Paradata in Documentation Standards and Recommendations for Digital Archaeological Visualisations

Lisa Börjesson, Olle Sköld, Isto Huvila

Abstract

Digitalisation of research data and massive efforts to make it findable, accessible, interoperable, and reusable has revealed that in addition to an eventual lack of description of the data itself (metadata), data reuse is often obstructed by the lack of information about the data-making and interpretation (i.e. paradata). In search of the extent and composition of categories for describing processes, this article reviews a selection of standards and recommendations frequently referred to as useful for documenting archaeological visualisations. It provides insight into 1) how current standards can be employed to document provenance and processing history (i.e. paradata), and 2) what aspects of the processing history can be made transparent using current standards and which aspects are pushed back or hidden. The findings show that processes are often either completely absent or only partially addressed in the standards. However, instead of criticising standards for bias and omissions as if a perfect description of everything would be attainable, the findings point to the need for a comprehensive consideration of the space a standard is operating in (e.g. national heritage administration or international harmonisation of data). When a standard is used in a specific space it makes particular processes, methods, or tools transparent. Given these premises, if the standard helps to document what needs to be documented (e.g. paradata), and if it provides a type of transparency required in a certain space, it is reasonable to deem the standard good enough for that purpose.

Keywords: Archaeology, FAIR, Heritage Data, Metadata, Paradata, Provenance, Standards, Transparency, Visualisations

DOI 10.14361/dcs-2020-0210
Introduction

Archaeology and cultural heritage preservation are heterogeneous disciplines. Global, local, formal and informal metadata standards and efforts to establish new standards hold them together and regulate activities from fieldwork documentation (e.g. Archaeology Data Service and Digital Antiquity, 2009-), to architectural reconstructions and visualisations (e.g. López et al., 2018), and virtualisation of heritage environments and intangible history (Papadopoulos and Schreibman, 2019). While many standardisation initiatives have focused on description of archaeological information objects like geographic datasets or visualisations, both archaeological (Bentkowska-Kafel and Denard, 2012; Richards-Rissetto and von Schwerin, 2017) and transdisciplinary (Pasquetto et al., 2019) research emphasise the parallel importance of documenting how these information objects were created. As a complement and in parallel to ‘metadata’, archaeologists dealing with visual data are increasingly using the term paradata to denote the provenance and processing history of visual objects. The term paradata is commonly used of data that aims to provide transparency of the operational and intellectual processes preceding image creation (Sköld et al., work in progress; Bentkowska-Kafel and Denard, 2012). “Naive” use of the term paradata has been criticised (Havemann, 2012), i.e. an assumption that paradata can give an exhaustive account of provenance and processing imparting a digital representation as neutral “redo property”. In parallel, the absence and need of standards and comprehensive guidelines for documenting paradata has been recognised as a problem (Nicolucci, 2012; Borrero and Stroth, 2020). Even if no data description can recount all details, well-structured process descriptions are vital to maintaining insight into the production of visualisations. However, considering that several existing heritage visualisation-related standards and recommendations stipulate documentation of paradata-like information to a varying degree and extent, a question arises of what they do as they are clearly not deemed good or comprehensive enough.

To increase the understanding of how contemporary documentation schemes represent paradata, this article reviews a selection of major standards and recommendations frequently referred to as useful for documenting archaeological visualisations in search of how they document processes i.e. make explicit and implicit stipulations relating to paradata. The article inquires into what process descriptions the standards support, how processes are represented, and what space i.e. context the standards are encompassing. Building on the long tradition of critical and practical knowledge organisation research in information science, a critical close reading (DuBois, 2003) of a selection of documentation schemes in this article has two aims. First, it aims to provide insight into (1) how current standards can be employed to document provenance and processing history (i.e. paradata) to inform (as urged by Hansson et al., 2020) the development of existing and new metadata and paradata standards, and second, to attain a (2) critical under-
standing of what aspects of the processing history can be made transparent using the reviewed standards, and which aspects are pushed back or hidden.

**Politics of metadata and paradata**

A long line of earlier research has investigated the political nature of organising and describing knowledge (Leazer and Montoya, 2020). This subsumes also the creation of metadata and paradata, and the development and use of related standards. In contrast to the currently unpopular epistemic aspirations to neutral and universal description of knowledge, contemporary research, especially from the 1990s onwards, has disclosed the implicit and explicit subjectivity of any attempts to describe and organise knowledge (Smiraglia, 2014). Especially Olson’s (e.g. 1994; 2002) groundbreaking work played a key role in uncovering bias and marginalising effects of knowledge organisation systems exposed the politics of organising knowledge and paved the way to study intersectionality in information and knowledge organisation and description (Fox, 2016). As the studies of folksonomies show (Adler, 2009; Gartner, 2016), even seemingly factual or neutral, open and ‘democratic’ schemes of organising knowledge are political (Hjørland, 2020), and follow the logic of the politics of formalism outlined by Star (1995): they are abstractions, simplifications, and incorporate choices of what is important and what can be discarded. Such formal arrangements can support existing practices but they have also capability to upgrade and redefine the socio-technical infrastructure of activities (Millerand and Bowker, 2008) with very real consequences (Moncrieffe and Eyben, 2007).

Beyond the generic factors outlined by Star (1995), there are several practical reasons why knowledge organisation systems, including metadata and paradata schemes, exercise their politics as they do. One is that language (in a broad sense cf. e.g. Manovich, 2001) is political. Even if concrete aspects of things tend to be less controversial to describe than abstract ones, it is difficult to find words that are and would remain neutral and reasonably resilient to change (Radio, 2018). The contemporary tendencies to mix different descriptive facets in metadata schemes (Gnoli, 2012; Radio, 2018) complicates the matter even further by obscuring what aspects of things are described.

