

A Multiple Perspective Account of Digital Curation for Cultural Heritage: Tasks, Disciplines and Institutions

Tugce Karatas
tugce.karatas@unito.it
Università di Torino
Torino, Italy

Vincenzo Lombardo
vincenzo.lombardo@unito.it
Università di Torino
Torino, Italy

ABSTRACT

Cultural heritage management is a multiple-perspective enterprise where several disciplines and practices contribute to successful dissemination and communication. Digital data in support of cultural heritage management are addressed by the digital curation process, which has been emerging to account for the diversity of disciplinary communities and cultural heritage organizations. Digital curation addresses the diversity of participating skills and practices by working on the relationship between the cultural heritage objects and their digital counterparts. In particular, the innumerable initiatives for providing access to cultural heritage data are ideally coordinated by digital curation and are part of the process since the beginning. However, some thorough reflections on its role and implementation in cultural heritage institutions yet lack. In this paper, we provide a survey of the digital curation process, by unpacking the component curatorial tasks, with the solutions that have been proposed in the literature and in the application projects to account for the multiple perspectives at hand.

CCS CONCEPTS

• **Human-centered computing** → Collaborative interaction; Human computer interaction (HCI); • **Information systems** → Data management systems; Digital libraries and archives; • **Social and professional topics** → Management of computing and information systems.

KEYWORDS

digital curation; cultural heritage; interdisciplinary process; project workflow

ACM Reference Format:

Tugce Karatas and Vincenzo Lombardo. 2020. A Multiple Perspective Account of Digital Curation for Cultural Heritage: Tasks, Disciplines and Institutions. In *Adjunct Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '20 Adjunct)*, July 14–17, 2020, Genoa, Italy. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3386392.3399277>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

UMAP '20 Adjunct, July 14–17, 2020, Genoa, Italy

© 2020 Association for Computing Machinery.

ACM ISBN 978-1-4503-7950-2/20/07...\$15.00

<https://doi.org/10.1145/3386392.3399277>

1 INTRODUCTION AND BACKGROUND

Cultural heritage is a very comprehensive field, which, in its broader meaning, includes tangible, intangible, and digital heritages. It plays an important role in the representations of societies' identities, belongings and behaviours [23]. The management of the cultural heritage assets is a complex challenge that addresses many competences from diverse disciplines (such as, e.g., management, archaeology, chemistry), with information technologies (IT) that are more and more pervasive, transversely. The most frequent impacts of the digital technologies concern 3D reconstruction of physical items, data analysis and visualization, information representation and sharing; specialised workflows occur in measuring and analysing tangible assets through recording and diagnostic methods, respectively [18]. However, although such a pervasiveness of information technologies, yet cultural institutions, on the one hand, often neglect their importance [38], on the other, are challenged by the archiving and preservation of their digital assets, beyond the mere digitisation act [34]. The porting of IT functions to the specific nature of the cultural heritage has also forced institutions to face the acquaintance with a number of software environments, in order to manage the entire scope of the related digital materials [8]. Also, each digital material is related to some discipline, which brings one further perspective to the whole digital process.

In this context, digital curation has emerged as a viable solution in the coordination of the representation and management of the digital information related to cultural heritage: in particular, digital assets undergo selection, preservation, maintenance, collection and archiving, with the further requirement of adding value for subsequent exploitation [41]. However, while the research and practice of digital curation have been continuing to mature, relatively little empirical, comparative research and the consequent awareness on the tasks that compose the digital curation process has been achieved to date [8]. According to Harvey's cookbook [21], technical obsolescence or fragility, lack of resources, ignorance of good practices, and uncertainty over appropriate infrastructure constitute serious risks to the practice of cultural heritage specialists. Also, the responsibilities involved in digital curation are to be shared across different institutions and communities and are subject to continuous change, because it qualifies as an ongoing process, not a one-off action [1]. The corresponding chain of activities require the appropriate financial, political, and technical infrastructures to ensure that digital curation can be continued over the long-term [36].

Recent research on conceptualisation and practice of digital curation shows that there are both overlaps and gaps between digital curation tasks and digital tools in various cultural heritage institutions [8, 34, 39]. In order to respond to such a requirement, some

researchers [2, 13] have attempted a systematization of best practices and have devised theoretical frameworks to model digital curation in detail. Here we adopt a more abstract point of view, with the goal of investigating how multi-disciplinary collaboration affects the individual tasks that compose digital curation.

