000), 1.

<sup>19</sup> Ibid.

<sup>20</sup> Keith W. Kintigh, "The Challenge of Archaeological Data Integration," in *Technology and Methodology for Archaeological Practice: Practical Applications for the Past Reconstruction*(BAR International Series S 2029, 2009), 1.

<sup>21</sup> Kintigh, McManamon, and Spielmann, "Enhancing Data Comparability and Enabling Synthesis with Tdar (the Digital Archaeological Record)," 1.

<sup>22</sup> Jane Hunter, "Enhancing the Semantic Interoperability of Multimedia through a Core Ontology," *Circuits and Systems for Video Technology, IEEE Transactions* 13, no. 1 (2003): 49.

synthetically.<sup>23</sup> Archaeologists should have confidence in the discovery, integration and preservation of their data online if terminology is standardised.<sup>24</sup> Without a defined and uniform terminology, there is the danger of disconnected site-specific digital data, with little or no relevance to other sites or projects.<sup>25</sup> Terminology standards offer a viable path to harnessing the potentials of computational archaeology.

Although the benefits are clear, archaeology as a discipline is caught between what it wants and what it needs. The discipline may require data standards for interoperability, but standards also challenge the idiosyncratic nature of archaeology. Many archaeologists feel that content standards could restrict intellectual inquiry and potential research avenues, and would curb the uniqueness of archaeological projects.<sup>26</sup> As a result, twenty years of scholarship and standards implementation have been unsuccessful for widespread content standards adoption. Unless the nature of practicing archaeologists is better understood, resistance and distrust in data standards will continue to challenge digital archaeology.<sup>27</sup> Any project that fails to take into account its core users and players is ultimately doomed to fail.

To conduct wider discussion and explore opinions on terminology standards, I distributed a survey titled "Vocabularies, Concepts and Standards" in June of 2013. A

---

<sup>23</sup> Kansa, "Openness and Archaeology's Information Ecosystem," 498.

<sup>24</sup> Heflin and Hendler, *Semantic Interoperability on the Web*, 1.

<sup>25</sup> Kintigh, "The Challenge of Archaeological Data Integration," 1.

<sup>26</sup> Julian D. Richards and Catherine Hardman, "Stepping Back from the Trench Edge: An Archaeological Perspective on the Development of Standards for Recording and Publication," in *The Virtual Representation of the Past*, ed. M. Greengass and L. Hughes (Farnham, Surrey and Burlington, USA: Ashgate Publishing Company, 2008), 168.

<sup>27</sup> Richards, "From Anarchy to Good Practice," 31.

range of insightful data was gathered from the 117 practicing archaeologists and students who participated. The survey had respondents share their experiences with fractured terminology, and how it has impacted fieldwork, research and publication. Participant opinions concerning the need for standardised terminology were also investigated, followed by questions that assessed the general willingness (or unwillingness) of archaeologists to compromise their practice for terminology standards. Additionally, the survey examined the development and use of preferred language, core concepts and professional vocabularies.

The survey dataset identifies varied understandings of terminology standards and, more importantly, a deep resistance to standards implementation. The results initially demonstrate widespread understanding of a lack of terminology standards. Upon further examination, problems in translating specific concepts suggest a potential lack of understanding among the survey group. Additionally, many respondents view the imposition of controlled terminology as a threat to the individuality and flexibility of their own projects. They believe that their projects are unique and deserve customised treatment; the mentality described by Julian Richards as the “not invented here syndrome.”<sup>28</sup> In contrast, some respondents exclude themselves from the use of fractured terminology all together. Instead, they blame others for idiosyncratic practices: those outside of their sub-disciplines or specific government legislation are primary targets. Furthermore, the survey reflects a “we can work with it attitude.” That is, archaeologists are willing to work with fractured terminology rather than move towards a solution. This inertia towards content standards is discussed in depth throughout chapter three.

---

<sup>28</sup> Ibid., 27.

Based on the current field of knowledge and results of the “Vocabularies, Concepts and Standards” survey, it is clear that the majority archaeologists are unwilling to adopt rigid content standards that could change and limit archaeological practice. Thus, content standards need to embrace diversity of data, research questions and archaeological practice. Instead of prescribing a common schema for capturing the meaning of data, semantic interoperability should be harnessed through the use of descriptive ontologies. An ontology is “the representation of meaning of terms in vocabularies and the relationships between those terms.”<sup>29</sup> In other words, ontologies take into account dynamic, versatile terms used by archaeologists. Ontologies as a solution will be discussed in detail in the concluding chapter.

Idiosyncratic terminology is an enormous problem for semantic interoperability that will require engagement and collaboration by the whole archaeological community. Archaeologists can continue to scatter data across an ever-proliferating array of un-integrated databases. They will end up with an “archaeological tower of babel” as a result.<sup>30</sup> It is only through wider discussion, cooperation and agreement within the archaeological community that data will become commensurate for advanced synthesis and accessibility.

---

<sup>29</sup> Deborah McGuinness and Frank van Harmelen, “Owl Web Ontology Language,” *W3C Recommendation* 10, no. 2004-03 (2004).

<sup>30</sup> Richards, “From Anarchy to Good Practice,” 27.

# CHAPTER ONE: DIGITAL ARCHAEOLOGY AND DATA ACCESSIBILITY

## 1.1. THE CRISIS OF DATA ACCESSIBILITY

Propelled by converging information technologies, archaeology stands on the brink of major changes in the creation, use and preservation of archaeological data. A revolution in mobile devices has emerged, extending from Apple products to Google's open-source Android operating system. Open-source environments for data management such as geographic information systems (GIS) and computer-aided design (CAD) have also matured.<sup>31</sup> Additionally, high-quality online archives and publication services have been implemented for digital archaeology. Placed together, technological improvements and the creation of high-quality digital spaces have laid the foundation for the development of archaeological information systems. Information systems are the networks of hardware and software used to create, edit, analyse, archive and manage

---

<sup>31</sup> Ross et al., "Creating Eresearch Tools for Archaeologists," 109.

The Open Geospatial Consortium (OGC) is an open-source GIS project. The OGC are an international GIS standards organisation encouraging the development and implementation of geospatial content and services, GIS data processing and data sharing. See <<http://www.opengeospatial.org>>

Star UML 5.0 is an open-source CAD project. The project aims to develop a fast, flexible, extensible and freely available unified modelling language (UML) and model driven architecture (MDA) platforms for diagram modelling. See <<http://staruml.sourceforge.net/en/>>



data for reusable datasets.<sup>32</sup> Archaeological information systems can improve archaeological research by facilitating the reuse, repurposing and reproducibility of data. Despite this potential, the discipline is threatened by the limited usefulness of archaeological data due to insufficient data interoperability.

‘Interoperability’ has three aspects, ‘technical’, ‘syntactic’ and ‘semantic’. Technical interoperability concerns the practical aspects of linking computer systems and services.<sup>33</sup> Machine-to-machine communication, for example, relies on technical interoperability. Syntactic interoperability is the communication and exchange of data via common formatting protocol and documentation.<sup>34</sup> Syntactic interoperability concerns, but is not limited to, data formats (for example, video and audio) and recording formats (for example, HTML).<sup>35</sup> In contrast, semantic interoperability concerns the exchange of shared, unambiguous information between computer systems via data content.<sup>36</sup> More specifically, data is exchanged based on the definition of data elements. Unlike syntax, which refers to grammar and the formal rules for defining sets of data, semantics define the meaning and use of data.<sup>37</sup> Shared vocabularies are crucial to semantics; terminology and description define contents of a dataset.<sup>38</sup> Without semantic interoperability, data cannot be meaningfully compared.

---

<sup>32</sup> Ibid., 2.

<sup>33</sup> Kubicek, Cimander, and Scholl, "Layers of Interoperability," 86.

<sup>34</sup> Ibid., 85.

<sup>35</sup> Ibid., 86.

<sup>36</sup> Heiler, "Semantic Interoperability," 271.

<sup>37</sup> Kubicek, Cimander, and Scholl, "Layers of Interoperability," 89.

<sup>38</sup> Heiler, "Semantic Interoperability," 271.

As indicated by Herbert Kubicek, "Only if the semantics of data sets are defined and shared can data be processed in one system, sent to another system, and there be automatically recognized and processed further."<sup>39</sup> If data is poorly defined and shared across computer systems, researchers searching for specific data will be returned with irrelevant results.<sup>40</sup> Currently, archaeological research is threatened by inadequate semantic interoperability.

The limited usefulness of archaeological data stems primarily from the idiosyncratic nature of archaeological practice, the diversity of data, and the lack of data standards within the discipline. More specifically, the individual and customised nature of data content hinders the exchange of shared meaning. Ultimately, the nature of archaeological practice as it is now limits the semantic interoperability of a dataset, and hinders the dissemination, discovery and preservation of data. Chapter one presents and discusses these hindrances to semantic interoperability.

## 1.2. THE IDIOSYNCRATIC NATURE OF ARCHAEOLOGY

Archaeology is a collaborative and multi-disciplinary field, and requires large datasets to be compiled and shared across teams of researchers. However, much of the data produced is idiosyncratic in nature. According to Julian Richards, archaeology is dominated by a belief that every project is, and should be, unique.<sup>41</sup> The belief stems

---

<sup>39</sup> Kubicek, Cimander, and Scholl, "Layers of Interoperability," 89.

<sup>40</sup> Veer and Wiles, "Achieving Technical Interoperability," 5.

Machine-readable data is a format, description or documentation of data understood by a computer

<sup>41</sup> Richards, "From Anarchy to Good Practice," 27.

from the diversity of data encountered on archaeological projects, and the association of archaeology with the small sciences. Small sciences are focused in decentralised institutions centring on case-specific research.<sup>42</sup> They exercise what Kansa and Bissell state as “customised methods and recording systems, and individually maintained data resources”.<sup>43</sup> As a result, small science disciplines do not generally adhere to specific or uniform methodological or recording standards. The association of archaeology with the small sciences arises out of emphasis on testing, scientific methodologies and the re-use of data, when compared to other humanities disciplines.<sup>44</sup> Archaeology straddles the humanities, social sciences and natural sciences, necessitating varied approaches to data management and documentation.<sup>45</sup>

Archaeological theory—for example, the idea that “data is not fact and therefore there are no objective records”—also influences idiosyncratic practice.<sup>46</sup> While the idea that data is not fact is perhaps extreme, there is a definite subjective aspect to the collection data. Archaeologists generally understand the collection and analysis of data as a practice shaped by the individual project and research question.<sup>47</sup> Data collection is often geared towards answering specific research question, and the results produced reflect this practice. Ad hoc practice creates disparate data that can have little or no

---

<sup>42</sup> Kansa and Bissell, "Web Syndication Approaches for Sharing Primary Data in "Small Science" Domains," 42.

<sup>43</sup> Ibid.

<sup>44</sup> Richards, "From Anarchy to Good Practice," 31.

<sup>45</sup> Kansa and Bissell, "Web Syndication Approaches for Sharing Primary Data in "Small Science" Domains," 42.

<sup>46</sup> Richards, "From Anarchy to Good Practice," 29.

<sup>47</sup> Ross et al., "Creating Eresearch Tools for Archaeologists," 110.

relevance beyond local levels of research.<sup>48</sup> Thus, disparate data hinders the shared, unambiguous meaning required for semantic interoperability. Disparate data cannot be “processed in one system, sent to another system, and there be automatically recognized and processed further”.<sup>49</sup> Data needs to share common meaning for semantic interoperability. The result, according to Kansa and Bissel, is a digital landscape of “data collections fragmented among various content silos”.<sup>50</sup>

Ad hoc recording methods are diverse and numerous, occurring within practices such as pen-and-paper recording. Although technology for digital data creation exists, hard-copy recording remains a common practice; paper is a dynamic, easy and low-cost method that can be customised to project aims. Pen-and-paper recording may be beneficial to the immediate project, but the practice frequently produces incomplete datasets. Pages can be lost, or deteriorate over time. Deficient datasets stem from diverse project methodologies and recording policies, as well as space constraints and the expense of producing complete catalogues.<sup>51</sup> Accurately digitising incomplete and diverse paper records is time-consuming, expensive and difficult to perfect. The digitisation of paper records is often performed in the absence of IT expertise.<sup>52</sup> Thus, the movement of idiosyncratic resources from hard copy to digital often produces poorly designed or implemented databases, unsuitable for wider dissemination. Even if

---

<sup>48</sup> Richards, "From Anarchy to Good Practice," 29.

<sup>49</sup> Kubicek, Cimander, and Scholl, "Layers of Interoperability," 89.

<sup>50</sup> Kansa and Bissell, "Web Syndication Approaches for Sharing Primary Data in "Small Science" Domains," 43.

<sup>51</sup> Dean R. Snow, Mark Gahegan, and Lee C. Giles, "Cybertools and Archaeology," *Science* 311, no. February (2006): 958.

<sup>52</sup> Ross et al., "Creating Eresearch Tools for Archaeologists," 108.

IT support is available the custom-made databases that emerge that are not interoperable with other projects. Repeated customisations essentially reinvent the wheel every time a new dataset is created, repeating the production of incompatible data.<sup>53</sup>

The creation of disparate data is not confined to paper records. Museums, government agencies, universities and individual researchers possess many diverse types of incompatible 'legacy data'. Legacy data refers to inherited or pre-existing data, much of which may be technologically obsolete or otherwise problematic, but is still of considerable value to researchers. It may include published and/or unpublished hard copy reports (with associated images, maps, and catalogues), digital multimedia, spreadsheets, custom databases, GIS files and geo-databases, to name a few.<sup>54</sup> In the past, datasets were widely understood as the by-product of an individual or a project, and a 'means to an end.'<sup>55</sup> As a result, individualised data structures and documentation became common practice. The use of varying vocabularies and taxonomies also surfaced (see section 1.3 for detail on fractured terminology). Nowadays, legacy datasets are difficult to disseminate online. Computer-assisted interoperability demands standardised, universal data content and documentation. Datasets shaped by individual projects do not possess any universal structures or meaning required for data sharing. Legacy data, due to its idiosyncratic nature, is also at risk when it comes to digital preservation.

Valuable data can become isolated if it appears to have little or no relevance to other data. As material evidence is often permanently altered or even destroyed during

---

<sup>53</sup> Ibid.

<sup>54</sup> Snow, Gahegan, and Giles, "Cybertools and Archaeology," 958.

<sup>55</sup> Richards and Hardman, "Stepping Back from the Trench Edge," 168.

the process of archaeological investigation, the preservation of archaeological data is vital. Meckseper and Warwick encapsulate the importance of preservation:

Archaeology is a destructive process: the physical remains in the ground are destroyed through their excavation and lifting of material...the written record and publication have therefore always been seen as synonymous with the preservation of the archaeological record.<sup>56</sup>

That is, when a site is excavated, the published record and any underlying primary data are the only surviving elements of an archaeological site.<sup>57</sup>

Although publications and associated data are vital for preservation, Julian Richards stresses that archaeology is beginning to reach a publication and archival crisis point:

Full publication is becoming increasingly expensive and difficult to achieve, and excavation monographs are read by few people, and bought by even fewer. Meanwhile museum archives are also reaching breaking point; most are running out of storage space; few can provide facilities for access and almost all report low levels of usage. Digital technology now offers the means by which the crisis may at last be overcome.<sup>58</sup>

---

<sup>56</sup> Christiane Meckseper and Claire Warwick, "The Publication of Archaeological Excavation Reports Using Xml," *Literary and Linguistic Computing* 18, no. 1 (2003): 64.

<sup>57</sup> Alison Babeu, *Rome Wasn't Digitised in a Day: Building a Cyberinfrastructure for Digital Classicists*, (Washington: Council on Library and Information Resources, 2011), <http://www.clir.org>. 67.

<sup>58</sup> Julian D. Richards, "Digital Preservation and Access," *European Journal of Archaeology* 5, no. 343 (2002).

Thus, the preservation of legacy data online is beneficial to the archaeological record (if it is accurately defined and shared).

The digital revolution invites archaeologists to share and preserve their data in digital form. In order to participate in this digital movement, archaeologists need to produce more comprehensive records, which would be suitable for digital exchange and sharing. Archaeologists can produce such records, but only with the help of data standards.<sup>59</sup> Data standards define universal documentation and representational structures that allow for shared meaning between data. Thus, data standards are vital aspects of dataset interoperability within computational archaeology.

### 1.3. DATA STANDARDS IN COMPUTATIONAL ARCHAEOLOGY

Data standards represent agreed documentation and representational elements of data.<sup>60</sup> More specifically, standards are universal rules and guidelines for the creation, description and management of data.<sup>61</sup> They ensure others understand the meaning of data as it is intended. Thus, data standards increase the potential for wide dissemination and assured preservation of archaeological data online.<sup>62</sup> Explained earlier, discovery

---

<sup>59</sup> Kintigh, "The Challenge of Archaeological Data Integration," 3.

<sup>60</sup> Baines and Brophy, "What's Another Word for Thesaurus? Data Standards and Classifying the Past," 237.

<sup>61</sup> Tony Gill and Paul Miller, "Re-Inventing the Wheel? Standards, Interoperability and Digital Cultural Content," *D-Lib Magazine* January 2002.

<sup>62</sup> Baines and Brophy, "What's Another Word for Thesaurus? Data Standards and Classifying the Past," 237.

and access to relevant data relies on shared representations of data.<sup>63</sup> Three broad types of data standards are relevant to archaeology: technical standards, metadata standards and content standards.

### 1.3.1. TECHNICAL STANDARDS

Technical standards are the most basic data standards. Technical standards concern the established norms or requirements of hardware and software.<sup>64</sup> In the early days of archaeological computing, archaeologists often used incompatible computer systems. As a result, archaeologists commonly faced issues when opening or reading files created by other archaeologists.<sup>65</sup> Files created on one system were not compatible with another system. The basic technical incompatibility prevented reuse and sharing of data.

With the rise of the Microsoft and Apple duopoly, the problem of software and hardware compatibility largely disappeared. Most archaeologists now use either Windows or iOS operating systems, running on a PC or Apple computer.<sup>66</sup> With regards to application software, a few select products including the Microsoft Office suite, Adobe Creative Suite, have dominated the market. Specific software controls specific tasks; for example, AutoDesk controls computer aided design (CAD), and ESRI is the primary platform for geographic information systems (GIS).<sup>67</sup> These products are so ubiquitous

---

<sup>63</sup> Kintigh, McManamon, and Spielmann, "Enhancing Data Comparability and Enabling Synthesis with Tdar (the Digital Archaeological Record)," 1.

<sup>64</sup> Richards, "From Anarchy to Good Practice," 28.

<sup>65</sup> Ibid.

<sup>66</sup> Ibid.