Another parallel factor is the contextuality of all metadata and paradata. Contextuality pertains both to how knowledge organisation schemes often describe domain-specific matters and discourses (Szostak et al., 2016), and how different types of data (Jansson, 2018) and descriptive schemes themselves are contextual and incommensurate with each other (Jansson and Huvila, 2019). Further, the intended and potential scope of metadata or paradata is not necessarily understood well enough. For instance, personal descriptions of things are not necessarily meant to inform others (Feinberg, 2011). Moreover, descriptions are produced with different objectives in mind. Textbooks (e.g. Foulonneau and
Riley, 2008; Haynes, 2018; Gartner, 2016; Smiraglia, 2014) enumerate typically a fairly standard set of purposes for metadata, including resource identification and description, information retrieval, information resource management, management of information rights, supporting learning, research and working with information. Haynes (2018) introduces information governance as an additional category. The comparisons of folksonomies, and formal classification and indexing systems have made it more apparent than ever (e.g. Jansson, 2018; Adler, 2009) that similarly to how individuals and groups of users have different metadata (and paradata) needs (Hu et al., 2019), the explicit and implicit purposes of producing and using particular descriptors and descriptive schemes have repercussions on what is described and how.

As the unfolding of the 'hows' of how knowledge organisation systems exercise their politics inevitably indicates, the politics of paradata has repercussions beyond the technical viability of standards and descriptors. It also ties into the ethics, power and ownership of resources. Considering the intricacy and complexity of influences different actors, including institutions and individuals, have on practices of description and documentation (e.g. Mayernik, 2015; Bates, 2018), the key questions are who gets to organise and describe, who pays for the work, who gets access to the descriptions and who benefits from them. Moreover, an additional, equally relevant question is who owns the space that the paradata is meant to describe and who gets to decide how the descriptive standards are developed and put to use (Haynes, 2018).

Documentation of archaeological visualisations

Before engaging with standards and recommendations relevant to archaeological visualisations, we give a brief overview of archaeological visualisations, their documentation and role in archaeological work.

Different types of visualisations – site drawings, maps, plans, photographs, object illustrations, physical and digital models – have always been central to archaeological documentation and dissemination of knowledge (Moser, 2012; Watterson, 2015). Similarly, the acts of producing images and visualising from map making, drawing and photography to modelling have been at the heart of archaeological knowledge making (Morgan and Wright, 2018). Images and visualisations have a major impact on the knowledge about (pre-)history. Since it is not always possible to see or feel archaeological remains in person, an image can become, in a sense, more real than what it represents (Moser, 2012).

Many archaeological visualisations, whether they are 3D models (Champion, 2018) or other visual artefacts, lack both descriptive metadata and process information (Piccoli, 2017). The attention to different types of visual information has also been somewhat unevenly distributed in the literature. Spatial information and 3D has been discussed more than many other forms of visualisations such as
drawings or photographs (Huvila, 2019; Morgan and Wright, 2018). The general intricacy of describing visual material (Lim and Liew, 2011; Huvila, 2019) and the difficulty to align different metadata schemes and descriptive needs to each other (von Schwerin et al., 2016) pertains also to archaeological visualisations. Another problem is the uneven availability and slow emergence of guidelines and metadata schemes for describing visual artefacts (Huvila, 2017).

The particular relevance of paradata for archaeological visualisation stems from that the insights developed during the process of creating visualisations can be more important than the final product (Morgan, 2009). The rapid development and proliferation of new, especially digital methods adds to the importance of documenting not only what but also why and how (Edmond and Morselli, 2020). Paradata can help to understand what was done and how (Richards-Rissetto and von Schwerin, 2017; Edmond and Morselli, 2020), make projects and their outcomes sustainable (Edmond and Morselli, 2020), and verify hypotheses and the authenticity of models (Kastanis, 2019), and contribute to “informed, critical and qualified interpretation” (Ogleby, 2007). A parallel question to the need of paradata, is how and in which form it should be preserved. As a form of processual information, (Agrifoglio, 2015) paradata can be difficult to capture and document. Proposed approaches range from writing free text narratives, producing annotations, developing virtual research environments and using information visualisation (Richards-Rissetto and von Schwerin, 2017), to 3D scholarly editions (Papadopoulos and Schreibman, 2019), video diaries (Hodder, 2000), written reports (de Kleijn et al., 2016) and formal modelling of intellectual and argumentative processes (Doerr and LeBoeuf, 2007; Marlet et al., 2019). At the same time, the digitalisation of archaeological practices has invigorated the earlier interest in standardisation and formal description of archaeological reasoning and workflows (e.g. Ogleby, 2007; Giovannini, 2018; Kastanis, 2019 cf. e.g. Gardin, 1980).

Besides the formats of descriptions, opinions diverge also on whether the best approach to improve the quality of documentation is to provide guidance or to increase standardisation. Studies of archaeological metadata confirm that the level of standardisation of descriptions and vocabulary is low especially between, but to a certain extent also within, investigation projects (e.g. Pavel, 2010; Oikarinen and Kortelainen, 2013; Henninger, 2018). There have been recurrent calls for increased standardisation (e.g. Quintero and Eppich, 2016; Gunnarsson, 2020) and specific new standards (e.g. Ogleby, 2007; Bendicho, 2013). In parallel, however, others argue that the complexity of the heritage field, its interdisciplinary and rapid development means that standards become too rapidly obsolete (Addison, 2007).

A parallel problem to obsolescence and finding a balance between too much and too little standardisation (e.g. Huvila, 2012) is that standards are not always implemented and used consequently (e.g. Molenda, 2020; Maron and Feinberg, 2018) and according to directives. Moreover, creating metadata takes is also often more time-consuming than anticipated and it requires skills both in the subject matter (Rejdovianova et al., 2018) and subject description (Llebot and Tuyl, 2019).
Standards for documenting archaeological visualisations

As noted earlier in this article, the existing standards and recommendations pertaining to archaeological visualisations make occasional, albeit as a whole, somewhat unsystematic references to the documentation of paradata. After reviewing current major standards and recommendations, we identified three broad categories of such references: 1) charters and recommendations, 2) metadata schemes and standards, and 3) conceptual models, exemplified here by CIDOC Conceptual Reference Model (CIDOC-CRM) and its extensions. The following three sections describe these categories before the text proceeds to explicate what and how the standards aim to standardise and how they represent processes.