In fact, in the context of cultural heritage, curatorial work with digital technologies is often far less linear and visible than work with analogue collections and processes frequently cut across traditional organizational boundaries [9]. Workflows allow stakeholders to measure their progress, to discern vulnerabilities, to encourage documentation, to develop standards, and to identify tools and services; thus they help preserve data's authenticity, reliability, and usability [21, 22]. The contribution of this work is to describe the digital curation and highlight the issues on personalised data. Therefore, this research address a comprehensive list of tools, procedures, and standards that are in use for digital curation in order to take account of the multiple perspective aspect of the cultural heritage projects. It will discuss the preliminary findings and share some insights with a conclusion with the future work and planned steps in our research.

2 WHAT IS DIGITAL CURATION?

First used in 2001¹, digital curation encompasses the creation and management of digital assets during their entire lifecycle [25, 41]. It consists of "actively managing data [...] with the aim of supporting reproducibility of results, reuse of and adding value to that data, managing it from its point of creation until it is determined not to be useful, and ensuring its long-term accessibility and preservation, authenticity and integrity" (Digital Curation Center - DCC - website²). This definition of digital curation inevitably addresses digital material that is linked to many other resources, which are stored in a range of formats over a number of large datasets as well as the static documents handled by libraries and archives. It bridges research, practice, and training across nations, disciplines, institutions, repositories, and data formats [19, 24].

In recent years, some researchers have been modelling the digital curation workflows from different institutional contexts [8], as a way to assess the evolution of the e-resource management processes [4], or to evaluate existing e-resource services in order to inform the selection and implementation of new e-resource management software [15]. A workflow embraces many activities: appraisal, classification, indexing, cataloguing, and authority management; enhancement, presentation, publication, and dissemination, including modelling the user experience; finally, preservation and repository management [10]. To systematize goals and practices, a number of models have recently appeared in the literature from many institutions, such as Digital Curation Center Curation Lifecycle Model [22], I2S2 Idealized Scientific Research Activity Lifecycle Model [33], Data Documentation Initiative (DDI) Combined Life Cycle Model [6], ANDS Data Sharing Verbs³, DataONE Data Lifecycle [3],

Research360 Institutional Research Lifecycle [31], Capability Maturity Model for Scientific Data Management [11], UK Data Archive Data Lifecycle⁴.

Although they present some differences, these models largely overlap in the definitions of the tasks that compose the digital curation workflow. In this paper, we do not commit to any specific model, but assume an abstraction of the activities that is coherent with the several descriptions provided in most. Our abstraction is based on practical experiences and literature review across disciplines, in order to support truly interdisciplinarity.

3 TASKS OF THE DIGITAL CURATION PROCESS

Digital asset is tend to be used for a great variety of purposes in cultural heritage field by interdisciplinary users in various settings to be shared, accessed, analyzed, and stored using many different kinds of technology. This research presents an abstract workflow of digital curation, with the tasks and data formats that are concerned with the management and workflow of the digital assets. As can be seen in Fig. 1, as soon as some cultural heritage asset is acquired (e.g., a container with a cup shape in a trench of an archaeological site)⁵, the digital curation process builds some digital assets; these can be acquired from the asset (e.g., by photograph) or created from scratch (e.g., artistic 3D modelling for illustration purposes); this will produce some data, that we can generally name *raw data*, because they do not include a model yet (e.g., a digital photograph or X-RAY image). These data are enriched with metadata that reveal an interpretation of the asset at some level (e.g., region of the image, identified via a path joining the pixels, label with the tag "warrior"). Metadata can be revealing hidden knowledge about the raw data. Together with data they are processed to enrich the data (*processed data*) and used for the interpretation process to become a proper model of the cultural heritage domain (e.g., the wine cup is a Kylix of 7th century BC). These processes also proceed as a part of the repository of the data curation and preservation tasks, which also consist of the creation of a repository. The model and the relative knowledge about the domain at hand, can be used in the archives as a part of the data documentation and archiving (e.g., the digital image also receives the identifier of the physical Kylix). Also, it can become a publication and dissemination outcome (e.g., the digital image is part of a virtual collection exhibited through a website or a catalogue, and inserted in a scientific paper for a journal).