<sup>67</sup> Ibid.



that their file formats have become de facto standards (for example, DOCX, PDF, DWG, and shape files respectively). Their manufacturers have also begun to provide the export facilities in standard, open file formats such as DNG, KML, and WKT. Keyhole Mark-up Language (KML) and Well-known Text (WKT) are examples for GIS.<sup>68</sup> As proposed by Julian Richards, technical standards were universal formats that the archaeological community did not need to agree upon: the market has forced them upon the discipline instead.<sup>69</sup>

### 1.3.2. METADATA STANDARDS

Metadata, by contrast, has not benefited from the positive externalities of agreement within the software industry. Metadata, the “data about data,” concerns the discoverability and documentation of data, and constitutes a major factor in data’s usefulness.<sup>70</sup> During the 1990s, emphasis on standardised recording moved from technical standards to the succinct documentation of the content of archaeological data.<sup>71</sup> Such metadata proves to be structured information that clarifies the location and

---

<sup>68</sup> Open Geospatial, “Kml,” <<http://www.opengeospatial.org/standards/kml>> (Accessed: 29 October, 2013). ; Keyhole Mark-up Language is an XML notation for expressing geographic annotation and visualisation for Internet based maps such as Google Earth

ArcGIS Server 9.3 Help, “The Ogc Well-Known Text Representation of Spatial Reference Systems,” [http://webhelp.esri.com/arcgisserver/9.3/java/index.htm#geodatabases/the\\_ogc-607957855.htm](http://webhelp.esri.com/arcgisserver/9.3/java/index.htm#geodatabases/the_ogc-607957855.htm) (Accessed: 29 October, 2013). ; Well-known Text is a text markup language

<sup>69</sup> Richards, “From Anarchy to Good Practice,” 22.

<sup>70</sup> Ibid., 30.

<sup>71</sup> Ibid.

usefulness of data. For example, if we consider a library book to be data, then the author, title, subject and location of the item are the metadata. These cataloguing details allow the reader to evaluate the book's usefulness and locate it within the library. In other words, metadata allows researchers to locate, retrieve, bring together and evaluate data resources.<sup>72</sup>

Turning towards the digital space, the appropriate documentation of data using metadata standards caters to dataset discoverability.<sup>73</sup> By describing datasets in a format understood by a computer—machine-readable data—digital repositories, and in turn their users, can fully harness data.<sup>74</sup> The automated matching of well-documented metadata allows archives and repositories to return high-quality, relevant datasets. An example of a metadata standard is the Dublin Core ISO standard. The Dublin Core is a set of terms used to describe a wide range of resources, such as videos, images, books and objects, for the purpose of their discovery.<sup>75</sup> It includes, amongst other things, fifteen classic metadata elements:

1. Title
2. Creator
3. Subject
4. Description
5. Publisher
6. Contributor

---

<sup>72</sup> W. Kilbride, "The Danube in Prehistory in the Digital Age: Towards a Common Information Environment for European Archaeology," *Archeologia e calcolatori* 15(2004): 134.

<sup>73</sup> Richards, "From Anarchy to Good Practice," 30.

<sup>74</sup> N.I.S.O. Press, "Understanding Metadata," *National Information Standards* 20(2004): 2.

<sup>75</sup> ASIS&T, "Dublin Core Metadata Initiative: Making It Easier to Find Information," <http://dublincore.org> (Accessed: 5 October, 2013).

7. Date
8. Type
9. Format
10. Identifier
11. Source
12. Language
13. Relation
14. Coverage
15. Rights

Although each element is optional, the Dublin Core enhances synthesis between metadata through universal resource description.<sup>76</sup>

Metadata standards are generally flexible concerning the nature and quality of the underlying data. They are standards governing how data is described or documented, not standards governing the data itself. Thus, metadata standards do not control how something should be recorded on a content level, nor does it describe how data should be created.<sup>77</sup>

### 1.3.3. CONTENT STANDARDS

Data content standards go further than both technical and metadata standards. Content standards define basic requirements of the data itself. They indicate what information must be collected, demanding the use of a controlled vocabulary or specifying the format for recording a date, colour or measurement.<sup>78</sup> By controlling content elements, content standards facilitate semantic interoperability; datasets

---

<sup>76</sup> Dublic Core Metadata Initiative, "Dublic Core Metadata Element Set, Version 1.1," <http://dublincore.org/documents/dces/> (Accessed: 5 October, 2013).

<sup>77</sup> Richards, "From Anarchy to Good Practice," 30.

<sup>78</sup> Ibid., 28.

recorded and modelled to standard requirements share common meaning.<sup>79</sup> Shared and precise meaning enhances the comparability of datasets within digital spaces. Therefore, content standards are essential to the spread and preservation of data in computational archaeology. Content standards are also important for ensuring the archaeologists are recording the minimum amount of information required for a dataset.

Returning to the preservation of data, content standards can provide “quality control” of what information is recorded.<sup>80</sup> When material evidence is excavated, its context is destroyed; in this case, the minimum recording standards can ensure the collection of the most important information.<sup>81</sup> Sometimes, publications and underlying primary data are the only surviving elements of an archaeological site.<sup>82</sup> Content standards are therefore crucial to accurately record the documentation and preservation of human history. Once again, accurate documentation enables efficient semantic interoperability in digital spaces. In practice, a number of standards have been created for governing the technical, metadata and content rules for data, but none of the content standards have gained widespread use. This lack of data content standards impedes the semantic interoperability – the exchange of meaning - across archaeological datasets.

---

<sup>79</sup> Ibid.

<sup>80</sup> Ibid.

<sup>81</sup> Ibid.

<sup>82</sup> Babeu, *Rome Wasn't Digitised in a Day: Building a Cyberinfrastructure for Digital Classicists*. 67.

#### 1.3.4. THE LACK OF CONTENT STANDARDS

Although content standards are essential to the meaningful exchange of data, archaeology suffers from resistance to standardised practice. Resistance stems from the “not invented here syndrome.”<sup>83</sup> Furthermore, resistance arises within the glorification of individualised practice.<sup>84</sup> Frederick Limp argues:

The reward structure in archaeological scholarship provides a powerful disincentive for participation in the development of semantic interoperability and, instead, privileges the individual to develop and defend individual terms/structures and categories.<sup>85</sup>

As Limp indicates, archaeologists understand individualised practice as a key feature of intellectual enquiry. Subsequently, a culture of inertia has arisen. Archaeologists feel that content standards will hinder a culture in which data collection and analysis is shaped by the individual project.<sup>86</sup> The imposition of rigid processes of data collection, documentation and representation are understood to stifle intellectual enquiry; if a project is forced to adhere to an unsuitable standard for a research question or project circumstance, research is restricted.<sup>87</sup> Thus, content standards are stereotyped as

---

<sup>83</sup> Richards, "From Anarchy to Good Practice," 27.

<sup>84</sup> Frederick Limp, "Web 2.0 and Beyond, or on the Web Nobody Knows You're an Archaeologists," in *Archaeology 2.0: New Tools for Communication and Collaboration*, ed. Eric Kansa (Los Angeles: Cotsen Institute of Archaeology, 2011), 287.

<sup>85</sup> Ibid., 278.

<sup>86</sup> Richards, "From Anarchy to Good Practice," 27.

<sup>87</sup> Richards and Hardman, "Stepping Back from the Trench Edge," 168.

regulations that block research avenues.<sup>88</sup> Distrust in content standards continues to hinder the semantic interoperability of digital spaces.

Content standards are also impeded by the inherent nature of archaeological data. Archaeological research is incomparable among its humanistic neighbours in its reliance on varied types of data, interdisciplinary teams and methods. Projects usually produce what Harley et al. state as “a plethora of data types in archaeology, including three-dimensional artifacts, maps, sketches, moving and still images, flora and faunal assemblages, geological samples, virtual reconstructions, and field notes.”<sup>89</sup> The development of standards for recording and sharing these diverse types of data is a substantial undertaking.<sup>90</sup> Hypothetically, if a rigid standard format were required for data content in archaeology, most projects would fail to meet the standard because of their fragmented datasets.<sup>91</sup> Thus, diverse data hinders the agreement, implementation and success of content standards.

Although there are drivers behind diversity, the development and adoption of content standards must be pursued; content standards are vital for computational archaeology.<sup>92</sup> While metadata assists the discovery of data, content standards grant

---

<sup>88</sup> Limp, "Web 2.0 and Beyond, or on the Web Nobody Knows You're an Archaeologists," 226.

<sup>89</sup> Diane Harley et al., "Use and Users of Digital Resources: A Focus on Undergraduate Education in the Humanities and Social Sciences," (Center for Studies in Higher Education, UC Berkeley 2006), 30-31.

<sup>90</sup> Babeu, *Rome Wasn't Digitised in a Day: Building a Cyberinfrastructure for Digital Classicists*. 66.

<sup>91</sup> Kintigh, "The Challenge of Archaeological Data Integration," 6.

<sup>92</sup> Richards, "From Anarchy to Good Practice," 30.

access to data.<sup>93</sup> Implemented at the beginning of the data life cycle, content standards define universal parameters for data creation. Archives and repositories use these parameters to accurately compare resources and identify relevant data. Parameters also reduce the time spent mapping a dataset to an archival standard. As a result, content standardisation is essential to the future of computational archaeology. Synthetic digital archives and repositories have the potential to advance long-term, large-scale understandings of human history through the sophisticated integration of data.<sup>94</sup> Without content standards, the sophisticated identification of patterning and comparison of data beyond local stages of research cannot happen.<sup>95</sup> Content standards that control idiosyncratic, fractured terminology will be particularly beneficial to semantic interoperability.

#### 1.4. FRACTURED TERMINOLOGY

The use of fractured terminology represents a key challenge to semantic interoperability.<sup>96</sup> Fractured terminology is a result of the idiosyncratic nature of terminology. More specifically, single terms are having multiple definitions (polysemy) and multiple terms are used to convey the same definition (synonyms). For example, “artefact scatter” is a polysemous archaeological term. Burke, Smith and Zimmerman define the term broadly:

---

<sup>93</sup> Jon Holemen, Christian-Emil Ore, and Øyvind Eide, "Documenting Two Histories at Once: Digging into Archaeology" (paper presented at the CAA 2004, Beyond the Artefacts: Digital Interpretation of the Past, 13-17 April 2004), 5.

<sup>94</sup> Kansa, "Openness and Archaeology's Information Ecosystem," 498.

<sup>95</sup> Richards, "From Anarchy to Good Practice," 29.

<sup>96</sup> Richards and Hardman, "Stepping Back from the Trench Edge," 169.

Artefact scatters can range in size from a few artefacts to a high-density scatter containing a wide range of different artefact types...Stone artefact scatters occur both as surface concentrations of material and as stratified deposits...<sup>97</sup>

In contrast, the Frankston Bypass project report defines the term on very specific parameters:

Artefact Scatter: A scatter of stone artefacts which is defined as being the occurrence of one (1) or more items of cultural material within 100 linear metres, with a distance of no greater than 20m between each item.<sup>98</sup>

Evidently, the term “artefact scatter” is defined for individual projects, rather than widespread understanding. While fractured terminology may function for the individual, the data should be uploaded to a repository for wider use. At the repository, problems are encountered; the term is rendered as site or project-specific, resulting in conflicting definitions.<sup>99</sup>

Digital repositories draw comparisons between datasets using shared, defined definitions.<sup>100</sup> Data created using fractured terminology can commonly have little or no relevance to data outside of its localised sphere of research. Thus, fractured terminology

---

<sup>97</sup> Heather Burke, Claire Smith, and Larry Zimmerman, *The Archaeologist's Field Handbook*(Crows Nest: Rowman Altamira, 2009), 203.

<sup>98</sup> Andrew Long & Associates Pty Ltd: Archaeological and Heritage Consultants, "Appendix 6: Areas of Cut and Fill," in *Frankston Bypass EES Project*(Southern and Eastern Integrated S31 Transport Authority2008), 140.

<sup>99</sup> Kintigh, "The Challenge of Archaeological Data Integration," 1.

<sup>100</sup> Heiler, "Semantic Interoperability," 271.



hinders semantic interoperability. Comparisons between disparate datasets can be done, but only after extensive and time-consuming semantic mapping.<sup>101</sup>

Fractured terminology is a serious obstruction to the research potentials of digitally available, high-quality archaeological datasets.<sup>102</sup> Standardised terminology, implemented early in the data life cycle, will largely eradicate idiosyncratic terminology use. Terminology standards are sets of terms—common to domain-specific vocabularies—used for the creation of synthetic data.<sup>103</sup> By controlling terminology use, content standards define datasets by a uniform, shared meaning, aiding semantic interoperability.<sup>104</sup>

Although the benefits are clear, progress towards terminology standards in archaeology has been remarkably slow. Discussed earlier, the imposition of rigid documentation and representation processes are understood to restrict individuality and intellectual enquiry.<sup>105</sup> Further research is required to better understand, and potentially overcome, the widely suffered inertia. A survey titled “Vocabularies, Concepts and Standards” was distributed in June of 2013 in order to confront this problem. The survey sought to explore terminology use and varying opinions of terminology standards in the Australian archaeological community.

---

<sup>101</sup> Richards and Hardman, "Stepping Back from the Trench Edge," 186.

<sup>102</sup> Kintigh, McManamon, and Spielmann, "Enhancing Data Comparability and Enabling Synthesis with Tdar (the Digital Archaeological Record)," 3.

<sup>103</sup> Hunter, "Enhancing the Semantic Interoperability of Multimedia through a Core Ontology," 49.

<sup>104</sup> Richards and Hardman, "Stepping Back from the Trench Edge," 169.

<sup>105</sup> *Ibid.*, 168.

#### 1.4.1. VOCABULARIES, CONCEPTS AND STANDARDS SURVEY

The “Vocabularies, Concepts and Standards” survey, distributed between 12<sup>th</sup> of June and 8<sup>th</sup> of July 2013, explored archaeological understandings, experiences, practices and opinions vis-à-vis terminology standards. Both qualitative and quantitative in nature, the survey had respondents share their experiences with fractured terminology and how this impacted fieldwork, research and publication. Additionally, the survey investigated the development and use of preferred language, core concepts and professional vocabularies. From these experiences, opinions arose about the need for standardised terminology and the willingness of the archaeological community to compromise their own practices to produce semantically interoperable data.

The survey was active for five weeks and gathered data from 117 respondents; 88% of the group worked or studied in Australia at the time of the survey. In general, the survey delivers diverse, but insightful, data on archaeological terminology. Subject areas include, but are not limited to, the development of terminology, reference materials and the use of recording systems. Most importantly, the results signpost strong resistance and distrust of terminology standards. The “Vocabularies, Concepts and Standards” survey is presented and discussed in chapters two and three.

### 1.5 CONCLUSIONS

As illustrated by chapter one, archaeological research is severely hindered by the limitations of the data created by the discipline. The crisis stems from the idiosyncratic nature of archaeological practice, the diversity of data, and the lack of data standards within the discipline. More specifically, the ad hoc nature of data collection and its structure hinders semantic interoperability. To enable sophisticated

dissemination and preservation of archaeological data, it is clear that content standards need to be developed and adopted.

By defining the basic requirements of the data itself, data can be meaningfully and accurately shared. The implementation of terminology standards is particularly vital to this goal. Fractured terminology represents one of the key impediments to synthetic data. Julian Richards boldly states that if we fail to achieve standards implementation “then our data is damned and can safely deleted at the end of the project as it is of no use to man nor beast.”<sup>106</sup> Thus, the long-term integration of archaeological data through content standards is crucial for the future of computational archaeology.

---

<sup>106</sup> Richards, "From Anarchy to Good Practice," 33.

## CHAPTER TWO: THE “VOCABULARIES, CONCEPTS AND STANDARDS” SURVEY

Circulated between the 12<sup>th</sup> of June and 8<sup>th</sup> of July 2013, the “Vocabularies, Concepts and Standards” survey conducted a study of fractured terminology and standards. In particular, the survey explored archaeological understandings, experiences, practices and opinions regarding terminology standards. The survey was developed using LimeSurvey software, an open source survey application.<sup>107</sup> The software enables users to create and publish on-line surveys without knowledge of any programming. It also collects the responses, and provides basic statistics of the survey results. For my analysis, the survey responses were downloaded onto an Excel spreadsheet (see section 2.3 for results processing and analysis) With a total of 34 questions, the survey took respondents 20-25 minutes to complete. The survey was distributed under the umbrella of the Federated Archaeological Information Management Systems (FAIMS) project of the University of New South Wales.

The FAIMS project is in the final stages of constructing a comprehensive information system for archaeology, including a mobile field application. The project aims to improve the digital creation, exchange and preservation of archaeological data. The system allows data from field and laboratory work to be created using digital mobile devices, processed in local databases, and exchanged via an online repository.<sup>108</sup> One of the components of this infrastructure is a mobile application, which can create semantically interoperable datasets.<sup>109</sup> The mobile application streamlines the creation

---

<sup>107</sup> <https://www.limesurvey.org/en/>

<sup>108</sup> Ross et al., "Creating Eresearch Tools for Archaeologists," 113.

<sup>109</sup> Ibid.

and management of fieldwork data by standardising fieldwork documentation.<sup>110</sup> Using the application, the FAIMS project aims to render archaeological data commensurate for advanced synthesis and accessibility. The examination of terminology standards, therefore, is vital for the FAIMS project. The “Vocabularies, Concepts and Standards Survey” catered to FAIMS interests, and was therefore distributed to their contact list via email. The Australian Archaeological Association (AAA) and OzArch also promoted the survey.

Since the research supports the FAIMS project, the survey primarily targeted the Australian archaeological community. To reach the most Australian participants, the survey was promoted to members of Australian Archaeological Association (AAA) over Facebook and email, and to the followers of the OzArch Google group. Overall, 117 respondents—103 of which work or study in Australia—took part in the survey. Relative to similar surveys, a large dataset was obtained. For example, the UK HE Research Data Management Survey only received 38 responses.<sup>111</sup> As the foundation to chapter three—the analysis of results—this chapter presents limitations experienced during the creation and distribution of the survey, followed by an introduction to the survey respondents. It also illustrates the analysis methodology used during the results processing and analysis.

---

<sup>110</sup> Ibid.

<sup>111</sup> "UK HE Research Data Management Survey," Loughborough University Research Data Management Project, <https://docs.google.com/forms/d/1M2AY1y9pCv2OAibqtnyWJZnM6880fBBxLv5xcehbZlC/viewform> (Accessed: 13 October, 2013).

## 2.1. LIMITATIONS

Although a high quantity dataset was obtained, the data suffers from a few limitations. To begin, the eight-month constraint of an honours degree allows a limited amount of time to create and perfect a survey before it is distributed, completed and analysed. Although the survey was created with a fellow honours student, Georgia Burnett, it was the first time we had developed and conducted an academic survey. We were to learn how to create effective surveys, how to use the Lime Survey software and, most importantly, word questions appropriately to elicit the most informative answers; the way respondents answer questions affects the way data appears in the results.

At the time of circulation, we were unaware that various questions requested too much information. For example, question 14 asks respondents to cite specific resources used for education (see appendix 1: question 14). These questions proved to be a burden on the respondents. They tended to skip the questions as a result. Additionally, we anticipated that some participants, such as students, would be unable to answer all the questions. Thus, the questions were not compulsory. These two factors yield many incomplete surveys, and gaps in the data emerge. The statistical significance of the survey responses suffered, especially when the analysis of relationships between results was attempted (discussed in section 2.4).

Despite these shortcomings, many valuable insights emerged from the survey data with regards to archaeological terminology and data interoperability. These insights arose on the basis of a diverse and highly educated survey demographic.