Charters and recommendations

Much of the standardisation of archaeological visualisations conducted outside of individual organisations and information management systems has happened through charters and recommendations. The Venice Charter on the Conservation and Restoration of Monuments and Sites from 1964 notes that “all works of preservation, restoration or excavation” need to be documented in “analytical and critical reports, illustrated with drawings and photographs” including a documentation of “[e]very stage of the work of clearing, consolidation, rearrangement and integration, as well as technical and formal features identified during the course of the work” (Gazzola et al., 1964). Even if the text does not instruct how drawings and photographs should be produced or documented, it suggests that some types of drawings and photographs are a part of the apparatus of “analytical and critical” documentation of sites and monuments. Comparable general formulations on the role of illustrations can be found in several of the charters issued by the International Council on Monuments and Sites (ICOMOS) relating to conservation and restoration of sites and monuments (Petzet and Ziesemer, 2004).

Two recent documents originating from two international working groups, the London Charter (Denard, 2013) and Seville Principles (or the International Guidelines for Virtual Archaeology, International Forum of Virtual Archaeology, 2011; Bendicho, 2013) have focused specifically on computer-based archaeological and heritage documentation and visualisations. The objectives of the London Charter include an aim to “[e]nsure that computer-based visualisation processes and outcomes can be properly understood and evaluated by users“ (The London Charter Organisation, 2009). The Seville Principles underlines similarly the importance of “scientific transparency” i.e. how visualisations need to be testable and confirmable or falsifiable by others. The document adds that “the incorporation of metadata and paradata is crucial to ensure scientific transparency of any virtual archaeology project. Paradata and metadata should be clear, concise and easily available. Besides, it should provide as much information as possible. The
scientific community should contribute with international standardization of metadata and paradata” (International Forum of Virtual Archaeology, 2011).

A fourth document, *ICOMOS Charter for the Interpretation and Presentation of Cultural Heritage Sites* (Silberman, 2008) notes that “the information sources on which [...] visual renderings are based should be clearly documented and alternative reconstructions based on the same evidence, when available, should be provided for comparison” and that “[i]nterpretation and presentation programmes and activities should also be documented and archived for future reference and reflection” (Silberman, 2008, 8). Here can also be mentioned the manifest (Frischer et al., 2002) for a Cultural Virtual Reality Organization (CVRO), a proposed organisation to, among other issues, develop “aesthetic, scientific, and technical standards for cultural virtual reality models” and a common “philology”, or conventions and semantics, for metadata on archaeological 3D visualisations. Photography and drawings lack similar charters, conceivably at least partly because of the longer history of these media types that precedes the era of cultural heritage charters.

Similarly to other charters in the cultural field (Luxen, 2004), all of the above mentioned charters are open to multiple interpretations. Moreover, they overlap with each other and provide only limited practical guidance. For example, none of the texts goes into specific detail in what metadata items should be included to achieve the desired level of documentation. Perhaps unsurprisingly, as Sköld et al. (work in progress) show, the references to the charters have a tendency to remain as fairly generic and the proposed means to implement them vary considerably. At the same time, the charters are empathetically political documents, or as Wells (2007) remarks, discursive acts, that construct very particular types of significances in that their focus is on describing a desirable future state of affairs rather than providing technical guidance.

Besides the collective efforts to write charters, several individual authors have produced lists of technical and policy recommendations based on their practical or scholarly work. There is a relatively comprehensive body of literature on archaeological illustration with such advice (incl. Dorrell, 1994; Adkins and Adkins, 1989; Baltzavias, 2006) – even if as Morgan and Wright (2018) justly remark that despite its central role in archaeological work, visual media is often marginalised in archaeological contexts. A part of these recommendations enumerate principles of how visualisations should be documented whereas others advocate for the adoption of specific standards (e.g. for photographs, Atarashi et al. 2000; Toffalori 2016, or spatial data, Shaw et al. 2009; McKeeag et al. 2020, 2019). Champion (2018) suggests that an infrastructure for virtual heritage models should provide documentation of 1) data accuracy, 2) format limitations (i.e. known limitations of used digital formats), 3) provenance, 4) community protocols (who gets to access the material and who decides), 5) authenticity, 6) cultural presence (cultural significance and value of the original site or artefact), 7) evaluation data, and the 8) purpose of creating a model. In addition to dedicated texts, comparable recom-
mendations can be found in the extensive archaeological handbook literature (e.g. Corsi et al., 2016; Barratt, 2016; Dorrell, 1994; Drewett, 1999).

**Metadata schemes and standards**

Despite the asserted importance of adequate well-structured metadata, only a few general, widely adopted standards specific to archaeological visualisations exist. This is especially obvious when compared to the general profusion of metadata standards in the cultural field (cf. Skinner, 2014).

Many popular standards used to describe archaeological visualisations cover archaeology and heritage, and its digital and non-digital representations in broader terms, or focus on the documentation of heritage rather than its representations (e.g. Toffalori, 2016; Ryan, 2001; Signore, 2009; Aloia et al., 2017b). Many archaeological repositories use also generic standards, such as the Dublin Core metadata framework (e.g. Miller, 1999; Atarashi et al., 2000; Kulasekaran et al., 2014) for documenting digital objects, popular geoinformation standards (De Roo et al., 2013; Shaw et al., 2009) for spatial data, and various schemes for documenting cultural work, such as Getty vocabularies (Baca and Gill, 2015), the British museum documentation standard SPECTRUM (McKenna and Patsatzi, 2007), the Categories for the Description of the Works of Art (Harpring, 2019), and the Visual Resources Association (VRA) Core (Library of Congress, 2014) standard for the description of works of visual culture and images that document them. The Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images (CCO) (Baca et al., 2006) specifies a guideline for compiling a cataloguing record that can be implemented in standardised metadata schemes such as Dublin Core or VRA CORE.