As depicted in Fig.1, now we go through the six common tasks for the data management cycle from the cultural heritage asset to the final outputs of the digital curation process, namely:

- (1) Data Creation or Acquisition
- (2) Data Processing and Modelling
- (3) Data Interpretation
- (4) Data Curation and Preservation
- (5) Data Documentation and Archiving
- (6) Data Publication and Dissemination

¹"Digital Curation: digital archives, libraries and e-science" seminar, <http://www.ariadne.ac.uk/issue/30/digital-curation>, (last visited on 3 April 2020)

²<http://www.dcc.ac.uk> (last visited on 3 April 2020)

³<https://www.ands.org.au/about-us/governance> (last visited on 13 April 2020)

⁴Dan Crane and Isabel Chadwick's Practical strategies for Research Data Management slide in 2017 http://www.open.ac.uk/blogs/the_orb/wp-content/uploads/2017/10/slides-RDM_PracticalStrategiesForRDM_2017-11.pdf (last visited on 3 May 2020)

⁵Even if the UNESCO classification includes three types of CH (tangible, intangible, digital), we do not make a strict point and address tangible heritage only to provide an example.

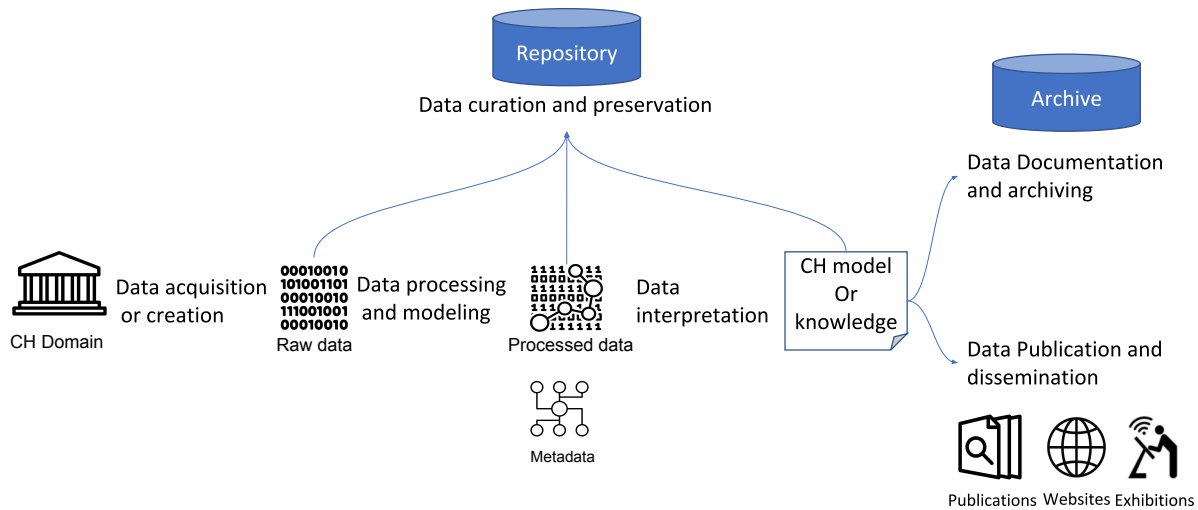


Figure 1: An abstract schema of the workflow of digital curation.

3.0.1 Data acquisition or creation. The first step of the digital curation pipeline is the process of collecting raw data, including what data are acquired, how, and why [35]. Data acquisition brings data that have been created by a source outside some organization into the organization, for production use. This means that a number of activities, supported by tools, must be carried out, namely identifying, sourcing, understanding, assessing, and ingesting raw data. Instead, data creation is the process that samples signals that measure real world physical conditions and converts the resulting samples into digital numeric values. The data acquisition or creation is typically interdisciplinary account: depending on the specific material, we have different tools and workflows, which all converge to produce raw data. It is of paramount importance that the specific competences ensure the best quality of data, which are certainly processed afterwards. It is here that data start their life cycle; algorithms and instruments must be annotated for subsequent reference and for processing and interpretation purposes. Such technical metadata are stored with raw data for possible revisions, in case of acquisition errors.

For example, the Virtual Electronic Poem (VEP) project realised the virtual reconstruction of a 1958 event (a case of intangible heritage) which was originally designed by Le Corbusier for Philips pavilion at the Brussels 1958 World’s Fair [29, 30]. The multimedia show consisted of visual effects, a film conceived by Le Corbusier, two music pieces composed by Varèse and Xenakis, respectively. The Poème Electronique was written on scores and each visual score was called minutage. Designed by Le Corbusier and Jean Petit, it was split into two pages; one for the ambiances (with their duration) and one for the other visual effects, updated on a second by second basis. The aural “control” score, of which only 30 seconds are left, indicated the routing of music tapes into loudspeakers with signals to activate loudspeaker groups before sound routing.