## 2.2. SURVEY RESPONDENTS

In an aim to characterise the survey population, the survey asked demographic questions about the age, location, education and working environments of the respondents. The simplicity and ease of the questions allowed for every respondent to answer the demographic questions. Question styles ranged between multiple-choice and open-ended text boxes (see appendix 1: question 1-11).

With a sizeable Australian demographic within the survey group (88%) and a dramatic expansion of the cultural heritage management (CHM) sector in Australia over the last four decades, it is not a surprise that 40% of respondents are CHM workers.<sup>112</sup> The remainder of the respondents include students (12%), university employees (34%), employees of government agencies (8%) and “other” (6%). “Other” includes people who identified themselves as information technology workers, miners and publishers (see appendix 2: section 2.1, table 1).

A majority of respondents are highly educated, with 42% accredited with a PhD and 21% with a master’s degree (see appendix 2: section 2.1, table 2). Of these two qualifications, 22% of those with a doctorate and 52% of those with a masters degree work in heritage management, while 67% of those with a PhD and 8% of those with an MA work in a university setting. Of those without a higher education degree, 28% of survey population hold an honours degree. In contrast, those with a bachelors degree represent 7% (see appendix 2: section 2.1, table 2). Of those with an honours degree, 67% work in cultural heritage management, 9% are still studying and 15% work in government agencies.

---

<sup>112</sup> Sean Ulm, Stephen Nichols, and Cameo Dalley, "Mapping the Shape of Contemporary Australian Archaeology: Implications for Archaeology Teaching and Learning," *Australian Archaeology* 61(2005): 11.

As well as education, the survey asked for how many years respondents have worked in their profession (see appendix 1: question 5). Most archaeologists have claimed 6-10 years of experience in the field (71%), followed by 1-5 years of experience (17%). Respondents with over 20 years experience in the field represent 13% of the contributors and those with 16-20 years 9% (see appendix 2: section 2.1, table 3). These participants have expertise in a range of geographic areas, chronological periods, sub-disciplines, methods/techniques and site types/material culture.

Question 7 asks, “In what geographic area(s) would you consider yourself an expert?” Out of 105 respondents, answers range from simply listing a country to listing states, cities and specific regions. Australian geographic areas are the most common, including the Kimberley, the Hunter Valley and Pilbara. For international participants, examples include Thrace, the United Arab Emirates, Cyprus, Greece, Papua New Guinea, the Pacific and Egypt (see spread sheet: columns Q-S for question 7).

Through question 8, it is evident the survey respondents have expertise in a range chronological periods (see appendix 1: question 8). Of the 104 participants, the most common response is the Holocene period (32%), the Pleistocene (24%) and the Late Holocene (17%). The Bronze Age, Colonial period (Australia) and Historic period (Australia) combined totalled 8% of the responses (spread sheet: columns T-V for question 8). A total of 112 respondents provided as many as three sub-disciplines of expertise.

Moving onto question 9, the survey asks, “In what sub-discipline(s) would you consider yourself an expert?” The most common sub-discipline is Indigenous archaeology (48%), followed closely by Historical archaeology (29%) and Marine archaeology (5%). These sub-disciplines are conventional for Australian archaeology.<sup>113</sup>

---

<sup>113</sup> Ibid., 14.



A further 32% of respondents say their expertise is in CHM; these professionals still study indigenous, historical and maritime archaeology. Other areas of expertise include the Archaeological Sciences, Classical archaeology, Industrial archaeology and Lithology (see spread sheet: columns W-Y for question 9).

A total of 107 respondents answered question 10, “In what method(s) or technique(s) would you consider yourself an expert?” The most common responses were surface survey (61%) and excavation (57%). The remaining response groups—methods and techniques such as GIS, landscape archaeology, research, lithic analysis and use wear analysis—appeared in no more than 11% of the total respondents (see spread sheet: columns Z-AD, question 10). In contrast to methods and techniques, the final demographic question had respondents list site types and/or material culture.

Question 11 asked, “In what site type(s) and/or material culture(s) would you consider yourself an expert?”. Of the 100 participants, lithic analysis was the most common response (43%). Rock shelters, ceramics and shell middens were regular responses as well, ranging between 18-21% of the survey group for each. Rare responses include urban sites, faunal remains and shipwrecks (see spread sheet: columns AE-AI, question 11). Although questions 1 to 11 were generic and boring for some respondents, the questions were vital for characterising the body of survey respondents.

Questions 1 to 11 identify a diverse and highly qualified survey group. Heritage management workers represent the largest group at 40%, but university and government agency employees are adequately represented as well. Respondents are highly educated, with 63% of them either holding a PhD or Masters degree. The demographics indicate wide exposure to both academic archaeology, and fieldwork practices; fractured terminology is mutual to all archaeological environments. Those within the academic sector are more likely to be exposed to archaeological literature,

while those in CHM are more likely to be exposed to terminology used within fieldwork documentation.

The students should not be forgotten either. Representing 12% of the survey population, students are particularly useful for understanding the development of core concepts and terminology. These respondents are currently building and developing their use of concepts and terminology. Therefore, students are more likely to be exposed to resources that build archaeological vocabularies.

The “Vocabularies, Concepts and Standards” survey respondents are a diverse and insightful group. Varied ages, experience, education and expertise allow for the exploration of diverse understandings of terminology and standards. As an analysis of the “research/teaching corpus” and “concepts and terminology” questions will show, the survey respondents are particularly insightful for exploring social factors as impediments to content standards in archaeology. These insights arose the basis of a manual spreadsheet analysis.

### 2.3. RESULT PROCESSING AND ANALYSIS

The results of the survey surfaced on the basis of manual spreadsheet analysis of individual questions and a query technique based on a set of hypotheses. The Excel spreadsheet can be downloaded from <<http://hdl.handle.net/102.100.100/13674>>. Each column within the spreadsheet represents an individual survey question. Within each column, the responses to the question reside. For example, question 1 “In which industry have you been primarily employed?” has five answers; “university”, “heritage management”, “government institution”, “I am still studying” and “other” (see appendix 1: question 12). Within the spreadsheet, column B contains the 117 responses to

question 1. Using these columns, individual and cross-question analyses were conducted.

Individual questions were analysed by counting different entities within each column. For example, question 12 – “Do any reference materials directly influence the concepts and terminology of your publications or reports?” - is a “yes” or “no” question. To analyse question 12, one can simply count how many answered “yes” (49) and how many answered “no” (13) within the column cells (see spreadsheet: column AC). These figures are then translated into percentages based on how many people participated in the question.

These analyses were coupled with more complex queries about how answers to questions relate to one another. It is important to observe connections between demographic samples; we can learn a lot about respondent understandings, attitudes and opinions. Demographics include, but are not limited to, employment, age, education and opinion. An example of how we can predict answers to question based on previous knowledge includes:

Respondents who are currently studying are more willing to compromise over the use of terms and concepts in their own practice to achieve a reconciliation of terms

Cross referencing questions 1 and 26 can help hypothesise the answer a majority of students will give to question 26 – “Are you willing to compromise over the use of terms and concepts in your own practice to achieve a reconciliation of terms?” (see appendix 3 for all hypotheses). In this case, we hypothesise that students *are* more willing to compromise terms. To test this, a manual cross-referencing technique within the spreadsheet was implemented. The method for this process is as follows.

To prove or disprove hypotheses made about demographic groups, entities within one column of the spreadsheet were compared to entities within another column. That is, one set of answers was compared with another. For example, the following hypothesis requires the comparison of the answers from question 24 and 25:

A majority of respondents who consider there to be a lack of standardised vocabulary in archaeological literature/reports also consider the impact of the lack of standardisation to be a major issue.

By cross referencing those who said “yes” in column BU to those who said “major” in column BV, we can understand the relationship between the two measured phenomena. In this example, the results were hung; 49% on both sides (1 respondent did not consider there to be an impact), and hypothesis was technically disproved (see chapter three, section 3.1 for further analysis of this result). A series of patterns emerged from the survey data using this query technique. The strength of these relationships was statistically tested using Pearson’s chi-square test of independence.

## 2.4. PEARSON’S CHI-SQUARE TEST OF INDEPENDENCE

While hypothesis analysis can identify insightful patterns concerning archaeological attitudes and understandings, it is important to evaluate their reliability. Significant relationships may be clear, but are they strong, moderate or weak relationships? <sup>114</sup> Using Pearson’s chi-square test of independence, the statistical significance of thirteen trends have been tested (see appendix 3 for hypotheses). Pearson’s chi-square test is a statistical analysis that calculates the distance between,

---

<sup>114</sup> "Statistical Significance," StatPac, <http://www.statpac.com/surveys/statistical-significance.htm> (Accessed: 15 October, 2013).

and independence of, two variables.<sup>115</sup> That is, chi-square tests the association between two measured phenomena. Chi-square proves or disproves the null hypothesis (the theoretical assumption that two measured phenomena share no relationship).<sup>116</sup> Thus, the null hypothesis was used as the baseline assumption for the results of the survey.

To prove or disprove the null hypothesis, assumptions about the survey results were tested using chi-square via a 2 x 2 contingency table (see appendix 3 for chi-square tests). A contingency table concerns the frequency distribution of data.<sup>117</sup> In the case of the survey, contingency tables incorporate the data of question variables. For example, question 12, "Do any reference materials directly influence the concepts and terminology of your publications or reports?", allows for a "yes" or "no" answer. Therefore, variables are the sample groups who answered "yes" and those who answered "no." Contingency tables cross-classify this data according to two categories, X and Y, each of which have two subcategories.<sup>118</sup> In other words, the components of two questions are incorporated within a 2 x 2 table. The chi-square method tests the distance between these variables, and therefore the relationships between them.

---

<sup>115</sup> David W. Stockburger, "Chi-Square and Tests of Contingency Tables," Missouri State University, <http://www.psychstat.missouristate.edu/introbook/sbk28m.htm> (Accessed: 15 September, 2013). ; A contingency table is a type of frequency table that displays the frequency distribution of two variables.

<sup>116</sup> J.H. McDonald, *Handbook of Biological Statistics: Basic Concepts of Hypothesis Testing (2nd Ed.)*(Baltimore, Maryland: Sparky House Publishing, 2009).

<sup>117</sup> Frequency Distribution: The organization of raw data in table form with classes and frequencies

<sup>118</sup> Richard Lowry, "2x2 Contingency Table," <http://www.vassarstats.net/tab2x2.html> (Accessed: 15 September, 2013).

Chi-square presents a P value of each contingency table query. A P-value is the estimated probability of rejecting the null hypothesis.<sup>119</sup> In other words, a P-value is the extent to which two measured phenomena share a relationship. If the P-value is less than 0.05 then the null hypothesis is rejected and the sample gives reasonable evidence that a relationship exists. A P-value of 0.05 implies that there is a 5% chance of being incorrect.<sup>120</sup> If the P-value is more than 0.05 then it is considered “statistically insignificant.”<sup>121</sup> Statistical insignificance means that the data is somewhat less reliable, but not unusable.

The chi-square analyses, conducted on thirteen survey hypotheses, proved the null hypothesis to be correct in twelve out of thirteen cases (see appendix 3 for hypotheses). In other words, most questions are independent and share no convincing relationships; they are statistically insignificant. Insignificant data stems from an incomplete dataset caused by respondents skipping questions (see chapter 2: section 2.1 for dataset limitations). The larger and more complete a dataset is, the more likely it is that statistically significant relationships will arise. Statistical issues have called for the revision and redistribution of the survey in 2014 to overcome problems experienced during the 2013 circulation.

Despite statistical issues, the patterns within the survey data are not without value. Statistical significance does not dictate a “meaningful” or “important” difference. It is for the researcher to assess the data’s relevance to a real world problem.<sup>122</sup>

---

<sup>119</sup> "P Values," [http://www.statsdirect.com/webhelp/content/basics/p\\_values.htm](http://www.statsdirect.com/webhelp/content/basics/p_values.htm) (Accessed: 10 October, 2013).

<sup>120</sup> Ibid.

<sup>121</sup> Ibid.

<sup>122</sup> Ibid.

Although relationships within in the data are statistically insignificant, they should not be dismissed. Some participants responded to every question and offered their own insight where room for comments was provided, while others skipped questions altogether. Therefore, the analysis will explore the way individuals and groups responded; it is not a statistical exercise. If we assess the survey responses on the basis of their content, a range of insightful data can be gleaned from the survey results.

With a total of 34 questions, the survey was split into three sections; “demographic”, “research/teaching corpus” and “concepts and terminology.” These sections gathered specific information about terminology use and terminology standards. Chapter two introduces the survey respondents via the “demographic” questions.

## CHAPTER THREE: SOCIAL FACTORS AND TERMINOLOGY STANDARDS

Social factors are key impediments to the adoption of content standards in archaeology. A current lack of agreement, leadership and cooperation within the archaeological community continues to hinder the implementation of standards.<sup>123</sup> The benefits of content standards are clear, yet they are persistently repelled; content standards enable integrated, synthetic and interdisciplinary research online by representing data within a universal schema.<sup>124</sup> Through a selection of data from the “Vocabularies, Concepts and Standards” survey, chapter three analyses socio-technic challenges to standards. The survey indicates a number of social factors in operation, including a divergence of archaeological attitudes and understandings of terminology standards. On the basis of this analysis, chapter three assesses the extent to which common ground can be met between interoperability and idiosyncratic practice.

Initially, the results demonstrate widespread understanding of a lack of terminology standards. Upon further examination, problems in translating specific concepts suggest a potential lack of understanding concerning the need for terminology standards. The “not invented here” syndrome also arises; several respondents believe their own projects are unique and deserve unique treatment.<sup>125</sup> The “not invented here syndrome” has been acknowledged but understudied by researchers.<sup>126</sup>

---

<sup>123</sup> Kintigh, McManamon, and Spielmann, "Enhancing Data Comparability and Enabling Synthesis with Tdar (the Digital Archaeological Record)," 3.

<sup>124</sup> Ibid.

<sup>125</sup> Richards, "From Anarchy to Good Practice," 27.

<sup>126</sup> Ibid.



Previously undetected attitudes are also evident. The survey reveals that many survey respondents exclude themselves from the use of non-standard terminology. Instead of protecting their right to use idiosyncratic terminology, they blame others for disparate practice. Participants also blame government legislation for terminology use, thereby excusing themselves from disparate practice. Another novel attitude is a “we can work with it” mentality, where archaeologists are willing to live with problematic terminology. Although these attitudes pose a challenge for the implementation of standards, the survey is optimistic for the future of content standards as well. Participants are deeply concerned about idiosyncratic terminology and willing to compromise their practice in order to achieve a reconciliation of terminology. Through an exploration of the “research/teaching corpus” and “concepts and terminology” questions, chapter three assesses the extent to which some of the challenging social factors can be overcome.

### 3.1. PARTICIPANT UNDERSTANDINGS OF FRACTURED TERMINOLOGY

If we want to look at how different archaeologists understand terminology and standards, the “Vocabularies, Concepts and Standards” survey is particularly insightful. A lack of agreement, leadership and cooperation within the archaeological community is strongly associated with the awareness of standards. If archaeologists are unaware of idiosyncratic terminology and its effect on digital archaeology, they are less likely to cooperate with standards. If archaeologists are aware, they may be more inclined to adopt a solution. To unpack different understandings of terminology standards within the survey responses, this section focuses on question 24.

Of every question in the survey, question 24—“Do you think there is a lack of standardised vocabulary in archaeological literature/reports?”—sheds the most light on

respondent perception of standards. The question forms two important demographic populations for chapter three; a group who believes the discipline needs standards, and a group that does not. These demographics can be compared to a variety of different questions to grasp how respondents understand terminology standards.

Of the 63 participants who answered question 24,, 70% claim there is a lack of standardised vocabulary in archaeological literature and/or reports. A further 17% claim there is no lack of standardised vocabulary, and 13% are not sure (see appendix 2: section 2.1, table 11). The results indicate a highly educated and standard-aware population, who have personally experienced the problems of idiosyncratic terminology in the course of their careers. What we do not grasp from question 24 is how serious they believe the issue is: cross-referencing question 24 with question 25 answers this query.

Question 25—“If you answered yes to the question above, how serious would you consider the impact of the lack of standardisation on archaeology as a discipline?”—is informative about the how serious the issue is to respondents. The calculated results are even, with 49% claiming a lack of standards has a “major” impact and 49% claiming a lack of standards “minor” has impact on the discipline. Only one honours student said there is no real impact concerning a lack of data standards (see appendix 2; section 1, table 12). Through a cross-reference of question 24 and 25, it is evident that 50% of those who agree there is a lack of standardised vocabulary also consider the impact to be a major issue. Evidently, most participants are aware of the lack of controlled terminology use, and the group largely understands the seriousness of the issue. These results are both compounded and disproved when question 24 is cross-referenced with other survey questions.

A comparison of questions 24 and 32 indicates confusion about the lack of controlled terminology in archaeology. We will begin with an individual overview of

question 32, then move into a comparison. Question 32 was an array style inquiry in which participants state if they strongly agree, agree, moderately agree, moderately disagree, disagree or strongly disagree with four statements about core concepts and terminology. The statements are as follows:

- **32a**—I think that archaeologists across all sub-disciplines share common core concepts and terminology
- **32b**—When I read archaeological literature across all sub-disciplines, I see a commonly shared bank of concepts and terminology
- **32c**—I think that archaeologists in my sub- discipline(s) share common core concepts and terminology
- **32d**—When I read archaeological literature in my sub-discipline(s), I see a commonly shared bank of concepts and terminology

The statements and results of question 32 comprise figures 1 to 4 and will be discussed in two groups: those who *agreed* to any extent and those who *disagreed* to any extent. As identified by figures 1 to 4, the majority of participants agree with every statement.

Illustrated within figure 1, 88% of respondents agree with the statement, “I think that archaeologists *across all sub-discipline(s)* share common core concepts and terminology.” Similarly, figure 2 identifies that 87% agree with the statement, “When I read archaeological literature *across all sub-disciplines*, I see a commonly shared bank of concepts and terminology.” These results can be understood as mistaken considering each sub-discipline of archaeology is characterised by diverse knowledge, expertise, skills, problems, challenges, inquiries, approaches and research areas. Although discipline does share core concepts and terminology, for example the term “context” or “site”, each sub-discipline operates with its own definitions. Terminology is commonly disparate. On reflection, the survey statement needed to be more specific. It should have

stated, “I think that archaeologists *across all sub-discipline(s)* share common core terms and definitions,” rather than “core concepts and terminology.”

Figure 1. Question 32a

*I think that archaeologists across all sub-disciplines share common core concepts and terminology*

Response	#	%
Strongly Agree	7/64	11
Agree	19/64	30
Moderately Agree	30/64	47
Moderately Disagree	3/64	5
Disagree	3/64	5
Strongly Disagree	2/64	3

Figure 2. Question 32c

*I think that archaeologists in my sub- discipline(s) share common core concepts and terminology*

Response	#	%
Strongly Agree	7/63	
Agree	26/63	
Moderately Agree	25/63	
Moderately Disagree	2/63	
Disagree	2/63	
Strongly Disagree	1/63	

Question 32 also asked respondents about common core concepts and terminology within sub-disciplines. Illustrated in figure 3, 92% of respondents agree with the statement, “I think that archaeologists *in my sub-discipline(s)* share common core concepts and terminology.” Similarly, figure 4 identifies that 95% of the respondents agree with the statement, “When I read archaeological literature *in my sub-discipline(s)*, I see a commonly shared bank of concepts and terminology.” Respondents

believe they do not practice idiosyncratically within their own sub-disciplines. Thus, question 32 shows a significant lack of awareness about the disparity of terminology.