Many of the archaeology specific metadata specifications, for instance, the British MIDAS (English Heritage, 2012) and Italian PICO (Scuola Normale Superiore di Pisa, 2007, 2011), are national standards and related to country-specific sites and monuments records (Ronzino et al., 2013). Some others are best characterised as proposals, such as the Cultural Heritage Markup Language (CHML), a formal language developed for documenting 3D reconstructions and reconstruction processes (Hauck and Kuroczyński, 2015), CHARM, a conceptual model of visual representation (Gonzalez-Perez et al., 2012; Apollonio and Giovannini, 2015), and the Extended Matrix/Framework (EM and EMF) for formal documentation of scientific processes underpinning archaeological virtual reconstructions (Demetrescu and Fanini, 2017). The large majority of the schemes are, however, developed and adopted by individual projects and repositories (Richards, 2009; McKeague et al., 2019; Carlisle and Lee, 2016). Despite the on-going standardisation work, influences that traverse from project and country to another (e.g. as vividly described in the work of Pavel, 2010), and occasional links to recommendations, charters and earlier standards (e.g. Dieckmann et al., 2010; von Schwerin
et al., 2016), archaeological documentation practices have been described for a good reason as parochial (Aitchison, 2017) rather than highly standardised.

Despite the prevalence of local variation, there are examples of successful archaeology-related metadata standardisation initiatives. One of them, explicitly described as successful (Champion, 2018), is the CARARE metadata scheme (Fernie et al., 2013). It is based (Fernie et al., 2013) on the British MIDAS heritage documentation standard (English Heritage, 2012), CIDOC-CRM (Doerr et al., 2007), Europeana Data Model (EDM) and CIDOC LIDO (Coburn et al., 2010) metadata harvesting scheme. Rather than being a foundational metadata scheme, it was developed for facilitating metadata aggregation to Europeana with a focus on collections rather than individual items (Papatheodorou et al., 2011). The ARIADNE catalogue model is another scheme describable as an interoperability standard with an explicit aim of bridging between local content standards. It was developed for the ARIADNE infrastructure to facilitate metadata aggregation relating to archaeological collections and focuses on resource-level metadata and authorship (publisher, creator, owner, responsible). While ARIADNE focuses on archaeological data, the broadly speaking comparable PARTHENOS Registry Data Model has a broader scope. It aims to bridge infrastructures and builds on existing standards in linguistics, humanities, heritage, history, archaeology and related fields including such generic registry standards as W3C DCAT and ISO-Cat, and schemes implemented by, for instance, CLARIN, META-SHARE, EHRI, LRE MAP, DARIAH, and CENDARI (Aloia et al., 2017b) projects. The resulting semantic framework, the PARTHENOS Entities, functions as a target data model for mapping metadata from multiple source infrastructures (Durco et al., 2018).

Many of the internationally influential standards are similar in that they focus on archaeological sites and monuments and their representations beyond visualisations. Many of them are interoperability schemes rather than content standards that echoes the shift of general focus in standardisation from aspirations to develop global documentation schemes to data interoperability and aggregation (e.g. Richards, 2009; Doerr et al., 2007; Aloia et al., 2017a) and mappings between different local standards and schemes (e.g. Meghini et al., 2017). As Richards (2009) points out, this does not, however, imply that a certain level of agreement on content level descriptions would not be necessary.

In addition to the published standards, the literature contains many proposals and enumerations of desirable metadata elements. A widely cited collection of guidelines is the Guides to Good Practice published by the UK-based Archaeology Data Service (ADS) in collaboration with the US-based DigitalArchaeological Record (tDAR) (Archaeology Data Service and Digital Antiquity, 2009–). Others include, for instance, Addison’s (2007) suggestion to include information on recording device parameters, data manipulation devices, environmental conditions, submitter, author and sponsor, date and location of data capture in virtual heritage metadata. Moreover, there is a sundry recommendations and proposals
of metadata schemes developed and implemented for specific archives or repositories (e.g. Polig, 2017; Ryan, 2001; von Schwerin et al., 2016).

Unsurprisingly, considering the ambiguity of views regarding the standardisation of archaeological documentation, the existing schemes have not avoided critique. Much of the critique focuses, however, either on details, or is expressed as a generic criticism of the limits of standardisation. For instance, Champion (2018) argues that the CARARE standard lacks an element for describing the cultural significance of documented assets included in his own list of recommendations, and explicit consideration of indigenous perspectives. Another comparable point of critique is the preferred naming of metadata elements (Hu et al., 2019) and the (lack of) consistency of how they are used (e.g. Carlisle and Lee, 2016).

**Conceptual models: CIDOC-CRM and extensions**

The third category of standards for the documentation of archaeological visualisations is conceptual models. Due to the dominant position of the CIDOC Conceptual Reference Model (CIDOC-CRM) in archaeology and heritage, it is used in the following to exemplify this category. CIDOC-CRM is a formal ontology (in the sense of Guarino, 1999) developed under the auspices of International Committee for Documentation (CIDOC) of the International Council of Museums (ICOM) to facilitate data exchange in the cultural heritage domain. The model defines relationships rather than terms to support integrating existing and future metadata structures and data schemes (Doerr et al., 2007). The version 1.0 of the CRM (Crofts et al., 1998) based on CIDOC information categories (Grant et al., 1995) was published in 1998. A standardisation process was started in 2000 and the model was accepted in 2006 as an ISO standard ISO21127:2006 (Doerr et al., 2007).

As the aim of CIDOC-CRM is to function as a general ontology of the cultural heritage sector, it has been complemented during the years with extensions to cover more specific aspects of heritage, including archaeological excavations (CRMarchaeo, Doerr et al., 2018), provenance metadata (CRMdig Theodoridou et al., 2010; Doerr et al., 2016), cross-research-infrastructure metadata (CRMpe Bruseker et al., 2017b), metadata about scientific observation, measurements and processed data in descriptive and empirical sciences (CRMsci, Doerr et al., 2014), about argumentation and inference making in descriptive and empirical sciences (CRMinf, Stead and Doerr, 2015), and ancient textual editions (CRMtex, Doerr et al., 2020).