Figure 3. Question 32b

*When I read archaeological literature across all sub-disciplines, I see a commonly shared bank of concepts and terminology*

Response	#	%
Strongly Agree	5/62	8
Agree	20/62	32
Moderately Agree	29/62	47
Moderately Disagree	5/62	8
Disagree	3/62	5
Strongly Disagree	0/62	0

Figure 4. Question 32d

*When I read archaeological literature in my sub-discipline(s), I see a commonly shared bank of concepts and terminology*

Response	#	%
Strongly Agree	7/62	11
Agree	27/62	44
Moderately Agree	25/62	40
Moderately Disagree	3/62	
Disagree	0/62	0
Strongly Disagree	0/62	0

A lack of general understanding, and an element of confusion, is also evident by cross-referencing questions 32a and 32b with question 24. Of those who answered both questions 24 and 32a, 84% both agree there is a lack of standardised vocabulary and archaeologists across all sub-disciplines share common core concepts and terminology (see appendix 3: hypothesis 9 for chi-square). These responses are contradictory. Participants acknowledge the lack of terminology standards in archaeology, but also

believe the discipline shares common terminology. A similar result is evident in a comparison of questions 24 and 32b. A high 81% of respondents who agree with a lack of standardised vocabulary also see commonly shared core concepts and terminology when they read archaeological literature across the discipline (see appendix 3: hypotheses 10 for chi-square). In theory, this response is contradictory. Respondents who state “yes” in question 24 should theoretically disagree with the statements in 32a and 32b. It is possible that respondents translated questions incorrectly. To overcome this potential issue during the next distribution of the FAIMS survey, the questions and detail given will be more explicit in the next circulation. Another possibility is that responses may have arisen out of little understanding of content standardisation.

A comparison between questions 24 and 27 also indicates contradictory responses. Question 27 asks, “To what extent would you consider your own lexicon to be part of a wider, accepted vocabulary within your sub-discipline(s)?” Illustrated in figure 5, 94% of respondents said that either “all” or “most” of the terms within their own lexicons are part of a wider accepted vocabulary. Of those who answered questions 24 and 27, 84% of respondents who say there is a lack of standardised vocabulary consider their own lexicon to be part of a wider, accepted vocabulary within their sub-disciplines (see appendix 3: hypothesis 6 for chi-square). Respondents either misunderstood the difference between “terminology standards” and “wider, accepted vocabularies,” or they do not consider themselves as part of the problem. The self-exclusion of participants from idiosyncratic terminology use will be discussed in depth throughout section 4.5. Largely, questions 24 and 27 identify misunderstandings in just how common idiosyncratic terminology use is. In contrast, questions 24 and 26 provide strong evidence for wide understanding of the issue and more importantly, a willingness to change.

Figure 5. Question 27

Response	#	%
All of the terms within my own lexicon are part of a wider, accepted vocabulary	18/61	29
Most of the terms within my own lexicon are part of a wider, accepted vocabulary	39/61	64
Some of the terms within my own lexicon are part of a wider, accepted vocabulary	3/61	5
Few of the terms within my own lexicon are part of a wider, accepted vocabulary	1/61	2
None of the terms within my own lexicon are part of a wider, accepted vocabulary	0/61	0

Question 26—“Are you willing to compromise over the use of terms and concepts in your own practice to achieve a reconciliation of terms?”—offers valuable insight into the willingness of respondents to adopt terminology standards. Of the 65 respondents who participated, 69% claimed they would compromise their personal use of terms for a reconciliation of terms. In contrast, respondents who will not compromise their use of terms make up 14%. Those who are “not sure” represent 17% of the survey group (see appendix 2: section 2.1, table 13). The optimism of this result is heightened when questions 24 and 26 are compared.

Through an analysis of question 24 and 26, it is clear that 68% of respondents who are willing to compromise over their use of terms also believe a lack of standardised vocabulary is prevalent within archaeology (see appendix 3: hypothesis 5 for chi-square). A community who identify a lack of terminology standards, while showing a willingness to change, can only be beneficial when implementing standards.

Simultaneously, we cannot expect that *every* archaeologist is willing to change his or her practice. The “not invented here syndrome” impedes the adoption of content standards, even if archaeologists are aware of the negative effects of idiosyncratic practice.

### 3.2. THE “NOT INVENTED HERE” SYNDROME

The “not invented here” (NIH) syndrome, also known as the “special snow flake syndrome,” is an understanding that every archaeological project is unique and therefore deserves unique treatment.<sup>127</sup> Archaeologists tend to invent new practices for individual needs rather than adopt pre-existing ones.<sup>128</sup> Constantly “reinventing the wheel” is problematic because it leads to the creation of disparate resources. The literature has acknowledged the NIH syndrome, but as a key impediment to the adoption of standards the attitude requires further research. The NIH syndrome manifests within questions 18 and 26.

Question 18 asks, “Do you design or manage recording systems? If yes, consider your key field and attributes. Which of the following statements best describes your situation?” The statements and results are illustrated in figure 6. A total of 38% of respondents say they create new recording systems. The significance of this outcome is that survey respondents believe their projects are unique and deserve unique treatment. If the respondents create new recording systems, rather than adopting pre-existing ones, they generate disparate resources. Without established recording

---

<sup>127</sup> Ibid.

<sup>128</sup> Scott Byrant, “Overcoming the ‘Not Invented Here’ Syndrome: Experience with Sourcing Educational Multimedia Developed Elsewhere” (paper presented at the Flexibility: The Next Wave, Proceedings of ASCILITE’98, 1998), 121.



systems—and content standards—repositories do not accurately compare and identify relevant data.<sup>129</sup>

We can find a bit of consolation in the remaining 30% of respondents who claim they borrow from a well-established system developed elsewhere. These respondents use pre-existing recording systems for their research. In other words, they may create resources that can be meaningfully shared through semantic interoperability online. Our relief is diminished once again when we look at question 26.

Figure 6. Question 18

Response	#	%
I use a well-established system developed elsewhere	12/86	14
I borrow from a well-established system developed elsewhere	26/86	30
I use a less-established system developed elsewhere	3/86	4
I borrow from a less established system developed elsewhere	6/86	7
I create new recording systems	33/86	38
I inherit recording systems from an unknown source	6/86	7

Discussed earlier, question 26 asked respondents, “Are you willing to compromise over the use of terms and concepts in your own practice to achieve a reconciliation of terms?” A total of 69% of respondents stated they would compromise their personal use of terms (see appendix 2, section 2.1, table 13). Although the answers to the question indicate a clear willingness to adopt standards, the “not invented here”

<sup>129</sup> Richards, “From Anarchy to Good Practice,” 30.

syndrome prevails in the participant comments.

Along with a basic “yes”, “no” or “not sure” inquiry, question 26 allowed for participants to comment. Of the 65 people who participated, 20 people added further comments. Several respondents were enthusiastic about changing their practice. For example, one person said, “Yes, for sure,” while another stated, “Providing I am sure that other companies are using the same terms and concepts, it would not be an issue” (see appendix 2: section 2.3, Q 26, ID 82 and 10). One participant also highlighted the lack of agreement about terminology standards, a major theme of the thesis; “Happy to use an accepted and clear term where it can be shown to be more widely adopted—however, consensus is often lacking” (see appendix 2: section 2.3, Q 26, ID 49). Although encouraging comments are apparent, the NIH syndrome is also evident.

Participant ID 116, who was “not sure” whether to compromise his/her practice, is clearly fearful of the stereotyped “one-size-fits-all straightjacket” of standards.<sup>130</sup> The respondent said:

In theory yes, but in practice I suspect that there would be areas in which I do not want to use a standardised vocabulary that departs from usage in my subfield or geographic area, or that does not accurately describe my material (especially in terms of chronological periods or ceramic types). (See appendix 2: section 2.3, ID 116.)

The participant fears his/her own unique data will be affected by standardisation. Although standards are clearly beneficial to digital archaeology, the participant does have legitimate reasons to be cautious. Content standards can impose rigid processes of data collection, documentation and representation. If a project is forced to adhere to an

---

<sup>130</sup> Bryn Nelson, “Empty Archives,” *Nature* 461, no. 10 (2009): 162.

unsuitable standard for a research question or project circumstance, research can become restricted.<sup>131</sup> Various other participants have similar outlooks to participant ID 116.

Respondents state, “Depends on what I am being asked to compromise, and how much I disagree with the terms and concepts”, “If I feel it is appropriate” and “I have worked collaboratively with other researchers and have shifted definitions and terms if it makes sense and is for the betterment of the project” (see appendix 2: section 2.3, ID 15, 107 and 3). Finally, participant ID 116 was the most sincere:

I think it's worth reiterating -- because I suspect I'm not alone in this -- that although I favour standardisation in theory and am strongly convinced of the importance of consistency in terminology for data sharing, I think that in practice it's unrealistic to expect archaeologists across the discipline to adopt standardized terminologies imposed from elsewhere unless there is substantial and immediate benefit (just look at the different terms different projects insist on applying for the idea of archaeological context). (See appendix 2: section 2.3, Q. 35, ID 116.)

Although these comments reject rigid content standards, they offer important insights. The comments indicate that in order to successfully implement content standards, the standards need to allow for the contextual and individual nature—that is, the personal idiosyncrasies—of archaeological projects. Only then will archaeologists, such as these respondents, be happy to adhere to content standards. Exactly how this may be achieved will be introduced in section 3.7. Simultaneously, the self-exclusion of participants from idiosyncratic practice is challenging for any standards implementation.

---

<sup>131</sup> Richards and Hardman, "Stepping Back from the Trench Edge," 168.

### 3.3. SELF-EXCLUSION

It is undeniable that archaeology is poorly supplied with controlled vocabularies, thesauri or ontologies.<sup>132</sup> At domain specific levels, there are standard terms but they only relate to discrete phenomena.<sup>133</sup> Outside of these domains, a wide variety of terminologies are used, causing confusion and misunderstanding. At the same time, archaeologists deny that they themselves are contributing to the problem. The survey revealed that many survey respondents exclude themselves from the use of non-standard terminology. They blame other archaeologists—specifically, those outside of their own sub-discipline(s)—for disparate practice, but exclude themselves from the need to adopt standards. Such self-exclusion is a significant concern for implementing any form of content standards. Simultaneously, the respondents confirm the proliferation of different terminologies. They identify that there are differences in the terms we use; there are a lot of competing standards, and it results in confusion. We will begin by exploring the understood proliferation of terminologies—conveyed within questions 30 and 31—and move into the opposing social factor of self-exclusion.

Question 30 asks, “How often do you disagree with the terminology used in fieldwork, research or writing in your chronological periods, methodological practices, sub-disciplines and outside their sub-disciplines?” As identified by figures 7 to 10, the respondents primarily “often” and “always” disagree with terminology used in their chronological periods (56%), methodological practices (65%), sub-disciplines (63%) and out-side of their disciplines (65%). Without asking for specific details, question 30 is sufficient indication of personal experiences with idiosyncratic terminology. If the

---

<sup>132</sup> Kilbride, "The Danube in Prehistory in the Digital Age: Towards a Common Information Environment for European Archaeology," 136.

<sup>133</sup> Ibid.

survey group predominantly “often” and “sometimes” disagree with terminology, they recognise that fractured terminology arises within many aspects of their archaeological practice. The results of question 31 convey similar inferences.

Figure 7. Question 30a

*In your chronological period (s)*

<b>Response</b>	<b>#</b>	<b>%</b>
Always	0/57	0
Often	9/57	13
Sometimes	28/57	42
Rarely	14/57	21
Never	6/57	9

Figure 8. Question 30b

*In your methodological period(s)*

<b>Response</b>	<b>#</b>	<b>%</b>
Always	0/61	
Often	7/61	
Sometimes	36/61	
Rarely	16/61	
Never	2/61	

Figure 9. Question 30c

*In your sub-discipline(s)*

<b>Response</b>	<b>#</b>	<b>%</b>
Always	0/61	0
Often	6/61	19
Sometimes	36/61	59
Rarely	17/61	28
Never	2/61	3

Figure 10. Question 30d

*Outside of your sub-discipline(s)*

Response	#	%
Always	0/56	0
Often	5/56	9
Sometimes	31/56	55
Rarely	14/56	35
Never	5/56	9

The respondents also identify the proliferation of diverse terminology in their responses to question 31. Question 31 states, “When I work with other colleagues, I have to learn new terms to fit in with their workflow.” Of the 61 participants, most respondents allege it is “often” or “sometimes” necessary to learn new terms (67%). A further 7% claim they “always” need to learn new terms (see appendix 2: section 1, 24). Once again, question 31 is a sufficient indication of personal experiences with idiosyncratic terminology. If the survey group primarily “often” and “sometimes” have to learn new terms, the respondents once again confirm the proliferation of different terminologies. The respondents also blame other archaeologists—principally outside of their own sub-discipline(s)—for disparate practice, and exclude themselves from the need to adopt standards. Participant self-exclusion begins with a comparison of question 24 and 32.

Question 24--“Do you think there is a lack of standardised vocabulary in archaeological literature/reports?”—and question 32 once again shed light on social impediments to standards (see appendix 1: question 24). Question 32 was an array style query in which participants state if they strongly agree, agree, moderately agree, moderately disagree, disagree or strongly disagree with four statements about core concepts and terminology (see appendix 1: question 32). The statements are as follows:

- **32a**—I think that archaeologists across all sub-disciplines share common core concepts and terminology
- **32b**—When I read archaeological literature across all sub-disciplines, I see a commonly shared bank of concepts and terminology
- **32c**—I think that archaeologists in my sub- discipline(s) share common core concepts and terminology
- **32d**—When I read archaeological literature in my sub-discipline(s), I see a commonly shared bank of concepts and terminology

Once again, the analysis of these queries will be discussed in two groups: those who *agreed* to any extent and those who *disagreed* to any extent. Statements 32c and 32d indicate self-exclusion by the participants from the wider problem of fractured terminology.

Of those who answered questions 24 and 32c, 91% of respondents agree that there is a lack of standardised terminology but also believe that archaeologists in their own sub-disciplines share common core concepts and terminology (see appendix 3: hypothesis 11 for chi-square). Furthermore, a comparison of questions 24 and 32d indicate that 93% of participants who recognise a lack of standardised terminology also agree that when they read archaeological literature in their own sub-disciplines, they see a commonly shared bank of concepts and terminology (see appendix 3: hypothesis 12). Evidently, respondents acknowledge that idiosyncratic terminology is apparent in archaeology, but do not see it within their own sub-disciplines. On a micro level, many archaeologists feel that they practice appropriately; they understand their own research data and how it relates to their own localised sphere of research. However, understanding project specific data does not suggest that their own sub-disciplines share common core concepts and terminology. Most projects use their own systems of

data creation, organisation and analysis.<sup>134</sup> If core vocabularies were used and adhered to, archaeological data would benefit from greater means of interoperability. Participant self-exclusion via the blame of others emerges from the comparison of questions 16 and 17.

Archaeology is a collaborative discipline requiring datasets and reports to be shared across teams of researchers. Consequently, meaningful definitions should be provided for others in order to understand both interpretations and raw data. Question 16 asked respondents if they provide meaningful definitions for key terms in their publications and/or datasets (see appendix 1: question 16). Illustrated in figure 11, 74% stated that they either “often” or “always” provide meaningful definitions for key terms. A further 20% claimed they “sometimes” provide meaningful definitions, while 6% “rarely” provide definitions. When these results are combined with that of question 17, participants are shown to believe it is other archaeologists who do not define their terminology.

Figure 11. Question 16

Response	#	%
Always	30/74	40
Often	25/74	34
Sometimes	15/74	20
Rarely	5/74	6
Never	0/74	0

Question 17 asked respondents if *other* archaeologists provide meaningful definitions for key terms in their publications and/or datasets (see appendix 1: question

<sup>134</sup> Kansa and Bissell, "Web Syndication Approaches for Sharing Primary Data in "Small Science" Domains," 42.



17). The results illustrated in figure 12, indicate that only 27% claimed that other archaeologists “always” or “often” provide meaningful definitions for key terms. In its place, 58% of respondents claimed that other archaeologists only “sometimes” provide definitions. Comparing questions 16 and 17, 55% of respondents who claim they “often” or “always” provide meaningful definitions for key terms say that other archaeologists only “sometimes” or “rarely” provide meaningful definitions. As the only statistically significant relationship (calculated using Pearson’s chi-square test of independence), the blame of others for a lack of defined terminology is clearly evident (see appendix 3: hypothesis 16 for chi-square). If archaeologists exclude themselves from poor practice, while attributing poor practice to others, the discipline will continue to suffer a lack of semantic interoperability. Archaeologists will reject terminology standards if they do not believe they need standards. Blame of others is also apparent within the cultural heritage management and government institution demographics of the survey.

Figure 12. Question 17

Response	#	%
Always	0/74	0
Often	20/74	27
Sometimes	43/74	58
Rarely	10/74	14
Never	1/74	1

### 3.4. GOVERNMENTAL CONTROL OF TERMINOLOGY USE

The legislative feature of cultural heritage management and governmental agencies has led to a shared understanding among various survey participants that the government controls their use of terminology. For example, one participant stated, “I work in government - the terms I use are fixed and defined by legislation and cannot be changed (easily)” (see appendix 2: section 2.2, Q 26, ID 76). Another participant agreed,

stating:

While it is not ideal, and practitioners should in fact have the freedom to undertake archaeological investigation according to methods that they can reasonably justify, consulting practice is increasingly prescribed by Codes and Guidelines. (See appendix 2: section 2.3, Q 35, ID 6.)

Various respondents believe it is the government's responsibility to control terminology use, and furthermore, it is their role to change archaeological practice within their systems. In reality, cultural heritage management and government employees have more freedom with terminology use than they believe. The Australian government is an adequate example of this freedom.

In many areas of Australian cultural heritage management and government agencies, standard measures and processes are widely used and promoted. They encourage the predictability needed for quality, sharing and transparency of archaeological data.<sup>135</sup> These standards concern, for example, the reporting of dates or the collection of data samples.<sup>136</sup> Although this is the case, the Australian government recognises that archaeology cannot be amendable to the type of absolute, objective standards that apply to other industries, such as the biomedical sciences. It is understood that standards can give "emphasis to 'objective' scientific measures and too little to cultural values and difficult-to-measure processes such as individual and community attachment to place and land".<sup>137</sup> Consequently, standardised typologies do

---

<sup>135</sup> Denis Byrne, "Standards for Archaeological Practice in Aboriginal Heritage Management," ed. NSW National Parks and Wildlife Service (Hurstville: NSW National Parks and Wildlife Service, 1997), 3.

<sup>136</sup> Ibid.

<sup>137</sup> Ibid., 4.

not exist within Australian governmental archaeology.<sup>138</sup> Thus, participant ID 32 can be considered mistaken when he/she stated, “terminology and concepts are clearly prescribed by the state regulator (OEH) in their guideline documents” (see appendix 2: section 2.2, Q. 26, ID 32). The lack of terminology control within the Australian government can be exemplified through the Victorian state government.

The Victorian government designed a number of component forms intended for the documentation of archaeological finds. These guides manage how archaeologists conduct investigations and interpret results, but they do not maintain a strict use of terminology. Consider the “artifact scatter” component form in appendix 4. The archaeologist can individually nominate any sized area of scatter, its dimensions and density. They can then freely label the material and artifact type based on how they define the components. Without defined and prescribed terms, the form does not significantly restrict vocabulary use. As it stands, government legislation controls the documentation of metadata more so than it does the documentation of content.

### 3.5. NOBODY’S PROBLEM

Whilst examining the issue of data accessibility in the article “Empty Archives,” one of the interviewees states, “It’s everybody’s problem and therefore nobody’s problem.”<sup>139</sup> The quote is incredibly relevant to the issue of data accessibility in archaeology. If data standards implementation were exclusive to only one sub-discipline, a solution would be less difficult to discover and implement. Unfortunately, data inaccessibility is a problem of the entire discipline. The survey results identify a “we can work with it” attitude, in which archaeologists who do experience difficulties

---

<sup>138</sup> Ibid., 3.

<sup>139</sup> Nelson, “Empty Archives,” 162.

with terminology are willing to live with the problem. The “we can work with it” attitude has arisen from the larger issue Nelson highlights, where the problem is mutual across the discipline and seems too large to be able to deal with.

Question 25—“How serious would you consider the impact of the lack of standardisation on archaeology as a discipline?”—is useful for exemplifying the “we can work with it” mentality (see appendix 1: question 25). One individual states, “Not a major issue, more of a sometimes issue” (see appendix 2: section 2.3, Q. 25, ID 35). Another respondent adds:

Heritage consultants in Western Australia produce large data sets each year. While many concepts are shared, terminologies can be troublesome. Harnessing the value of these data sets in research is made problematic, but not impossible, because of a lack of standardisation.

(See appendix 2: section 2.3, Q 25, ID 86.)

Evidently, some respondents feel that although harnessing the value of datasets through terminology can be difficult at times, it is not impossible. This “we can work with it” attitude is an impediment to standards implementation. Archaeologists who are only occasionally affected by fractured terminology will not see an urgent need to standardise practice. Thus, the attitude fosters the creation of disparate data and continues to scatter data across un-integrated archives and repositories.<sup>140</sup> While identifying a “we can work with it attitude”, question 25 also provides for additional insights on the importance of terminology standards to archaeology.

---

<sup>140</sup> Ibid., 160.

### 3.6. HOW SERIOUS IS A LACK OF STANDARDS?

So far chapter three has primarily explored social factors that continue to challenge the implementation of standards. In contrast, there were a number of supportive responses manifest within the results. Question 25—“How serious would you consider the impact of the lack of standardisation on archaeology as a discipline?”—sheds the most light on standards-aware and enthusiastic respondents (see appendix 1: question 25).

Generally, a number of respondents are aware of the negative effects of fractured terminology. For example, one respondent is aware that fractured terminology presents a barrier to the spread and comparison of data:

A lack of standardisation produces disparate recording systems which are hard to understand/reconcile by people evaluating, comparing or analysing data collected in the field or by research. (See appendix 2: section 2.3, Q 25, ID 108.)

Participant ID 58 is also aware of these translation issues. More specifically, the respondent understands that access to data doesn't necessarily mean access to the meaning of the data:

Confusion over the meaning of certain terms prevents other archaeologists from fully understanding the nature of their assumptions. (See appendix 2: section 2.3, Q 25, ID 58.)

As a result of project-specific terminology, archaeologists generally need to understand terms and concepts foreign to their own before they can comprehend the data itself. Amongst this confusion, a respondent understands that data can be isolated or lost; “In a development context, the lack of standardisation affects what people identify as worth

saving.” (See appendix 2: section 2.3, Q 25, ID 40.) For digital repositories, the exchange of information must occur meaningfully and accurately via semantic interoperability.<sup>141</sup> If accurate meaning is not controlled, data will become isolated and site-specific, with little or no relevance to other data.

Other respondents were more specific, drawing connections to digital archaeology: “Lack of standardisation is an issue in linking data from different datasets with different platforms so that they can be discovered or combined for new analysis” (see appendix 2: section 2.3, Q 25, ID 116). The respondent observes how serious a lack of terminology standards can be to digital archaeology.<sup>142</sup> More specifically, the respondent has identified the need for semantic interoperability to improve analysis. Another respondent specifically highlights the need for standardised terminology in Australian archaeology:

In Australia, a standardised reporting and recording system as well as vocabulary is urgently needed. Likewise an online repository for archaeological data is required. More importantly, a mandate that requires all the documentary and cultural archives to be deposited and made publically available is required. (See appendix 2: section 2.3, Q 25, ID 79)

The respondent identifies the various goals of the Federated Archaeological Information Management Systems (FAIMS) project. Using a custom-made field application, the project aims to incorporate a descriptive but controlled vocabulary for data creation. Furthermore, the project aims to preserve the data within a digital repository. The results of question 34 are optimistic for the goals of the FAIMS project as well.

---

<sup>141</sup> Heiler, "Semantic Interoperability," 271.

<sup>142</sup> Kubicek, Cimander, and Scholl, "Layers of Interoperability," 85.

The final question of the survey, asks, “When does the standardisation of terminology become most important?” Illustrated in figure 13, respondents selected one of six options. The results indicate that the majority of respondents identify data creation as a crucial time for terminology standards (61%). A high preference for “at data creation” by a largely Australian respondent group is an optimistic result for the FAIMS project. The project aims to standardise fieldwork practice, and indeed terminology, at data creation through their mobile application. Standardised terminology at data creation will assist in rendering archaeological data commensurate for advanced synthesis and accessibility online.

Figure 13. Question 34

Response	#	%
At Data Creation	38/62	61
At Data Analysis	5/62	8
During Writing	12/62	19
During Editing	1/62	2
At Publication	3/62	5
Other	3/62	5

Although the survey data indicates resistance to the implementation of standards, it is evident that a series of positive and insightful opinions were obtained from the results as well. These opinions give projects such as FAIMS hope that archaeologists are beginning to understand the importance of standards. Simultaneously, strong resistance to terminology standards manifests within the survey population. The need for contextually aware content standards, rather than rigid processes of documentation and recording, is obvious.

### 3.7. DISCUSSION

The “Vocabularies, Concepts and Standards” survey provides contradictory results. While the respondents understand the need for terminology standards, majority communicate a deep resistance to terminology standards. Furthermore, they are distrustful to any standards that do not cater to their individual circumstances and needs.<sup>143</sup> These contradictory attitudes represent a key social impediment to the adoption of standardised vocabularies. Inertia overcomes the leadership and cooperation required for terminology standards, which is a considerable concern for the future of digital archaeology.<sup>144</sup>

The survey results clearly communicate the challenging and unavoidable nature of social factors. Social resistance to standards is extremely difficult to overcome. So far there have been relatively few success stories for the use of standards in archaeology; the bio-archaeological standard for recording human skeletal remains is an exception rather than a rule.<sup>145</sup> The stalemate situation will continue if common ground, is not met between interoperability and archaeological practice. The final solution should balance the desires of each party, just as a survey participant points out:

I'd be careful on how standards get implemented. Standards can be hugely helpful, but they need to be contestable. Publishing, revising, contesting, etc. explicit machine-readable standards needs to be a much more regular part of research. It requires new modes of scholarly communication, since books and journals can't do it. Standards are not

---

<sup>143</sup> Kintigh, McManamon, and Spielmann, "Enhancing Data Comparability and Enabling Synthesis with Tdar (the Digital Archaeological Record)," 3.

<sup>144</sup> Ibid.

<sup>145</sup> Ibid.



something that should ever be frozen for all time, since that would mean we've run out of interesting research problems. (See appendix 2: section 2.3, Q 35, ID 112.)

How do we come up with a standard that is not “frozen for all time”? It is the proposal of this thesis that an archaeological ontology is a flexible standard that will allow individualised practice and semantic interoperability to co-exist. Ontologies are “the representation of the meaning of terms in vocabularies and the relationships between those terms.”<sup>146</sup> In other words, ontologies take into account an array of terms—basic, technical and dynamic—used by archaeologists and the relationships between those terms. They also represent terms in relation concepts such as classifications, typologies or domains. Mapping idiosyncrasies in terminology to core ontologies during data creation can allow for both freedom of expression and semantic interoperability. As a solution for contextually aware content standards, ontologies will be explored in greater detail in the concluding chapter.

Having assessed the survey results, it is clear that various attitudes aggravate the crisis of data accessibility. More specifically, the survey reveals the challenges Australian digital archaeology faces, as 88% of the survey group work or study in Australia. That said, Australian archaeology has a few considerable advantages in comparison to other places around the globe; it has a collectively small professional body and a high awareness of the lack of standards. Furthermore, there are only a handful of training institutions and a small body of archaeologists working within the government. Therefore, the potential to consult, agree upon, implement, enforce and regulate contextually aware, individualised standards is promising.

---

<sup>146</sup> McGuinness and Harmelen, "Owl Web Ontology Language."

## CONCLUSION AND PROSPECTS FOR FUTURE RESEARCH

The creation of the World Wide Web has had an overwhelming impact on the way archaeological information can be shared, stored and reused. With increasing quantities of archaeological information becoming available online, Martin Doerr asserts that there is now an “increasing demand for global searches, comparative studies, data transfer and data migration”.<sup>147</sup> Comparative studies are possible if data is structured and defined using data standards.<sup>148</sup> If data standards are not utilised, the accurate exchange and spread of information is hampered. In practice, a number of standards have been created for governing the technical, metadata and content rules for archaeological data, but content standards have not gained widespread support or use. Semantic interoperability, which relies on content standards, is significantly compromised as a result. Identified by the “Vocabularies, Concepts and Standards” survey, a key impediment to the adoption of content standards is resistance within the archaeological community itself.

Although survey respondents express concern for the effects of idiosyncratic terminology, the majority view terminology standards as a threat to their practice. Content standards are perceived as constraints on the data collection and analysis culture of an individual project.<sup>149</sup> Investigated throughout chapter three, resistance surfaces within various attitudes and understandings of terminology standards.

---

<sup>147</sup> Martin Doerr, "The Cidoc Conceptual Reference Model: An Ontological Approach to Semantic Interoperability of Metadata," *AI Magazine* 24.32003, 1.

<sup>148</sup> Baines and Brophy, "What's Another Word for Thesaurus? Data Standards and Classifying the Past," 237.

<sup>149</sup> Richards, "From Anarchy to Good Practice," 27.

Participants either adhere to the “not invented here syndrome”—the formulation that every project is unique and deserves unique treatment—or blame others for idiosyncratic practice. Moreover, participants declare they can continue to work with fractured terminology, even if it is occasionally troublesome. Although these attitudes are problematic for content standards, the disinterest in standards felt by the survey respondents is not without justification. So far, efforts to create semantically interoperable data have concentrated on the standardisation of shared data structures and controlled terminology.<sup>150</sup> These standards have enforced rigid processes of data collection, documentation and representation.<sup>151</sup> Many archaeologists understand these data standards as restrictive policies, which stifle intellectual enquiry and research avenues.<sup>152</sup> Nevertheless, archaeologists still need to find a coherent way of dealing with the two-edged sword of terminology use.

Given the realities identified by the “Vocabularies, Concepts and Standards” survey, it is clear that content standards need to embrace the diversity of data, research questions and archaeological practice. The content standards need to allow for personal idiosyncrasies, while adhering to an enforced standard. Instead of *prescribing* a common schema for capturing the meaning of data, semantic interoperability should be harnessed through the use of descriptive ontologies. By defining a useful and open ontology, the universe of variables concerning terminology use can be embraced alongside freedom of practice.

---

<sup>150</sup> Doerr, "The Cidoc Conceptual Reference Model," 1.

<sup>151</sup> Richards and Hardman, "Stepping Back from the Trench Edge," 168.

<sup>152</sup> *Ibid.*

Ontologies are descriptive, representational vocabularies often exclusive to a domain or subject matter.<sup>153</sup> More specifically, ontologies are networks or hierarchies of terminology shaped using objects, properties of objects and relations between objects.<sup>154</sup> Conceptualisations—groupings, classes and classifications—of terminology are captured by ontologies as well. For example, ontologies can group a network of terms as specific to the domain of lithics or material type such as obsidian.<sup>155</sup> These vocabularies are different to the standard thesaurus. Although thesauri group terms together according synonymous relationships, their static, hierarchical nature does not capture the relational aspect of ontologies. As a result, thesauri classify terms and list synonyms, but do not allow for nuances related to how synonymous terms can exist as valid alternatives. Ontological vocabularies are descriptive standards that account for polysemous, synonymous and conceptual entities of terms. By both accurately describing data and adhering to mutual definition, ontologies facilitate synthetic data sharing between researchers and computer systems.<sup>156</sup>

When used as a content standard, ontologies do not impose the rigid processes of data collection, documentation and analysis that archaeologists are generally afraid of. Ontologies map the structure of knowledge, rather than dictate structure.<sup>157</sup> Varied

---

<sup>153</sup> Balakrishnan Chandrasekaran, John R. Josephson, and V. Richard Benjamins, "What Are Ontologies and Why Do We Need Them?," *Intelligent Systems and Their Application* 14, no. 1 (1999): 20.

<sup>154</sup> Ibid.

<sup>155</sup> Ibid.

<sup>156</sup> Noy F. Natalya and Deborah L. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology," (Stanford: Stanford University, 2001), 1.

<sup>157</sup> Chandrasekaran, Josephson, and Benjamins, "What Are Ontologies and Why Do We Need Them?," 21.

terminologies can be mapped against one another with the help of an ontology, and better understood as a result. As a result, ontologies grant archaeologists the freedom of practice they wish for and sufficient access to the discourses they part of.<sup>158</sup>

Many disciplines have developed standardised ontologies, and use them to share and interpret information. The Unified Medical Language System (UMLS) is a leading example. Intended for the biomedical sciences, the UMLS is a comprehensive ontology of controlled vocabularies. The biomedical sciences also practice idiosyncratically via the terminologies used in information sources.<sup>159</sup> Unless these sources are both human and machine-readable, they have little application to digital spaces and consequently, little relevance to user inquiry.<sup>160</sup> Donald Lindberg summarises objectives of the UMLS program:

The objective of this program is to solve what is the most fundamental barrier to the application of computers in medicine; namely, the lack of a standard language in medicine. We will attempt to build that vocabulary, a language that will cross between the biomedical literature and the observations on the patient, as well as the educational applications in the school, a language which allows those areas to be interrelated.<sup>161</sup>

The overarching goal of the UMLS is to enhance information systems for the biomedical

---

<sup>158</sup> Baines and Brophy, "What's Another Word for Thesaurus? Data Standards and Classifying the Past," 239.

<sup>159</sup> Betsy L. Humphreys et al., "The Unified Medical Language System: An Informatics Research Collaboration," *Journal of the American Medical Informatics Association* 5, no. 1 (Jan/Feb 1998): 1.

<sup>160</sup> Ibid.

<sup>161</sup> Ibid., 2.

sciences so that users can discover and use relevant information, even if the data is technically disparate.<sup>162</sup> The UMLS program is closely bound to the issues digital archaeology faces; the meaning of data must be first harnessed before users can access relevant information. Ontologies for archaeology have been created, but they are primarily intended for cultural heritage management, and are not suitable for field data collection. The CIDOC Conceptual Reference Model (CRM) project is a key example.

Intended for cultural management within museums, the CIDOC-CRM analyses common conceptualisations behind content and metadata structures to enable the spread, access and integration of data.<sup>163</sup> The project views ontologies as important entities for “dealing within knowledge from the past in any domain”.<sup>164</sup> The ontology itself is a “top-level” ontology; that is, it follows a top-down implementation pattern.<sup>165</sup> Although the CIDOC-CRM is a good example of how ontologies have been harnessed for archaeology, the ontology itself is not intended for data creation; it was designed to annotate pre-existing data (museum objects, for example) for discovery purposes. As a result, any implementation to standard archaeological field data would not be wise. The ontology requires too high a standard than that which is reasonably expected from a standard field project.

However, ontologies applied during fieldwork documentation are significantly more beneficial to semantic interoperability. Fieldwork ontologies eliminate the time spent semantically mapping pre-existing data and furthermore, data is defined

---

<sup>162</sup> Ibid., 9.

<sup>163</sup> Doerr, "The Cidoc Conceptual Reference Model," 1.

<sup>164</sup> Ibid.

<sup>165</sup> Ibid., 2.

accurately. The Federated Archaeological Information Management Systems (FAIMS) project aims to make fieldwork ontologies a reality for Australian archaeology.

With an objective to cater to both idiosyncratic practice and sophisticated semantic interoperability, the FAIMS project will incorporate ontologies within their custom-built fieldwork application. These ontologies will be specific to Australian archaeological domains such as Historical, Indigenous, Maritime and Industrial archaeology. Terminology will be descriptively mapped and defined by users of the application during data creation. As a result, FAIMS clients will be able to represent data in ways that are useful to their own projects and digital repositories. For the FAIMS project, ontologies are not implicit representations of data, but a flexible tool that can make an effort towards reconciling the “archaeological tower of babel” currently in existence in Australian archaeology.<sup>166</sup>

The future of this thesis aligns within the objectives of the FAIMS project. As a pilot project for descriptive ontologies I plan to generate a contextually aware, human-mediated reconciliation of terminology. To do this, implicit epistemologies of archaeological terms will be explored within the context of their texts. Text is a rich source of information for interpreting and describing the behaviour, values, norms and practices of a writer or their subject.<sup>167</sup> Thus, text analysis is useful for extracting how terminology is used and defined. Within the future research, archaeological texts will be analysed using computer-assisted-textual-analysis (CATA). CATA is a computerised method for discovering meaning in text through the automatic analysis of machine-

---

<sup>166</sup> Richards, "From Anarchy to Good Practice," 27.

<sup>167</sup> Melina Alexa, *Computer Assisted Text Analysis Methodology in the Social Sciences*(Mannheim: ZUMA Arbeitsbericht, 1997), 3.

readable text.<sup>168</sup> To generate a contextually aware reconciliation of terms, concordance analysis—cross-references between different parts of a sentence or phrase—will be conducted using recursive partitioning text analysis.

Recursive partitioning is a text-mining process designed to grasp the contextual nature of small portions of text.<sup>169</sup> Sequences of text, relevant to a term under investigation, are collapsed into smaller, summarised accounts of the larger text.<sup>170</sup> In other words, passages of text are translated into one-line summaries. These summaries provide contextual information about the term under investigation. One cannot understand the use of a term from the term alone. The relationship between a term and the text around a term is important for definition. For example, “artefact scatter” could be contextualized in a paragraph as a small area of artefact concentration, a large or dispersed area of concentration, an entity within a unit of a site, or an entire site. The ties between terms and definitions are not explicit, but implicit, linked semantically through the sentences and paragraphs of the text.<sup>171</sup> The context of one term can be compared with other uses of the term, in order to understand the diversity of terminology use and meaning among archaeologists.