Close to how paradata is defined in the literature, Niccolucci and Felicetti (2018) propose a combination of CRMsci, CRMdig and CRMpe for documenting heritage science i.e. scientific activities in support of conservation, access and interpretation of cultural heritage. Guillem and colleagues (2015) put forward a comparable CIDOC-CRM based scheme for documenting reconstruction and reconstruction-related argumentation processes, and Amico et al. (2013), a combi-
Paradata in Documentation Standards

Many ontologies and documentation schemes (e.g. Aloia et al., 2017a; Henninger, 2018; Felicetti et al., 2013; Marlet et al., 2019) from different heritage-related disciplines has been mapped into CIDOC-CRM (Moraitou et al., 2018). The model has also been used as a basis for developing new ones (e.g. Padfield et al., 2019; Giovannini, 2018; Moraitou et al., 2018; Felicetti et al., 2013). Moraitou and colleagues note that the current proliferation of extensions and alternative mappings means that it is possible to select the most useful semantic representations among several options. They argue that this can improve the quality of preserved information and facilitate its reuse. At the same time, the burgeoning of alternatives means that it is both more difficult and important to have an overview of available options (Moraitou et al., 2018).

Formal ontologies such as the CIDOC-CRM have multiple benefits when compared to protocol-based knowledge organisation systems such as metadata schemes (Bruseker et al., 2017a). Even if they do not per se solve the intricacies of describing and naming things, they can improve machine readability of descriptions and the transparency of ontological assumptions of how concepts are related to each other. The typical critique of CIDOC-CRM and formal ontologies tend to relate to their complexity and the inconsistencies in how they are interpreted and applied in practice (e.g. Bruseker et al., 2017a; Nussbaumer and Haslhofer, 1999, also e.g. Peponakis, 2012).

Processes in the standards

After a brief review of major categories of standards and recommendations relevant to archaeological visualisations, we proceed to inquire into how processes are represented in a selection of these guidelines (for the sake of simplicity, referred hereafter collectively as ‘standards’). In comparison to dedicated standards for representing scientific processes like CERIF (Jörg et al., 2012) and OBOE (Madin et al., 2007) and comparable discipline-specific schemes (e.g. Niccolucci and Felicetti, 2018), the above reviewed archaeological standards engage with processes on multiple levels of specificity. The Table 1 synthesises observations on a set of standards and guidelines pertinent to archaeological visualisations, how they engage politically with processes and what means they provide for documenting processes in the form of paradata categories. As a limitation, it is necessary to emphasise that the selection does not cover all conceivable standards or guidelines that could be used for the purpose. The review is also limited to providing an overview of the general approach of the standard rather than a detailed list of all metadata elements or options. In spite of these provisions, we argue that the enumeration provides a useful glimpse of how contemporary standards and guidelines expound processes. The model for analysis also provides a basis for scrutinising additional schemes and their intrinsic and extrinsic politics.
Table 1: Standards relating to archaeological visualisations and how they engage with process metadata.

<table>
<thead>
<tr>
<th>Standard or guideline</th>
<th>How processes are represented</th>
<th>What space the standard encompasses</th>
<th>Examples of references to other standards and guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>London Charter (The London Charter Organisation, 2009)</td>
<td>Documentation of decisions, development of documentation strategies, sources, paradata, relationships between research sources, implicit knowledge, explicit reasoning, and visualisation-based outcomes, description of methods</td>
<td>Computer-based visualisation methods in research, communication and preservation of cultural heritage</td>
<td>ADS and tDAR Guides to Good Practice, CVRO</td>
</tr>
<tr>
<td>ADS &amp; tDAR Guides to Good Practice (Archaeology Data Service and Digital Antiquity, 2009-)</td>
<td>Project and resource-level documentation and administrative metadata, file-level metadata according to the specific type of visualisation e.g. workflow documentation, creator, instruments, purpose, software and devices, results</td>
<td>Archaeology, especially parties intending to deposit data in ADS or tDAR</td>
<td>Multiple incl. CARARE, CIDOC-CRM (incl. extensions), Dublin Core</td>
</tr>
<tr>
<td>Standard or guideline</td>
<td>How processes are represented</td>
<td>What space the standard encompasses</td>
<td>Examples of references to other standards and guidelines</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------------------</td>
<td>------------------------------------</td>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>CDWA (Harpring, 2019)</td>
<td>Creator, creation date, place, cultural context of creation, conservation/treatment and condition/examination history, (historical) events associated with the object, critical responses, ownership/collection, exhibition/loan and cataloguing history</td>
<td>Works of art</td>
<td>E.g. VRA Core, CCO</td>
</tr>
<tr>
<td>VRA Core</td>
<td>Incl. agent, cultural context, source, date (of creation, design, production, presentation, performance, construction, or alteration), description, location, relations, techniques</td>
<td>Works of visual culture as well as the images</td>
<td>CDWA, CCO, Dublin Core</td>
</tr>
<tr>
<td>EM/EMF (Deme trescu, 2015; Deme trescu and Fanini, 2017)</td>
<td>Schemaless semantic graph (of a reconstruction process)</td>
<td>Virtual reconstructions</td>
<td>E.g. CIDOC-CRM, CHARM (as related work)</td>
</tr>
<tr>
<td>CHARM (Gonzalez-Perez et al., 2012; Gonzalez-Perez, 2018)</td>
<td>Performative entities, manifestations, occurrences (activities - processes, actions, projects, tasks)</td>
<td>Cultural heritage</td>
<td>E.g. CIDOC-CRM</td>
</tr>
<tr>
<td>CHML (Hauck and Kuroczyński, 2015)</td>
<td>Historic events, research activities</td>
<td>3D reconstructions</td>
<td>Multiple XML-based (incl. RadianceML)</td>
</tr>
<tr>
<td>Standard or guideline</td>
<td>How processes are represented</td>
<td>What space the standard encompasses</td>
<td>Examples of references to other standards and guidelines</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------------------</td>
<td>------------------------------------</td>
<td>----------------------------------------------------------</td>
</tr>
<tr>
<td>CARARE (Fernie et al., 2013)</td>
<td>Activities, collections, heritage assets and their constituents and relations to each other</td>
<td>Heritage assets, in version 2.0 compatibility with Europeana Data Model</td>
<td>E.g. MIDAS, CIDOC-CRM, CRMdig, LIDO</td>
</tr>
<tr>
<td>CIDOC-CRM (Doerr et al., 2007)</td>
<td>Events (incl. activities) and their constituents and related entities</td>
<td>Integration, mediation and interchange of heterogeneous cultural heritage information (in museum context)</td>
<td>No explicit references in the definition</td>
</tr>
<tr>
<td>CRMsci (Doerr et al., 2014)</td>
<td>Activities, alterations, beginnings of existence and their constituents/related entities</td>
<td>Integrating metadata about scientific observation, measurements and processed data in descriptive and empirical sciences</td>
<td>Multiple incl. CIDOC-CRM, OBOE,</td>
</tr>
<tr>
<td>CRMinf (Stead and Doerr, 2015)</td>
<td>Activities, alterations, beginnings of existence and their constituents/related entities</td>
<td>Integrating metadata about argumentation and inference making in descriptive and empirical sciences</td>
<td>Multiple incl. CIDOC-CRM, CRMsci</td>
</tr>
</tbody>
</table>