---

<sup>168</sup> R. P. Weber, "Computer-Aided Text Classification : Advantages , Difficulties , and Types," *Qualitative Sociology* 126-147(1984): 127.

<sup>169</sup> Brian Ballsun-Stanton, "Asking About Data: Exploring Different Realities of Data Via the Social Data Flow Network Methodology" (The University of New South Wales, 2012), 409.

<sup>170</sup> Vineet Chaoji, Apirak Hoonlor, and Boleslaw K. Szymanski, "Recursive Data Mining for Role Identification," in *Proc. 5th International Conference on Soft Computing as Transdisciplinary Science and Technology*(Paris, FranceOctober 27-31, 2008), 3.

<sup>171</sup> Eleanor T. Lewis, Jana Diesner, and Kathleen M. Carley, "Using Automated Text Analysis to Study Self-Preservation Strategies," in *Computational Analysis of Social and Organizational Systems (CASOS) Conference*(Pittsburgh PA2001), 1.



The recursive partitioning analysis will be coupled with a close engagement with the archaeological community. Further insight into the nature, understanding, acquisition and use of terminology by practicing archaeologists is important. As identified by the “Vocabularies, Concepts and Standards” survey, resistance is a key impediment to terminology standards. A reconciliation of terminology conducted alongside an engagement within the archaeological community will assist any effort to create an effective ontology. Engagement with practicing archaeologists will also assist areas in need of further research from the “Vocabularies, Concepts and Standards” survey; for example, understandings of the lack of terminology standards. Although there are plans for an improved and more concise survey, the “Vocabularies, Concepts and Standards” survey is an insightful dataset for terminology standards research.

Using the “Vocabularies, Concepts and Standards” survey results, this thesis investigated an important debate on terminology standards and synthetic research. Although idiosyncratic practice is a key feature of archaeology, it is also an enormous problem for data interoperability. Idiosyncratic terminology, in particular, hinders the meaningful exchange of information. To overcome project specific data, it is clear by the survey results that archaeology requires descriptive ontologies. These ontologies need to take into account contextual nature of archaeological practice and enquiry. Only then may the wider archaeological community be willing to adopt terminology standards. By reconciling descriptive ontologies, digital data will be rendered commensurate for advanced synthesis. Consequently, archaeological data will be more readily accessible as a result of semantic interoperability.



# APPENDIX

## APPENDIX ONE – SURVEY FORM

### FAIMS Project: Vocabularies, concepts, and standards survey



This survey will investigate what sources from which archaeologists draw their preferred language, core concepts, and professional vocabularies. More specifically, it is designed to gather information about influences on the use of language and vocabulary through questions regarding the literature produced and consumed by archaeologists. The results of this survey will be published in two honours theses, due in November 2013, and will contribute towards other publications of the FAIMS project (journal articles and book chapters).

Thank you for your time.

You are invited to participate in a vocabularies, concepts and standards survey produced by the Federated Archaeological Information Management Systems (FAIMS) project based in the School of Humanities, Faculty of Arts and Social Sciences, at the University of New South Wales (UNSW).

This survey has been created by Corinne Softley and Georgia Burnett, Honours Students at UNSW, in conjunction with the FAIMS project and under the supervision of Dr Shawn Ross, Senior Lecturer at UNSW, and Dr Adela Sobotkova, Research Fellow at UNSW. FAIMS is an eResearch Tools project funded by National eResearch Collaboration Tools and Resources (NeCTAR) program in Australia. Its aim is to develop digital infrastructure that would allow archaeological data to be created digitally and archived to online repositories. The project was launched on 5 June 2012 and is due to be completed by December 2013.

This survey is designed to gather information about what influences the development and use of preferred language, core concepts, and professional vocabularies among practicing archaeologists. More specifically, it is designed to gather information about influences on the use of language and vocabulary through questions regarding the literature produced and consumed by archaeologists. This data will be used to assess the relationship between field methods and practices, academic and commercial archaeological publication, and the emergence of archaeological vocabularies. **The results of this survey may be deployed in publications.**

We would appreciate your participation, and hope to collect responses from a broad range of archaeologists in a variety of different professional settings and sub-disciplines. There are no expected costs to the participant. Your consent is confirmed by the completion and return of this survey under the National Statement on Ethical Conduct in Research Guidelines, section 1.9.

This survey will take twenty to twenty-five minutes to complete, and is **active until 5th July 2013**. Participation is anonymous and all questions are non-mandatory, though we encourage you to be as thorough as possible. This survey has been approved under the Human Research Ethics Advisory Panel B of the University of New South Wales (reference number: 13 040).

Complaints may be directed to the Ethics Secretariat, The University of New South Wales, SYDNEY 2052 AUSTRALIA (phone 9385 4234, fax 9385 6648, email [ethics.sec@unsw.edu.au](mailto:ethics.sec@unsw.edu.au)). Any complaint you make will be investigated promptly and you will be informed of the outcome.

A summary of the research findings can be obtained from the chief research investigator on request. If you have any questions, please feel free to ask us. If you have any additional questions later, we will be happy to answer them.

Thank you for your time.

#### CONTACT DETAILS:

**Miss Georgia Burnett**

Email: [z3332180@student.unsw.edu.au](mailto:z3332180@student.unsw.edu.au)

**Dr Shawn Ross**

Email: [shawn.ross@unsw.edu.au](mailto:shawn.ross@unsw.edu.au)

**Miss Corinne Softley**

Email: [z3293545@student.unsw.edu.au](mailto:z3293545@student.unsw.edu.au)

**Dr Adela Sobotkova**

Email: [a.sobotkova@unsw.edu.au](mailto:a.sobotkova@unsw.edu.au)

There are 35 questions in this survey

## Demographic

### 1 In which industry have you been primarily employed? Please choose one of the following answers.

Please choose **only one** of the following:

- ☐ University
- ☐ Heritage Management
- ☐ Government Institution
- ☐ I am still studying
- ☐ Other

For example, you may have worked five years in a university setting, but since you have worked three years in heritage management. In this case, your principal occupation would be at the university.

### 2 What is the highest qualification you hold? Please choose one of the following answers.

Please choose **only one** of the following:

- ☐ Graduate Diploma
- ☐ Bachelors
- ☐ Honours
- ☐ Masters
- ☐ Doctorate
- ☐ Other

For example, if you have completed your Bachelors with Honours but are currently working on your Doctorate, your highest qualification is Honours.

### 3 What is the institution where you completed your highest qualification? Please enter in the text box below.

Please write your answer here:

For example, if you studied at the Australian National University for your Bachelors and La Trobe University for your Masters, then the institution of your highest qualification would be La Trobe University.

### 4 When did you receive your highest qualification? Please enter the year in the box below.

Please write your answer here:

**5 How many years have you worked in your profession?  
Please choose one of the following answers.**

Please choose **only one** of the following:

- ☐ I am currently studying.
- ☐ 1-5
- ☐ 6-10
- ☐ 11-15
- ☐ 16-20
- ☐ 20+

Please indicate the number of years you have worked as a professional archaeologist since you received your first qualification (e.g. Bachelors).

**6 What is the country and state of your organisation?  
Please enter in the text box below.**

Please write your answer here:

Please do not use abbreviations i.e. New South Wales, Australia *not* NSW, Australia.

**7 In what geographic area(s) would you consider yourself an expert?  
Please list up to three geographic areas.**

Please write your answer(s) here:

Please specify a region and country. For example: The Kimberley, Australia.

**8 In what chronological period(s) would you consider yourself an expert?  
Please list up to three chronological periods.**

Please write your answer(s) here:

For example: Pleistocene, Late Holocene, Neolithic, Bronze Age, etc.

**9 In what sub-discipline(s) would you consider yourself an expert?  
Please list up to three sub-disciplines.**

Please write your answer(s) here:


For example: Indigenous, Maritime, Historical, Industrial, Classical, Cultural Heritage Management, Archaeological Sciences, etc.

**10 In what method(s) or technique(s) would you consider yourself an expert?  
Please list up to five methods or techniques.**

Please write your answer(s) here:


For example: excavation, remote sensing, surface survey, geophysics, archaeobotany, zooarchaeology, etc.

**11 In what site type(s) and/or material culture(s) would you consider yourself an expert?  
Please list up to five site types and/or material cultures.**

Please write your answer(s) here:


For example: rock art, lithics, ceramics, faunal remains, graffiti, rock shelters, shell middens, etc.

## Research/Teaching Corpus

**12 Do any reference materials directly influence the concepts and terminology of your publications or reports?**

**Please select yes or no. Please name your most influential reference material in the comment section.**

Please choose **only one** of the following:

- ☐ Yes  
☐ No

**13 If you answered yes to the question above, what reference materials most influence the concepts and terminology that you use?**

**Please list up to five reference materials.**

**Only answer this question if the following conditions are met:**

Answer was 'Yes' at question '12 [03]' (Do any reference materials directly influence the concepts and terminology of your publications or reports? Please select yes or no. Please name your most influential reference material in the comment section. )

Please write your answer(s) here:


To the extent possible, please use this format: Author (Year), Title, Resource Type.

**14 If you teach archaeology (at a university, as part of induction of new employees into a cultural heritage firm, or in other settings), what reference materials do you recommend to your students or trainees? Please name up to five reference materials. Leave blank if not applicable.**

Please write your answer(s) here:


To the extent possible, please use this format: Author (Year), Title, Resource Type.

**15 Have you seen other archaeologists using reference material which you consider to be outdated?**

**Please name up to five reference materials. Leave blank if not applicable.**

Please write your answer(s) here:


To the extent possible, please use this format: Author (Year), Title, Resource Type.

**16 Do you provide meaningful definitions for key terms in your publications and/or datasets?**

**Please choose one of the following answers.**

Please choose **only one** of the following:

- ☐ Always
- ☐ Often
- ☐ Sometimes
- ☐ Rarely
- ☐ Never



**17 Do other archaeologists provide meaningful definitions for key terms in their publications and/or datasets?**

**Please choose one of the following answers.**

Please choose **only one** of the following:

- ☐ Always
- ☐ Often
- ☐ Sometimes
- ☐ Rarely
- ☐ Never

**18 Do you design or manage recording systems? If yes, consider your key fields and attributes. Which of the following statements best describes your situation?**

**Please select up to three. Leave blank if not applicable.**

Please choose **all** that apply:

- ☐ I use a well-established system developed elsewhere
- ☐ I borrow from a well-established system developed elsewhere
- ☐ I use a less-established system developed elsewhere
- ☐ I borrow from a less-established system developed elsewhere
- ☐ I create new recording systems
- ☐ I inherit recording systems from an unknown source
- ☐ Other:

**19 If you answered the above question, and you borrow or adapt an existing system, please specify which system.**

Please write your answer here:

## Concepts and Terminology

### 20 What working environments have most influenced the development of your concepts and terminology?

Please number each box in order of preference from 1 to 7

<input type="text"/>	Institutional Training
<input type="text"/>	Fieldwork
<input type="text"/>	Individual Research
<input type="text"/>	Organisation/Company
<input type="text"/>	Professional Mentor
<input type="text"/>	Government Institution
<input type="text"/>	Other

Please drag the following items to the righthand box, starting with your highest ranking item, moving through to your lowest ranking item. **Only rank items that apply.**

### 21 If you included "other", please specify.

Please write your answer here:

**22 What do you think are the most valuable source(s) for generating and reconciling core concepts and terminology?**

Please number each box in order of preference from 1 to 9

<input type="text"/>	Handbooks
<input type="text"/>	Textbooks
<input type="text"/>	Monographs
<input type="text"/>	Journal Articles
<input type="text"/>	Recording Forms
<input type="text"/>	Reports
<input type="text"/>	Conference Papers
<input type="text"/>	Workshops
<input type="text"/>	Other

Please drag the following items to the righthand box, starting with your highest ranking item, moving through to your lowest ranking item. **Only rank items that apply.**

**23 If you included "other", please specify.**

Please write your answer here:

**24 Do you think there is a lack of standardised vocabulary in archaeological literature/reports?**

**Please choose one of the following answers.**

Please choose **only one** of the following:

- ☐ Yes
- ☐ No
- ☐ Not sure

**25 If you answered yes to the question above, how serious would you consider the impact of the lack of standardisation on archaeology as a discipline?**

**If you did not answer yes to the question above, please select 'no answer.'**

**Optional: can you think of a specific example where a lack of standardisation has been a significant issue? Please enter in comment section.**

Please choose **only one** of the following:

- ☐ Major
- ☐ Minor
- ☐ No real impact

Make a comment on your choice here:

**26 Are you willing to compromise over the use of terms and concepts in your own practice to achieve a reconciliation of terms?**

Please choose **only one** of the following:

- ☐ Yes
- ☐ No
- ☐ Not sure

Make a comment on your choice here:

**27 To what extent would you consider your own lexicon to be part of a wider, accepted vocabulary within your sub-discipline(s)?**  
**Please choose one of the following answers.**

Please choose **only one** of the following:

- ☐ All of the terms within my own lexicon are part of a wider, accepted vocabulary.
- ☐ Most of the terms within my own lexicon are part of a wider, accepted vocabulary.
- ☐ Some of the terms within my own lexicon are part of a wider, accepted vocabulary.
- ☐ Few of the terms within my own lexicon are part of a wider, accepted vocabulary.
- ☐ None of the terms within my own lexicon are part of a wider, accepted vocabulary.

**28 Please think of a disagreement you have had with other archaeologists over the use of a term. What was the term? Please list up to five terms.**

Please write your answer(s) here:

**29 Please describe in a few words an example where a disagreement over the definition of a term had a significant impact on research, analysis, or understanding. Please name the term, and provide a brief description. Leave blank if not applicable.**

Please write your answer here:

**30 How often do you disagree with the terminology used in fieldwork, research or writing:**

**Please choose one of the following answers for each question. \***

Please choose the appropriate response for each item:

	Always	Often	Sometimes	Rarely	Never	No answer
In your chronological period(s)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
In your methodological practice(s)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
In your sub-discipline(s)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Outside your sub-discipline(s)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**31 When I work with other colleagues, I have to learn new terms to fit in with their workflow.**

**Please choose one of the following answers.**

Please choose **only one** of the following:

- ☐ Always
- ☐ Often
- ☐ Sometimes
- ☐ Rarely
- ☐ Never

**32 Please select your stance on the following statements.  
Please choose one of the following answers for each statement.**

Please choose the appropriate response for each item:

	Strongly Agree	Agree	Moderately Agree	Moderately Disagree	Disagree	Strongly Disagree
I think that archaeologists <b>across all sub-disciplines</b> share common core concepts and terminology.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When I read archaeological literature <b>across all sub-disciplines</b> , I see a commonly shared bank of concepts and terminology.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I think that archaeologists <b>in my sub-discipline(s)</b> share common core concepts and terminology.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When I read archaeological literature <b>in my sub-discipline(s)</b> , I see a commonly shared bank of concepts and terminology.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**33 What three resources do you most frequently use to define concepts and/or terms?**

**Please list up to three resources. Please be as specific as possible.**

Please write your answer(s) here:

<input type="text"/>	
	<input type="text"/>
	<input type="text"/>

Consider online resources, textbooks, manuals, journals, government regulations, word of mouth etc.

To the extent possible, please use this format: Author (Year), Title, Resource Type.



**34 When does the standardisation of terminology become most important?  
Please choose one of the following answers.**

Please choose **only one** of the following:

- ☐ At data creation
- ☐ At data analysis
- ☐ During writing
- ☐ During editing
- ☐ At publication
- ☐ Other

## Further Comments

**35 Thank you for participating in our survey. If you have any further comments please enter them in comment box below. If you wish for a reply, further information or to be updated on the progress of this project, please leave an email address.**

Please write your answer here:

## APPENDIX TWO – SURVEY RESULTS

### 2.1. QUANTITATIVE RESPONSES

\*All percentages rounded to the nearest whole number

\*Only quantifiable survey results included

TABLE 1.

**Question 1. In which industry have you been primarily employed?**

Response	#	%
Heritage Management	47/117	40
University	40/117	34
Government Institution	9/117	8
Still Studying	14/117	12
Other	7/117	6

TABLE 2.

**Question 2. What is the highest qualification you hold?**

<b>Response</b>	<b>#</b>	<b>%</b>
Graduate Diploma	1/117	1
Bachelors	8/117	7
Honours	33/117	28
Masters	25/117	21
Doctorate	49/117	42
Other	1/117	1

TABLE 3.

**Question 5. How many years have you worked in your profession?**

<b>Response</b>	<b>#</b>	<b>%</b>
I am currently studying	14/117	12
1-5 years	20/117	17
6-10 years	19/117	16
11-15 years	11/117	9
16-20 years	15/117	13
20+	38/117	33

TABLE 4.

**Question 12. Do any reference materials directly influence the concepts and terminology of your publications or reports?**

<b>Response</b>	<b>#</b>	<b>%</b>
Yes	49/62	79
No	13/62	21

TABLE 5.

**Question 15. Have you seen other archaeologists using reference material which you consider to be out-dated?**

<b>Response</b>	<b>#</b>	<b>%</b>
Yes	6/15	40
No	9/15	60

TABLE 6.

**Question 16. Do you provide meaningful definitions for key terms in your publications and/or datasets?**

<b>Response</b>	<b>#</b>	<b>%</b>
Always	30/74	40
Often	25/74	34
Sometimes	15/74	20
Rarely	5/74	6
Never	0/74	0

TABLE 7.

**Question 17. Do other archaeologists provide meaningful definitions for key terms in their publications and/or datasets?**

<b>Response</b>	<b>#</b>	<b>%</b>
Always	0/74	0
Often	20/74	27
Sometimes	43/74	58
Rarely	10/74	14
Never	1/74	1

TABLE 8.

**Question 18. Do you design or manage recording systems? If yes, consider your key fields and attributes. Which of the following statements best describes your situation? Please select up to three. Leave blank if not applicable.**

<b>Response</b>	<b>#</b>	<b>%</b>
I use a well-established system developed elsewhere	12/86	14
I borrow from a well-established system developed elsewhere	26/86	30
I use a less-established system developed elsewhere	3/86	4
I borrow from a less established system developed elsewhere	6/86	7
I create new recording systems	33/86	38
I inherit recording systems from an unknown source	6/86	7

TABLE 9.

**Question 20. What working environments have most influenced the development of your concepts and terminology?**

<b>Response</b>	<b>#</b>	<b>%</b>
Institutional Training	20/65	31
Fieldwork	17/65	26
Individual Research	10/65	15
Organisation/Company	4/65	6
Professional Mentor	10/65	15
Government Institution	2/65	3
Other	2/65	3

TABLE 10.

**Question 22. What do you think are the most valuable source(s) for generating and reconciling core concepts and terminology?**

<b>Response</b>	<b>#</b>	<b>%</b>
Handbooks	7/61	11
Textbooks	9/61	15
Monographs	8/61	13
Journal Articles	12/61	23
Recording Forms	9/61	15
Reports	4/61	6
Conference Papers	1/61	2
Workshops	5/61	8
Other	4/61	7



TABLE 11.