Due to their generic nature, it is not surprising that charters and recommendations often include generic formulations relating to the documentation of processes. The London Charter posits that “documentation of the evaluative, analytical, deductive, interpretative and creative decisions made in the course of computer-based visualisation should be disseminated in such a way that the relationship between research sources, implicit knowledge, explicit reasoning, and visualisation-based outcomes can be understood.” (The London Charter Organisation, 2009, 8). The document contains comparable provisions on the documentation of methods and development of documentation strategies (The London Charter Organisation, 2009) but even they remain – likely on purpose – on a fairly abstract level to make the document as a whole valid for diverse stakeholder groups (Denard, 2013). The Seville Principles are more specific in that “clear, concise and easily available” internationally standardised metadata and paradata are vital for ensuring scientific transparency. The text refers further to earlier
principles on how heritage sites should be documented and to a need to establish an international “database” of exemplary projects (International Forum of Virtual Archaeology, 2011).

Even if metadata standards and schemes tend to focus on objects (Signore, 2009), they provide varying means to represent processes. The ADS and tDAR Guides to Good Practice is a comprehensive document with somewhat varying instructions for the documentation of multiple types of visualisations and visual data. For example, according to guides, the documentation of vector-based graphics should contain information on provenance (origins of the image), software, purpose of creation, conventions used in the document, creator and creation date, whereas the guidelines to document 3D visualisations refer to CARARE and CRMdig standards and description of workflows. In addition to document-specific metadata, the Guides to Good Practice advise to include project and resource, administrative and file-level process metadata (Archaeology Data Service and Digital Antiquity, 2009-). Unlike many others, the CARARE standard incorporates a specific field for provenance information, instructed to be provided as a narrative (Fernie et al., 2013). CCO asks who created the work or what is its culture of origin, where and when it was created and discovered, where it has been located, where it is now, and what tools and techniques and creative activities (e.g. creation, design, execution) were involved (Baca et al., 2006). However, in the majority of the standards and repositories surveyed by the PARTHENOS project, metadata covers authors, agents or owners of collections (Aloia et al., 2017b) following the rationale of describing basic provenance data (creator, date, method, see Bizer et al., 2011) rather than means or expectations to elaborate processes in detail.

A common feature of the standards, independent of their emphasis of actors and contexts (e.g. CCO) or events (as in CENDARI CENDARI WP6, 2013a,b, and VRA CORE, Library of Congress, 2014), is their interest in the historical (e.g. CENDARI, VRA CORE) and curatorial processes (e.g. PICO, CCO) relating to archaeological sites, monuments and objects rather than to their representations. This impression is strengthened by review of cataloguing examples provided as a part of the documentation of the standards.

Ontologies unfold as a third distinct type of standard. Instead of elucidating what counts as significant process-related information, they provide guidance to approaching the ontological question of what is a process and how to represent it. CIDOC-CRM approaches the question through an events-based approach with Temporal Entity and Event hierarchies of classes that are used to represent events and activities. Together with classes referring to persistent items, time-spans, places, dimensions, spacetime volumes and primitive values they can be used to describe processes and their constituents. The core standard with its documentation examples does, much like CENDARI and CVRA CORE, emphasise historical events whereas many of the extensions focus specifically on scholarly processes. The work of, for instance, Guillem et al. (2015) and Niccolucci and Felicetti (2018) show that the model and its extensions provide versatile tools for representing
and describing archaeological scholarly processes – including those relating to the making and use of archaeological visualisations.

Discussion

This paper has inquired into a selection of recommendations, metadata schemes and ontologies pertinent to archaeological visualisations and unveiled a heterogeneous array of approaches and perspectives to document their provenance and processing history (i.e. paradata). They all enact a distinct take on the politics of what is important to describe and what can be left out. Both standards and charters have their own politics even if the latter are more explicitly political documents, commonly used as a basis for requirements engineering (Carrillo Gea et al., 2013) or developing standards (as the Seville principles explicitly suggest International Forum of Virtual Archaeology, 2011) – measures that are often experienced as neutral and factual (Bowker and Star, 2000; Star and Lampland, 2009).

To discuss the influence of politics and bias in the reviewed set of documentation standards, we use three criteria: 1) what is standardised i.e. the whatness of the objects they are helping to maintain, 2) what types of spaces the different standards create, and 3) what becomes standardised in the process. The first apparent observation is – considering the lively debate – the relative lack of dedicated standards for documenting archaeological visualisations and their related paradata. This can be perhaps explained partly in terms of the absence of a single domain pertaining to that field (Doerr, 2009; Hjørland and Hartel, 2003) and the consequential, relative lack of consensus on the ‘whatness’ (Denis, 2018) of archaeological visualisations and the significant aspects of their processing history. As a result, the schemes standardise or make recommendations on the standardisation of widely different aspects of visualisations and their processing history.

A lack of consensus on what archaeological visualisations are, is however, unlikely the only or decisive cause of the diversity. Another, perhaps a more pertinent explanation is the heterogeneity of who develops standards, for whom and for what purpose. Paradata, like metadata, are more likely to be of interest for specialists rather than for an average archaeologist who would probably be more inclined to stress the importance of tools instead (Benardou and Dunning, 2018). The needs and priorities of specific stakeholder groups are visible both in the more openly political recommendations and in the metadata standards that are often formulated in a more neutral, matter-of-factual tone. In infrastructural standards (e.g. CARARE, Guides to Good Practice) the focus is unsurprisingly on digital provenance (Ross, 2018) whereas curatorial standards (e.g. VRA Core, CDWA) put emphasis on the documentation of custodial processes. There is nothing inherently wrong with such emphases but they do make flagrantly visible the contextual and political nature of metadata schemes. In this respect, the calls for increased flexibility (e.g. Löwenborg, 2007) and standardisation (e.g. Quintero
Paradata in Documentation Standards

and Eppich, 2016) do not need to be as contradictory as they are portrayed by their proponents. The key question is how to navigate between standardisation and flexibility (Huvila, 2012) in relation to the purposes of the standards and their users – or what to standardise, and when (and how) to open up for diversity to facilitate specific types of future activities. As a whole, it is important to document the provenance and processing history in more detail and to think increasingly explicitly about who the likely users of information are (e.g. Edmond and Morselli, 2020). Similarly, an increased emphasis on managing the social aspects of archaeological information (Huvila, 2019) is undoubtedly useful – but perhaps even more important than to try to anticipate the future would be to be cognisant of the premises of existing descriptions and their guiding principles. Further, it would be of value to extend this critical awareness to rethinking and describing the theoretical foundations of current and emerging standards and data descriptions as enabling one type of description at one point in time.

A second remark that can be made on the basis of the review is that the standards create widely different spaces of operation. Some of the standards are or have been local to specific systems like particular national sites and monuments records, project databases, or digital libraries like Europeana. Choosing between standards and pairing up with a particular information infrastructure is a political choice (cf. Star, 1995; Hjørland, 2020) with consequences even if it would be backed by seemingly matter-of-factual arguments. Extending the emphasis of Moraitou et al. (2018) of the importance to be aware of CIDOC-CRM and in general of metadata schemes, it is equally important to be cognisant of what type of standards they are: recommendations, metadata or paradata schemes, or conceptual models, and to what type of space they are referring. A standard makes the documentation compliant to and at home in its space in a similar sense how Pétursdóttir (2020) comments the fate of archaeological objects in the teeth of heritage discourse.

A third question to ask a standard is what becomes transparent in the process of standardisation. If a standard is followed to the point – which is not always the case as they embed and are embedded in local practices (Maron and Feinberg, 2018) – it guides to provide a description of a specific perspective to a process rather than the process in its entirety. For example, the most of the reviewed standards represent processes as chains or networks of events and/or activities that essentially turn them to enumerations of distinct actions and incidents rather than, for instance, a continuum. Even when standards, like CRMarchaeo (Doerr et al., 2018), explain that they are neutral to accommodating different traditions and schools of thought, they are still imposing a lot – quoting Hanseth and Monteiro, “in complex and non-transparent ways” (Hanseth and Monteiro, 1997, p. 183). It is not a secret that standards impose perspectives – they are supposed to do so – but it is still worth being aware of how standards and their accompanying regimes not only enable specific trajectories but also constrain them (Fuenfschilling and Binz, 2018). Moreover, as the present review of standards shows, the transparency varies
depending on the type of the standard, its intended purpose, and the infrastructural and organisational context where it was developed.

As a whole, the analysis underlines the critical importance to understand standards, what they standardise, for whom, what is their context or space of standardisation and what becomes standardised and what is ‘maintained’ (Denis, 2018) when they are applied. Understanding metadata as a meshwork rather than a thing becomes increasingly significant in the on-going shift to entity rather than object-based linked metadata (Smith-Yoshimura, 2020) and links to the calls for a critical stance towards “naive” perceptions of paradata (Havemann, 2012) and metadata as complete and permanent representations of data and its processing history. Various strategies have been proposed to remedy the biases and shortcomings of existing catalogues and data descriptions. These include user-generated metadata (Conradi, 2010), automatic annotation (Foley et al., 2017) and, for instance, deliberate correcting of descriptions through ‘queering of the catalogue’ (Drabinski, 2013). Even if many of the proposed efforts are promising, remedies come easily with their own biases. However, if different types of descriptive data are treated from their distinctive premises (Jansson, 2018) and combined with an adequate understanding of the space where applied data standards are operating and what becomes transparent when they are used, combinations of diverse types of data descriptions can be expected to have real potential to add perspectives (Freund and Butterworth, 2008; Stvilia et al., 2012) to descriptions, and consequently, to contribute to the findability of described assets and the efficiency of information retrieval.