**Question 24. Do you think there is a lack of standardised vocabulary in archaeological literature/reports?**

<b>Response</b>	<b>#</b>	<b>%</b>
Yes	44/63	70
No	11/63	17
Not Sure	8/63	13

TABLE 12.

**Question 25. If you answered yes to the question above, how serious would you consider the impact of the lack of standardisation on archaeology as a discipline?**

<b>Response</b>	<b>#</b>	<b>%</b>
Major	22/45	49
Minor	22/45	49
No Real Impact	1/45	2

TABLE 13.

**Question 26. Are you willing to compromise over the use of terms and concepts in your own practice to achieve a reconciliation of terms?**

<b>Response</b>	<b>#</b>	<b>%</b>
Yes	45/65	69
No	9/65	14
Not Sure	11/65	17

TABLE 14.

**Question 27. To what extent would you consider your own lexicon to be part of a wider, accepted vocabulary within your sub-discipline (s)?**

Response	#	%
All of the terms within my own lexicon are part of a wider, accepted vocabulary	18/61	29
Most of the terms within my own lexicon are part of a wider, accepted vocabulary	39/61	64
Some of the terms within my own lexicon are part of a wider, accepted vocabulary	3/61	5
Few of the terms within my own lexicon are part of a wider, accepted vocabulary	1/61	2
None of the terms within my own lexicon are part of a wider, accepted vocabulary	0/61	0

TABLE 15.

**Question 30. How often do you disagree with the terminology used in fieldwork, research or writing:**

**Please choose one of the following answers for each question.**

*a. In your chronological period (s)*

<b>Response</b>	<b>#</b>	<b>%</b>
Always	0/57	0
Often	9/57	13
Sometimes	28/57	42
Rarely	14/57	21
Never	6/57	9

*b. In your methodological period(s)*

<b>Response</b>	<b>#</b>	<b>%</b>
Always	0/61	
Often	7/61	
Sometimes	36/61	
Rarely	16/61	
Never	2/61	

*c. In your sub-discipline(s)*

---

<b>Response</b>	<b>#</b>	<b>%</b>
Always	0/61	0
Often	6/61	19
Sometimes	36/61	59
Rarely	17/61	28
Never	2/61	3

---

*d. Outside of your sub-discipline(s)*

---

<b>Response</b>	<b>#</b>	<b>%</b>
Always	0/56	0
Often	5/56	9
Sometimes	31/56	55
Rarely	14/56	35
Never	5/56	9

---

TABLE 16.

**Question 31. When I work with other colleagues, I have to learn new terms to fit in with their workflow**

Response	#	%
Always	4/61	6
Often	11/61	18
Sometimes	30/61	49
Rarely	12/61	20
Never	4/61	7

TABLE 17.

**Question 32. Please select your stance on the following statements.**

**Please choose one of the following answers for each statement.**

- a. *I think that archaeologists across all sub-disciplines share common core concepts and terminology*

Response	#	%
Strongly Agree	7/64	11
Agree	19/64	30
Moderately Agree	30/64	47
Moderately Disagree	3/64	5
Disagree	3/64	5
Strongly Disagree	2/64	3

- b. *When I read archaeological literature across all sub-disciplines, I see a commonly shared bank of concepts and terminology*

Response	#	%
Strongly Agree	5/62	8
Agree	20/62	32
Moderately Agree	29/62	47
Moderately Disagree	5/62	8
Disagree	3/62	5
Strongly Disagree	0/62	0

- c. *I think that archaeologists in my sub- discipline(s) share common core concepts and terminology*

Response	#	%
Strongly Agree	7/63	
Agree	26/63	
Moderately Agree	25/63	
Moderately Disagree	2/63	
Disagree	2/63	
Strongly Disagree	1/63	

- d. *When I read archaeological literature in my sub-discipline(s), I see a commonly shared bank of concepts and terminology*

Response	#	%
Strongly Agree	7/62	11
Agree	27/62	44
Moderately Agree	25/62	40
Moderately Disagree	3/62	
Disagree	0/62	0
Strongly Disagree	0/62	0

TABLE 18.

**Question 34. When does the standardisation of terminology become most important?**

Response	#	%
At Data Creation	38/62	61
At Data Analysis	5/62	8
During Writing	12/62	19
During Editing	1/62	2
At Publication	3/62	5
Other	3/62	5

## 2.2. QUALITATIVE RESPONSES

### QUESTION 25

**If you answered yes to the question above, how serious would you consider the impact of the lack of standardisation on archaeology as a discipline?**  
**If you did not answer yes to the question above, please select 'no answer.' Optional: can you think of a specific example where a lack of standardization has been a significant issue? Please enter in comment section.**

Participant ID	Response
3	"Researcher (Marwick 2002) not defining artefact type that comprises 80% of his Pleistocene assemblage."
10	"Definitions surrounding archaeological "potential" vary from company to company, and between individuals."
12	"Biggest issue is when the writers fails to define their own terminology"
22	"It's difficult to quantify, though, because the question assumes that a difference in terminology equates to a difference in interpretation - whereas the interpretation can be poor or non existent even with the correct and accepted terminology. I've said major, but there is no way to know this for sure without an innovative way to measure the correlation."
24	"Moving from one country to another, where the practitioners used a completely different technical vocab to that used in Australian practice"
29	"It is a major impact because the communication with other sub-disciplines is affected and even misunderstandings in the sub-disciplines are common"
30	"I am currently revising my Glossary of rural fence terms to provide a consistent terminology for confused and confusing terms."



35	"Not a major issue, more of a sometimes issue"
38	"Lack of standardisation of naming stone types, even within the same report."
40	"In a development context the lack of standardisation affects what people identify as worth saving."
49	"Use of 'place' or 'site', stone material definitions; functional versus morphological stone tool categorisation; site type categorization"
58	Confusion over the meaning of certain terms prevents archaeologists from fully understanding the nature of their assumptions. For example, most archaeologists have a flawed definition of the term "deterministic".
67	"Terminology used in heritage management differs substantially from archaeological terminology. The absence of standard terminology for use in the museum sector is an issue, especially when curators and managers are from differing disciplines (e.g. archaeology, anthropology, geography and [increasingly] cultural heritage management."
68	"The lack of agreement stymies the development of core vocabulary for projects like FAIMS!"
76	"Archaeological Potential vs. Archaeological Significance"
78	"Standardized terminology is not needed (except for archiving data bases for universal storage/access) - on the proviso that terms are defined and the rationale for their use is explained"
79	"Leads to conflicting results, misinterpretation and reporting that is unable to be duplicated/reverse engineered"
82	"Databases for recording archaeological excavation results, including analysis of artefacts."
85	"Depends on clarity of definitions provided"
86	"Heritage consultants in Western Australia produce large data sets each year. While many concepts are shared, terminologies can be troublesome. Harnessing the value of these data sets in research is made problematic, but not impossible, because of a lack of standardisation."
93	"Geological terms for lithic raw materials are a problem -

	particularly between states.”
105	“Yes in regards to making decisions about site extents in situations with extensive background scatters.”
108	“Lack of standardisation produces disparate recording systems which are hard to understand/reconcile by people evaluating, comparing or analysing data collected in the field or by research.”
110	“Failure to record some information which is recorded by others - a standard recording format would trigger responses”
112	“Standards are useful, but need to be contestable and extensible. Should be formalized more (in a computing sense) so they can actually be used, adapted, contested, and revised in manner that is more clear and explicit. Now all that only happens vaguely (if at all) in prose in some closed-access non reusable journal.”
115	“Different definitions of "scatter" of lithics in Australian archaeology”
116	“Lack of standarisation is an issue in linking data from different datasets with different platforms so that they can be discovered or combined for new analysis”
117	“There is a tacit standard, and then there are books with standards.”

## QUESTION 26

**Are you willing to compromise over the use of terms and concepts in your own practice to achieve a reconciliation of terms?**

Participant ID	Response
3	"I have worked collaboratively with other researchers and have shifted definitions and terms if it makes sense and is for the betterment for the project."
10	"Providing I am sure that other companies are using the same terms and concepts, it would not be an issue."
12	"Only if I believe the standard terms are better and more easily understood"
15	"Depends on what I am being asked to compromise, and how much I disagree with the use of the terms and concepts"
29	"We should accept complementary definitions of terms and concepts and find reconciliation of terms on a higher level of abstraction"
30	"Confusion over nomenclature inhibits understanding, and may lead to incorrect conclusions."
32	"In most states (but particularly NSW where I work), terminology and concepts are clearly prescribed by the state regulator (OEH) in their guideline documents, so practicing archaeologists typically comply with these terms and concepts to ensure projects are compliant with legislative requirements."
49	"Happy to use an accepted and clear term where it can be shown to be more widely adopted - however, consensus is often lacking"
53	"However a lot of the terms I use that both archaeologist question are well accepted in chemistry, geology, geomorphology etc and in those cases I will not compromise"
70	"My employers/clients are not professionals and to assist them in understanding the significance of a site or object may require using common language rather than technical terms"
76	"I work in government - the terms I use are fixed and defined by"

	Legislation and can not be changed (easily)."
79	"As long as the terms achieved following compromise are accurate"
82	Yes, for sure."
86	"So long as I share the same goal I don't mind compromising. E.g. If the goal is to characterise a stone artefact assemblage I am happy to record orientated and/ or maximum dimensions."
93	"The terminology used is far less important than the underlying concepts. Unfortunately, the limited experience of many archaeologists means that often they don't really know what they are talking about, or don't realise the difference between the local and the general. Unfortunately the tendency for the loudest voices to prevail is unhelpful."
104	"It would depend on the different situations."
107	"If I feel it is appropriate"
112	"Yep, but see above"
116	"In theory yes, but in practice I suspect that there would be areas in which I do not want to use a standardized vocabulary that departs from usage in my subfield or geographic area, or that does not accurately describe my material (especially in terms of chronological periods or ceramic types)"

## QUESTION 29

**Please describe in a few words an example where a disagreement over the definition of a term had a significant impact on research, analysis, or understanding. Please name the term, and provide a brief description.**

Participant ID	Response
2	"Can't remember the exact phrasing but 'same as', i.e. 'same style as' had been misinterpreted as 'same vessel as', i.e. conjoined artefacts. This found its way into the report and it was claimed that the conjoins were distributed over several houseblocks and that all deposits were part of a 'single event'. Further checking revealed there were none or few actual joins and there was no reason to presume these separate fills were part of the same deposition event."
6	"Artefact scatter: Following comments on the issue at a AAA conferences and elsewhere by Colin Pardoe, scatter is widely used to denote something that is in fact the opposite - it is usually better described as an 'artefact concentration'"
10	"Other consultants labelled areas as being "Potential Archaeological Deposits" (PADs) interspersed with "Potentially Sensitive Areas" (PSAs). PADs have a defined meaning with the OEH and archaeologists in NSW, while no one I was in contact with had come across PSAs as a definition of area. It also caused complications when trying to explain to OEH why these areas of PSA had been created by other consultants."
14	"Disagreement about the use of specific terms has led to a major long-running debate in publications between myself (and a co-author) and another colleague working in the same field."
19	"Error used to express only precision, when accuracy is the more significant, larger term (and remains unexpressed)"
24	"When I started work in the UK I almost got fired because they thought I didn't know what I was talking about due to

	terminology differences”
30	“Problems with referees over my use of terms in paper submitted to journal”
38	“Opalite is regarded by geologists to be an ill-defined misnomer not to be used, yet it is still used in an archaeological context in several reports that I have recently dealt with.”
40	“Different definitions of a site will produce reports with different numbers of sites. This has management implications.”
43	“The use of the term burnishing or burnished is far too rigid- the disagreement was over the prescription of how early ceramicists achieved the smooth 'burnished' effect. As I am also a ceramicist, I understand that a potter can use virtually any smooth object (spoon, river stone, thumb/finger nail, glass etc...), and lubricating medium (oils, fats, water, slurry) to compact the surface of the clay to achieve a smooth and shined finish.”
45	“Artefact; how broad is the category of artefact? Do you have to be able to pick it up”
49	“Site and place have been open to interpretation, along with artefact scatter, isolated artefact, etc. many arguments with regulators, project sponsors, and aboriginal groups, about how to define site extents.”
53	“In the case of clasifing raw materials miss-attribution of a raw material (such as hornfelds for quartzite) can lead to a miss-attribution of likely source locations and hence influence the interpretation of a lithic dataset”
58	“In both the UK and Australia the use of the term "interpretive" implies a post-processualist agenda; where there is often a rejection of quantitative analysis and the term typically refers to the ephemeral (or non-quantifiable) part of the argument being made. In contrast, to me, interpretive implies that the author is making a detailed argument of causality and illustrates the concept without any reference to the post-processual/processual debate.”
60	“Most archaeologists don't know anything about lithology. Therefore their recording is meaningless.”

63	"Not so much how the terms are defined, but how they are applied in the field."
67	"When inadequate or incorrect terminology is used in a descriptive database, it becomes very difficult to conduct useful searches. There is no thesaurus common to all heritage institutions in Australia that connects and captures all terms."
70	"In designing a Standard Operating Procedure for identifying human bones / indigenous graves etc., my colleague chose the definition of "skeletal remains" believing that "human remains" indicated that there was a "body" involved. The conflict of terms certainly arose with the definitions of the Act (Criminal Code Act 1899 and Coroners Act 2003 (Qld))"
74	"The definition of miniature is often assumed not argued for resulting in confused use of the term in publications and in field work reports so that vessels are haphazardly defined as miniature or not."
76	"It didn't overly because I made them change their report to the NSW Heritage Council preferred term. They weren't happy, but I was."
79	"Stringer - it is a term used to describe a timber used in ship construction to provide longitudinal stability. Another archaeologist was using it to describe timber jetty support which is incorrect"
82	"Clients and some colleagues become very confused between an assessment of a site having 'archaeological potential' verses a site that has 'archaeological significance'. They think that both mean the same thing and using clearly defined government definitions (publications) doesn't always help. Can lead to re-writing sections of reports to be VERY clear."
87	"The scholar I was trying to communicate with kept saying "you mean a 'clean context'?" So I asked him to explain what he meant. He said he doesn't know if he has a "clean context" until he is months out of the field when he has a ceramicist come and look at his pottery. If she dates it all the same he designates the locus a "clean context." We define the locus by the stratification and read

	the pottery on site. It's defined by the stratification and the date results from ceramic analysis, not vice versa."
103	"Angular fragments in the Pilbara are in 90% of cases natural. I have had arguments with other consultants who use that term and their definition means that any piece of heat fractured stone with no flaking diagnostics is recorded as an artefact. This skews the results of their analysis"
105	"Where people had limited understanding of what a midden was and were not capable of distinguishing between anthropogenic and natural deposits containing shell. This is much about poor understanding of geomorphology as it is about standards though."
115	"Artefact scatter - people use different numbers and/or densities of artefacts leading directly to different interpretations of the field notes and/or report"
117	"Archaeologists sometimes think they are special and have a special vocabulary, but truly they are products of their institutions and training. Only time I have a problem is when they try to reinvent the wheel."



## QUESTION 34

**When does the standardisation of terminology become most important?**

Participant ID	Response
1	"All of the above, as they're distinct steps."
22	"All times - creation is critical for the recognition but since everything after that depends on it they can't be separated."
32	When seeking approvals from the regulator - non compliance with their prescribed terminology can lead to rejection of permit applications & can compromise projects"

## QUESTION 35

**Thank you for participating in our survey. If you have any further comments please enter them in comment box below. If you wish for a reply, further information or to be updated on the progress of this project, please leave an email address.**

Participant ID	Response
6	"While it is not ideal, and practitioners should in fact have the freedom to undertake archaeological investigation according to methods that they can reasonable justify, consulting practice is increasingly proscribed by Codes and Guidelines. My feeling in the structure of the questions is that this may have been underestimated."
24	"A major problem in Australian archaeology is the difference between training for indigenous and historical archaeology. Training in the lexicon of indigenous test pit excavation doesn't prepare you for open area historical site excavation and vice versa."
30	"I'm not sure if my response will help you. I'm not a professional archaeologist, but over the last decade or so, I've done a lot of historical archaeological research on rural fences. A major part of this has been to collect the vocabulary and record it so that we can get some standardisation, or at least, better understanding of the range of terms."
43	"I would like for a dictionary/lexicon of concepts and terms to be made available. Preferably a hardback book and an online interface/ advanced search engine function. If it is online it can be updated and amended as needed. Having it in book form is also very important."
70	"My job is mainly in the resource industry ensuring cultural heritage items and sites of significance are protected from damage during mining activities. I tend not to have to analyse

	<p>the object or site, but rather identify it as having significance. My reports to the Traditional Owners and management have to be modified to the respective audience. The emphasis is not what the object/site is but how we can preserve and protect it.”</p>
79	<p>“In Australia, a standardised reporting and recording system and vocabulary is urgently needed. Likewise an online repository for archaeological data is required. More importantly a mandate that requires all the documentary and cultural archive to be deposited and publically available is also require. I applaud the FAIMS project, however feel that this may not be fully utilised/the benefits fully realised unless the Government heritage agencies (C'wealth included) are on board with the project and specify use of the FAIMS repository as part of the consent conditions therefore obliging archaeologists to use FAIMS as occurs with EH/ADS in the UK.”</p>
110	<p>“I myself use, and I see other people use, teaching and reference materials created a long time ago - I don't necessarily consider these to be outdated. Some have never been updated, some remain standards, some are used to teach to start off the study of a subdiscipline. Adopting new terminology is not as important as being clear on synonyms between terminologies, therefore I might not change my terms as long as also include terms used by others, if I don't consider them better.”</p>
112	<p>“I'd be careful on how standards get implemented. Standards can be hugely helpful, but they need to be contestable. Publishing, revising, contesting, etc. explicit machine-readable standards needs to be a much more regular part of research. It requires new modes of scholarly communication, since books and journals can't do it. Standards are not something that should ever be frozen for all time, since that would mean we've run out of interesting research problems”</p>
116	<p>“I think it's worth reiterating -- because I suspect I'm not alone in this -- that although I favor standardization in theory and am</p>

strongly convinced of the importance of consistency in terminology for data sharing, I think that in practice it's unrealistic to expect archaeologists across the discipline to adopt standardized terminologies imposed from elsewhere unless there is substantial and immediate benefit (just look at the different terms different projects insist on applying for the idea of archaeological context). I think it's much better to foster a more distributed ecosystem but provide mappings to common concepts and/or common points of reference, so that datasets can be loosely coupled rather than directly integrated. The problem, to my mind, is that when we try to standardize, our field just ends up with a lot of competing and incompatible standards, to which different camps adhere."

## APPENDIX THREE – PEARSON’S CHI-SQUARE TEST OF INDEPENDENCE

### HYPOTHESIS 1

Question 1 and 27

**Respondents working in CHM and government institutions are more likely to  
consider that most, or all, of their own lexicon is part of a wider, accepted  
vocabulary within their sub-discipline(s)**

	All + Most	Few + Some	TOTAL
CHM/Gov. Inst.	28	1	29
Other Industry	29	3	32
TOTAL	57	4	61

- Chi-squared = 0.872
- Two-tailed P value = 0.3504
- The association between rows (groups) and columns (outcomes) is considered to be not statistically significant

## HYPOTHESIS 2

Question 1 and 27

**Respondents working in a university setting are more likely to consider that few, or some, of the their own lexicon is part of a wider, accepted vocabulary within their sub-discipline(s)**

	Few + Some	All + Most	TOTAL
University	2	19	21
Other Industry	2	38	40
TOTAL	4	57	61

- Chi-squared = 0.460
- Two-tailed P value = 0.4977
- The association between rows (groups) and columns (outcomes) is considered to be not statistically significant
-

### HYPOTHESIS 3

Question 1 and 27

**Respondents currently studying are more likely to consider that all of their own lexicon is part of a wider, accepted vocabulary within their sub-discipline(s)**

	All	Most + Few + Some	TOTAL
Currently Studying	1	5	6
Other Industry	17	38	55
TOTAL	18	43	61

- Chi-squared = 0.528
- Two-tailed P value = 0.4676
- The association between rows (groups) and columns (outcomes) is considered to be not statistically significant

## HYPOTHESIS 4

Question 24 and 16

**Respondents who often/always provide meaningful definitions for key terms in their publications and/or datasets are more likely to consider that there is not lack of standardized vocabulary in archaeological literature/reports**

	No	Yes + Not Sure	TOTAL
<b>Always + Often</b>	6	36	42
<b>Sometime + Rarely</b>	5	16	21
<b>+ Never</b>			
<b>TOTAL</b>	11	52	63

- Chi-squared = 0.881
- Two-tailed P value = 0.3479
- The association between rows (groups) and columns (outcomes) is considered to be not statistically significant



## HYPOTHESIS 5

Question 24 and 26

**Respondents who are willing to compromise over their use of terms and concepts  
to achieve a reconciliation of terms believe there is a lack of standardised  
vocabulary in archaeological literature/reports**

	<b>Lack of Standards</b>	<b>No Lack of Standards</b>	<b>TOTAL</b>
<b>Willing to compromise</b>	30	14	44
<b>Not willing to compromise + Not sure</b>	14	5	19
<b>TOTAL</b>	44	19	63

- Chi-squared = 0.191
- Two-tailed P value = 0.6623
- The association between rows (groups) and columns (outcomes) is considered to be not statistically significant

## HYPOTHESIS 6

Question 24 and 27

**Respondents who consider there to be a lack of standardised vocabulary in archaeological literature/reports think that not all their terms are part of a wider accepted vocabulary**

	Few + Some	All + Most	TOTAL
<b>Lack of Standards</b>	3	37	40
<b>No Lack of Standards</b>	1	18	19
<b>TOTAL</b>	4	55	59

- Chi-squared = 0.102
- Two-tailed P value = 0.7495
- The association between rows (groups) and columns (outcomes) is considered to be not statistically significant

## HYPOTHESIS 7

Question 24 and 30c

**Respondents who consider there to be a lack of standardised vocabulary in archaeological literature/reports are always/often in disagreement with regards to the terminology used within fieldwork, research or writing *within* their own-sub-disciplines terminology**

	Sometimes +		
	Always + Often	Rarely + Never + No Answer	TOTAL
<b>Lack of Standards</b>	3	37	40
<b>No Lack of Standards</b>	3	16	19
<b>TOTAL</b>	6	53	59

- Chi-squared = 0.969
- Two-tailed P value = 0.3249
- The association between rows (groups) and columns (outcomes) is considered to be not statistically significant

## HYPOTHESIS 8

Question 24 and 30d

**Respondents who consider there to be a lack of standardised vocabulary in archaeological literature/reports are always/often in disagreement with regard to the terminology used within fieldwork, research or writing *outside* of their own sub-discipline.**

	Sometimes +		
	Always + Often	Rarely + Never + No Answer	TOTAL
<b>Lack of Standards</b>	2	34	36
<b>No Lack of Standards</b>	4	15	19
<b>TOTAL</b>	6	49	55

- Chi-squared = 3.073
- Two-tailed P value = 0.0796
- The association between rows (groups) and columns (outcomes) is considered to be not statistically significant

## HYPOTHESIS 9

Question 24 and 32a

**Respondents who consider there to be a lack of standardised vocabulary in archaeological literature/reports agree that archaeologists across all sub-disciplines share common core concepts and terminology.**

	Agree	Disagree	TOTAL
<b>Lack of Standards</b>	36	7	43
<b>No Lack of Standards</b>	18	1	19
<b>TOTAL</b>	54	8	62

- Chi-squared = 1.423
- Two-tailed P value = 0.2329
- The association between rows (groups) and columns (outcomes) is considered to be not statistically significant

## HYPOTHESIS 10

Question 24 and 32b

**Respondents who consider there to be a lack of standardised vocabulary in archaeological literature/reports agree that when they read archaeological literature across all sub-disciplines they see a commonly shared bank of core concepts and terminology**

	Agree	Disagree	TOTAL
<b>Lack of Standards</b>	34	8	42
<b>No Lack of Standards</b>	19	0	19
<b>TOTAL</b>	53	8	61

- Chi-squared = 4.165
- Two-tailed P value = 0.0413
- The association between rows (groups) and columns (outcomes) is considered to be not statistically significant

## HYPOTHESIS 11

Question 24 and 32c

**Respondents who consider there to be a lack of standardised vocabulary in archaeological literature/reports do not consider their own-sub-disciplines to lack common core concepts and terminology**

	Agree	Disagree	TOTAL
<b>Lack of Standards</b>	39	4	43
<b>No Lack of Standards</b>	19	0	19
<b>TOTAL</b>	58	4	62

- Chi-squared = 1.889
- Two-tailed P value = 0.1693
- The association between rows (groups) and columns (outcomes) is considered to be not statistically significant

## HYPOTHESIS 12

Question 24 and 32d

**Respondents who consider there to be a lack of standardised vocabulary in archaeological literature/reports agree that when they read archaeological literature in their own sub-disciplines they see a commonly shared bank of concepts and terminology**

	Agree	Disagree	TOTAL
<b>Lack of Standards</b>	39	3	42
<b>No Lack of Standards</b>	19	0	19
<b>TOTAL</b>	58	3	61

- Chi-squared = 1.427
- Two-tailed P value = 0.2322
- The association between rows (groups) and columns (outcomes) is considered to be not statistically significant



## HYPOTHESIS 13

Question 16 and 17

**Respondents who claim they often/always provide meaningful definitions for key terms in their publications, say that other archaeologists sometimes/rarely provide meaningful definitions for key terms in their publications**

	Sometimes + Rarely + Never	Often + Always	TOTAL
<b>Often + Always</b>	35	19	54
<b>Sometime + Rarely + Never</b>	18	0	18
<b>TOTAL</b>	53	19	72

- Chi-squared = 8.604
- Two-tailed P value = 0.0034
- **The association between rows (groups) and columns (outcomes) is considered to be very statistically significant**

## APPENDIX FOUR – “ARTEFACT SCATTER” COMPONENT FORM

**Grid Coordinates\*** E      N      *\*All grid coordinates must be presented with reference to the MGA94 datum* ☐ Zone 54  
☐ Zone 55

**Context** ☐ Open ☐ Subsurface ☐ Overhang ☐ Cave

**Ground Surface Exposure** \_\_\_\_\_ % (*Percentage of ground surface visible*)

Area of Scatter Examined							m <sup>2</sup>
	All	Centre	NE	NW	SW	SE	

Average density \_\_\_\_\_m<sup>2</sup>      Maximum density \_\_\_\_\_m<sup>2</sup>      Minimum density \_\_\_\_\_m<sup>2</sup>

Number of Artefacts

Estimated number ☐ 10s ☐ 100s ☐ 1000s ☐ 10000s

Counted number \_\_\_\_\_

## COMPOSITION OF SCATTER

Predicted number \_\_\_\_\_  
across site

The following tables should be used to indicate the general composition of artefact scatters. Include a key for any abbreviations used. "Industry" refers to the method of production used in manufacture (i.e. battered, ground, pecked, flaked, etc). "Technological Class" refers to reduction type (e.g. flake core, core, retouched flake etc). "Artefact Type" refers to the typology (e.g. thumbnail scraper, backed blade etc).

**Other Cultural Material** *(Include non-lithic artefacts and faunal remains other than shell)*

[illegible]

**RETURN TO:** The Heritage Registrar, Aboriginal Affairs Victoria, PO Box 2392 MELBOURNE VIC 3001

**Lithic Artefacts** (Include flaked glass and ceramic, and grindstones).

[illegible]

(Show relationship between components and label accordingly.  
Use same horizontal scale for plan and profile. Show scale values.)

North

This image shows a full page of blank graph paper. The grid consists of thin, light gray horizontal and vertical lines that intersect to form a series of small, uniform squares across the entire surface. There are no margins, text, or other markings on the paper.

(Do not excavate. Use same horizontal scale for plan and profile. Show scale values.)

[illegible]

Depth/Height \_\_\_\_\_ m

---

---

---

---

---



## REFERENCES

- Alexa, Melina. *Computer Assisted Text Analysis Methodology in the Social Sciences*. Mannheim: ZUMA Arbeitsbericht, 1997.
- ASIS&T. "Dublin Core Metadata Initiative: Making It Easier to Find Information." <http://dublincore.org>. Accessed: 5 October, 2013.
- Babeu, Alison. *Rome Wasn't Digitised in a Day: Building a Cyberinfrastructure for Digital Classicists*. Washington: Council on Library and Information Resources, 2011. <http://www.clir.org>.
- Baines, Andrew, and Kenneth Brophy. "What's Another Word for Thesaurus? Data Standards and Classifying the Past." In *Digital Archaeology: Bridging Method and Theory*, edited by Thomas L. Evans and Patrick Daly, 236-50. London: Routledge, 2006.
- Ballsun-Stanton, Brian. "Asking About Data: Exploring Different Realities of Data Via the Social Data Flow Network Methodology." The University of New South Wales, 2012.
- Binding, Ceri, Keith May, and Douglas Tudhope. "Semantic Interoperability in Archaeological Datasets: Data Mapping and Extraction Via the Cidco Crm." In *Research and Advanced Technology for Digital Libraries*, 280-90. Berlin Heidelberg: Springer, 2008.
- Burke, Heather, Claire Smith, and Larry Zimmerman. *The Archaeologist's Field Handbook*. Crows Nest: Rowman Altamira, 2009.
- Byrant, Scott. "Overcoming the 'Not Invented Here' Syndrome: Experience with Sourcing Educational Multimedia Developed Elsewhere." Paper presented at the Flexibility: The Next Wave, Proceedings of ASCILITE'98, 1998.

- Byrne, Denis. "Standards for Archaeological Practice in Aboriginal Heritage Management." edited by NSW National Parks and Wildlife Service. Hurstville: NSW National Parks and Wildlife Service, 1997.
- Chandrasekaran, Balakrishnan, John R. Josephson, and V. Richard Benjamins. "What Are Ontologies and Why Do We Need Them?". *Intelligent Systems and Their Application* 14, no. 1 (1999): 20-26.
- Chaoji, Vineet, Apirak Hoonlor, and Boleslaw K. Szymanski. "Recursive Data Mining for Role Identification." In *Proc. 5th International Conference on Soft Computing as Transdisciplinary Science and Technology*, 218-25. Paris, France, October 27-31, 2008.
- Consultants, Andrew Long & Associates Pty Ltd: Archaeological and Heritage. "Appendix 6: Areas of Cut and Fill." In *Frankston Bypass EES Project*, 137-58. Southern and Eastern Integrated S31 Transport Authority, 2008.
- Doerr, Martin. "The Cidoc Conceptual Reference Model: An Ontological Approach to Semantic Interoperability of Metadata." *AI Magazine* 24.3, 2003.
- Geospatial, Open. "Kml." <<http://www.opengeospatial.org/standards/kml%3E>. Accessed: 29 October, 2013.
- Gill, Tony, and Paul Miller. "Re-Inventing the Wheel? Standards, Interoperability and Digital Cultural Content." *D-Lib Magazine*, January 2002.
- Harley, Diane, Jonathan Henke, Shannon Lawrence, Ian Miller, Irene Perciali, Daiv Nasatir, Charis Kaskiris, and Cara Cautista. "Use and Users of Digital Resources: A Focus on Undergraduate Education in the Humanities and Social Sciences." Center for Studies in Higher Education, UC Berkeley, 2006.
- Heflin, Jeff, and James Hendler. *Semantic Interoperability on the Web*. Maryland University College Park Department of Computer Science, 2000.

- Heiler, Sandra. "Semantic Interoperability." *ACM Computing Surveys* 27, no. 2 (1995): 271-73.
- Help, ArcGIS Server 9.3. "The Ogc Well-Known Text Representation of Spatial Reference Systems." [http://webhelp.esri.com/arcgisserver/9.3/java/index.htm - geodatabases/the\\_ogc-607957855.htm](http://webhelp.esri.com/arcgisserver/9.3/java/index.htm-geodatabases/the_ogc-607957855.htm). Accessed: 29 October, 2013.
- Holemen, Jon, Christian-Emil Ore, and Øyvind Eide. "Documenting Two Histories at Once: Digging into Archaeology." Paper presented at the CAA 2004, Beyond the Artefacts: Digital Interpretation of the Past, 13-17 April 2004.
- Humphreys, Betsy L., Donald A. B. Lindberg, Harold M. Schoolman, and G. Octo Barnett. "The Unified Medical Language System: An Informatics Research Collaboration." *Journal of the American Medical Informatics Association* 5, no. 1 (Jan/Feb 1998): 1-11.
- Hunter, Jane. "Enhancing the Semantic Interoperability of Multimedia through a Core Ontology." *Circuits and Systems for Video Technology, IEEE Transactions* 13, no. 1 (2003): 49-58.
- Initiative, Dublic Core Metadata. "Dublic Core Metadata Element Set, Version 1.1." <http://dublincore.org/documents/dces/>. Accessed: 5 October, 2013.
- Kansa, Eric. "Openness and Archaeology's Information Ecosystem." *World Archaeology* 33, no. 4 (2012): 498-520.
- Kansa, Eric, and Ahrash Bissell. "Web Syndication Approaches for Sharing Primary Data in "Small Science" Domains." *Data Science Journal* 9 (2010): 42-53.
- Kilbride, W. "The Danube in Prehistory in the Digital Age: Towards a Common Information Environment for European Archaeology." *Archeologia e calcolatori* 15 (2004): 129-44.



- Kintigh, Keith W. "The Challenge of Archaeological Data Integration." In *Technology and Methodology for Archaeological Practice: Practical Applications for the Past Reconstruction*, 81-86: BAR International Series S 2029, 2009.
- Kintigh, Keith. W., Francis P. McManamon, and Katherine Spielmann. "Enhancing Data Comparability and Enabling Synthesis with Tdar (the Digital Archaeological Record)." Paper presented at the 78th Annual Meeting of the Society for American Archaeology, Honolulu, Hawaii, April 3-7 2013.
- Kubicek, Herbert, Ralf Cimander, and Hans Jochen Scholl. "Layers of Interoperability." In *Organisational Interoperability in E-Government: Lessons from 77 European Good-Practice Cases*. Berlin: Springer, 2011.
- Lewis, Eleanor T., Jana Diesner, and Kathleen M. Carley. "Using Automated Text Analysis to Study Self-Preservation Strategies." In *Computational Analysis of Social and Organizational Systems (CASOS) Conference*. Pittsburgh PA, 2001.
- Limp, Frederick. "Web 2.0 and Beyond, or on the Web Nobody Knows You're an Archaeologists." In *Archaeology 2.0: New Tools for Communication and Collaboration*, edited by Eric Kansa. Los Angeles: Cotsen Institute of Archaeology, 2011.
- Lowry, Richard. "2x2 Contingency Table." <http://www.vassarstats.net/tab2x2.html>. Accessed: 15 September, 2013.
- Mattison, David. "The Digital Humanities Revolution." *Searcher* 14, no. 5 (2006): 25-34.
- McDonald, J.H. *Handbook of Biological Statistics: Basic Concepts of Hypothesis Testing (2nd Ed.)*. Baltimore, Maryland: Sparky House Publishing, 2009.
- McGuinness, Deborah, and Frank van Harmelen. "Owl Web Ontology Language." *W3C Recommendation* 10, no. 2004-03 (2004).

- Meckseper, Christiane, and Claire Warwick. "The Publication of Archaeological Excavation Reports Using Xml." *Literary and Linguistic Computing* 18, no. 1 (2003): 63-75.
- Natalya, Noy F., and Deborah L. McGuinness. "Ontology Development 101: A Guide to Creating Your First Ontology." Stanford: Stanford University, 2001.
- Nelson, Bryn. "Empty Archives." *Nature* 461, no. 10 (2009).
- "P Values." [http://www.statsdirect.com/webhelp/content/basics/p\\_values.htm](http://www.statsdirect.com/webhelp/content/basics/p_values.htm). Accessed: 10 October, 2013.
- Press, N.I.S.O. "Understanding Metadata." *National Information Standards* 20 (2004).
- Richards, Julian D. "Digital Preservation and Access." *European Journal of Archaeology* 5, no. 343 (2002): 343-66.
- . "From Anarchy to Good Practice: The Evolution of Standards in Archaeological Computing." *Archeologia e Calcolatori* 20 (2009): 27-35.
- Richards, Julian D., and Catherine Hardman. "Stepping Back from the Trench Edge: An Archaeological Perspective on the Development of Standards for Recording and Publication." In *The Virtual Representation of the Past*, edited by M. Greengass and L. Hughes, 101-12. Farnham, Surrey and Burlington, USA: Ashgate Publishing Company, 2008.
- Ross, Shawn, Adela Sobotkova, Brian Ballsun-Stanton, and Penny Crook. "Creating Eresearch Tools for Archaeologists: The Federated Archaeological Information Management Systems Project." *Australian Archaeology* 77 (2013): 107-19.
- Snow, Dean R., Mark Gahegan, and Lee C. Giles. "Cybertools and Archaeology." *Science* 311, no. February (2006): 958-59.

"Statistical Significance." StatPac, <http://www.statpac.com/surveys/statistical-significance.htm>. Accessed: 15 October, 2013.

Stockburger, David W. "Chi-Square and Tests of Contingency Tables." Missouri State University, <http://www.psychstat.missouristate.edu/introbook/sbk28m.htm>. Accessed: 15 September, 2013.

Sturges, Paul, and Anne Griffin. "The Archaeologist Undeceived: Selecting Quality Archaeological Information from the Internet." *Informing Science Journal* 6 (2003): 221-32.

"UK HE Research Data Management Survey." Loughborough University Research Data Management Project, <https://docs.google.com/forms/d/1M2AY1y9pCv2OAibqtntyWJZnM6880fBBxLv5xcehbZlc/viewform>. Accessed: 13 October, 2013.

Ulm, Sean, Stephen Nichols, and Cameo Dalley. "Mapping the Shape of Contemporary Australian Archaeology: Implications for Archaeology Teaching and Learning." *Australian Archaeology* 61 (2005): 11-23.

Veer, Hans van der, and Anthony Wiles. "Achieving Technical Interoperability." Sophia Antipolis Cedex, France: European Telecommunications Standards Institute, April 2008.

Weber, R. P. "Computer-Aided Text Classification : Advantages , Difficulties , and Types." *Qualitative Sociology* 126-147 (1984).