Finally, beyond elucidating what standards are aiming at documenting and where, there is one more question to explore that can provide insights into what is attainable with contemporary documentation standards: who decides what the standards’ aims and the spaces of operation are. A partial answer is that it is the standard developing organisations and communities. Currently, a large part of the standardisation effort is in the hands of multinational European projects. Even if the projects tend to be large, with up to hundreds of participating organisations and individuals, and comprehensive outreach programmes, the development is still driven and coordinated by a fairly small number of key actors with many possible voices that might be and are probably missing. In broader terms, metadata (Gartner, 2016) and paradata themselves are ideologies and political promoters of interoperability (DeNardis, 2011) – the desirability that resources are described and identifiable, information is retrievable, manageable and managed, and (re)used for multiple purposes. In a narrower sense, it is worth being observant to disciplinary variants of this broader political idea. In addition to assessing standards imported from neighbouring disciplines in the context where they are applied rather than from the perspective of the discipline of their origin (Huggett, 2013), it is useful to remember that the same should apply to standards developed within a neighbouring archaeological frame of reference. Otherwise a specific activity within archaeology (e.g. landscape archaeology or pottery studies) – and in a broader...
sense, any borrowing field – risks to become ‘colonised’ and subject to the politics of another activity (e.g. data curation or spatial analysis) rather than capable of developing their own metadata and paradata policies.

Conclusions

Before concluding remarks, it is appropriate to stress that the purpose of the present analysis was not to evaluate the schemes for eventual shortcomings but to investigate into the politics of paradata and the representation of specific facets of data. It is apparent that the primary concern of most of the reviewed standards is not to document information and data making processes. This explains why processes are often either completely absent or only partially addressed in the schemes and their accompanying documentation. At the same time, however, many of the standards like the CIDOC-CRM and its extensions incorporate structures to represent processual information to an extent that they can effectively substitute dedicated standards (e.g. Niccolucci and Felicetti, 2018). Moreover, as many of the reviewed schemes show, they can be combined with each other to achieve a broader coverage of details.

The analysis also suggests that the inescapably political nature of knowledge organisation deserves to be taken seriously when working with paradata. Instead of agonising that all paradata (and metadata) is biased, bad and needs to be fixed, it is more important to try to understand the contexts, choices and assumptions that underpin the development of specific descriptive schemes and their practical consequences. Accordingly, as the use of existing metadata (Feinberg, 2017) and paradata, also the use of metadata and paradata schemes should be sensitive to what they are and how they have been developed. If a standard is doing what it was expected to do (assuming that its users are following it to a reasonable degree) – for instance, facilitating information retrieval in a museum context or harmonising data for inclusion in a specific database – it is perhaps good enough and can be used productively – and ethically as Olson (2000) underlines – and does not need to be changed. As Hjørland (2020) reminds, to assume that standards can be corrected is to deny their political nature and to believe that there is an unbiased apolitical master standard to be found.

When considering the practice of documenting the processing history of archaeological visualisations and other data, the present review of standards shows also that there are limits to how far a single standard stretches. The explicit and implicit assumptions on what types of things or processes a standard should standardise, and for what purpose, set limits to its general applicability. A standard for documenting ‘activities related to digital reconstructions of cultural heritage artefacts’ turns a process into a series of ‘activities related to digital reconstructions of cultural heritage artefacts’. Besides the whatness of a standard, a comparable aspect is the space within which the standard is meant to operate and it aims
to cover and control, and when put in practice, where it is used. If a standard is developed to cover ‘museum knowledge’, its politics is geared to the very specific space of museum knowledge even when its outspoken aim would be to aim at a broader scope of relevance. Therefore a final question to ask before embracing a particular standard should be if it can be employed to document what needs to be documented (e.g. paradata), and if it provides the expected type of transparency.

Acknowledgements

We would like to thank two anonymous reviewers and the editors of this special issue for their invaluable comments on earlier versions of this text. This work has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme grant agreement No 818210 as a part of the project CAppturing Paradata for documenTing data creation and Use for the REsearch of the future (CAPTURE).

References


Archaeology Data Service and Digital Antiquity, eds. 2009-. Guides to Good Practice. York and Tempe, AZ. URL https://guides.archaeologydataservice.ac.uk/g2gp/Main.


Gonzalez-Perez, C. 2018. *Information Modelling for Archaeology and Anthropology*. Cham: Springer.


Marlet, O., E. Zadora-Rio, P.-Y. Buard, B. Markhoff, and X. Rodier. 2019. The archaeological excavation report of Rigny: An example of an interoperable logi-

Maron, D. and M. Feinberg. 2018. What does it mean to adopt a metadata stan-
dard?. *JDOC* 74 (4):674–691.


McKeague, P., R. van’t Veer, I. Huvila, A. Moreau, P. Verhagen, L. Bernard,
A. Cooper, C. Green, and N. van Manen. 2019. Mapping our heritage: Towards
a sustainable future for digital spatial information and technologies in eu-

McKenna, G. and E. Patsatzi, eds. 2007. *SPECTRUM: The UK Museum Documen-

Meghini, C., R. Scopigno, J. Richards, H. Wright, G. Geser, S. Cuy, J. Fihn,
B. Fanini, H. Hollander, F. Niccolucci, A. Felicetti, P. Ronzino, F. Nurra,
C. Papatheodorou, D. Gavrilis, M. Theodoridou, M. Doerr, D. Tudhope,
C. Binding, and A. Vlachidis. 2017. Ariadne: A research infrastructure for

Miller, P. 1999. The importance of metadata to archaeology: One view from within
the Archaeology Data Service. In *Proceedings of the CAA 1997*, eds. L. Ding-
wall, S. Exon, V. Gaffney, S. Laflin, and M. van Leusen, pp. 133–136. Oxford:
Archaeopress.

Paris: ÉAC.

Molenda, A. 2020. *The Use of Preservation Tools Among Dutch Heritage Organiza-
tions*. The Hague: DDHN.


Morgan, C. and H. Wright. 2018. Pencils and pixels: Drawing and digital media in

Morgan, C. 2009. (Re)Building Çatalhöyük: Changing Virtual Reality in Archae-


Sköld, O., Börjesson, L., Huvila, I. work in progress. Interrogating paradata.


Wells, J. 2007. The plurality of truth in culture, context, and heritage. City&Time 3 (2).