The recorded audio consisted on tapes stored in the archive of the Den Haag Royal Conservatory that were digitized as individual

tracks; the film, stored in Philips archived as a VHS tape was digitized too; the visual effects were created from the scratch, given the indications provided by the documentation available. Then, sounds and visuals were spatialized and temporized according to the control score, which assigned sounds to groups of loudspeakers in time (animated sound effect) and visuals to walls of the pavilion. Even the control score, retrieved on a small excerpt on a journal paper was reconstructed together with the pavillon structure (Fig.2) We can notice that audio and video technicians were in close contact with designers and programmers of the virtual exhibition, sharing formats and information about the digitized materials. The sharing of formats was of particular importance here, because of the software functioning.

3.0.2 Data Processing and Modelling. Data Processing and Modelling is the process of creating a conceptual model for the data to be stored in a database, together with the associations between different data objects and the rules (E-R Model and UML are common). The resulting metadata deal with attributes “that describe, provide context, indicate the quality, or document another object (or data) characteristics” [20]. In order to ensure correct use and interpretation of the data by its owners and users, different scientific communities use different sets of metadata standards, such as, e.g., EML (Ecological Metadata Language), FGDC (Federal Geographic Data Committee) standard, and ISO 19115 (International Organization for Standardization Geographic information metadata). Standardized structure and consistent metadata are usually in machine readable extensible markup language (xml) that can be represented in other human readable formats (e.g., html, pdf, etc.) [12]. In this context, metadata sharing of electronic resources, widely used for scientific articles, is achieved through the Digital Object Identifiers (DOI⁶), also popular for scientific primary data [7]. Usually, the basic level of metadata is entrusted to the Dublin

⁶<http://www.doi.org>, (last visited on 10 April 2020)

Ambiance		Sec	Volumés		Ecrans		Sec	Tritons	
Ambiance n°	Durée		Mat	Fem	Notes	Vision	Référence		
50	10"	446				Jeu des 2 enfants	T.274	446	7' 26"
		447						447	7' 27"
		448						448	7' 28"
		449						449	7' 29"
		450				Jeu des 2 enfants	T.274	450	7' 30"
		451						451	7' 31"
		452						452	7' 32"
		453				La main ouverte	T.275	453	7' 33"
		454						454	7' 34"
		455						455	7' 35"

Figure 2: The reconstruction of the minutage. A number of ambiances were deleted or simplified with respect to the original setting [29].

Core Metadata Element Set (DCMES⁷), which is a widely accepted standard, which covers all general requirements and information to describe a dataset. Basic information is typically Title, Creator, Subject (e.g., the name of the project), Description (abstract), Publisher, actual Contributor of some record, Date, Type of dataset, Identifier (e.g., DOI), primary Language.

Particularly related to cultural heritage is the comprehensive Europeana Data Model (EDM⁸), an ontology that enables cultural heritage institutions to structure collection data, so that they can be used by the data aggregator Europeana project. Europeana features over 58 million cultural heritage items from around 4,000 institutions across Europe⁹. This model is designed to provide a generic data model for the core categories (e.g., object type, media type, date, place) and be compatible with specific data models that are used by individual libraries, archives and museums. The generic data model consists of some existing top-level ontologies, such as OAI-ORE (Open Archives Initiative Object Reuse & Exchange¹⁰), Dublin Core¹¹, SKOS (Simple Knowledge Organization System¹²), and CIDOC-CRM¹³. Also EDM introduces its own concepts and properties to organise the descriptive metadata and digital representations of the cultural heritage object (CHO). The management of metadata also involves controlled vocabularies, authority files, and thesauri (such as, e.g., Geonames). The goal is to support effective exchange of knowledge and interoperability and keep a consistent global database including the several collections involved. Fig. 3 shows a fictitious example of the creation of the conceptual model of a cultural heritage object. Here this specific example, the data properties of "Mona Lisa" is shown. In order to be able to unambiguously refer to a set of Web resources, a new Resource is introduced that stands for the collection of other Resources. This new resource, named an Aggregation, has a URI and a machine-readable representation that provides details about the Aggregation. It lists the Aggregated Resource here as the OAI-ORE provides ore:aggregation relationship. This URI (*ex1:aggregationID1*) can be seen through the related web source (*http://www.../image/joconde/XXX.jpg*) which can be tagged through *edm:hasView*. Here, also, the resource is

linked through the provider relationship (*edm:dataProvider*) to the "Direction des musées de France", written in French (*@fr*). Dublin Core terms used for the authorship of the creator (*dc:creator*) which leads to a person specified as *E21* according to CIDOC-CRM. Classification into this that can be expressed with SKOS (*skos:prefLabel*). In particular, the relationship with CIDOC-CRM does not correspond to any real case, since the two models, which are informally claimed to be compliant, did not reach an effective integration yet, for the exploration of researchers, administrators and the public.

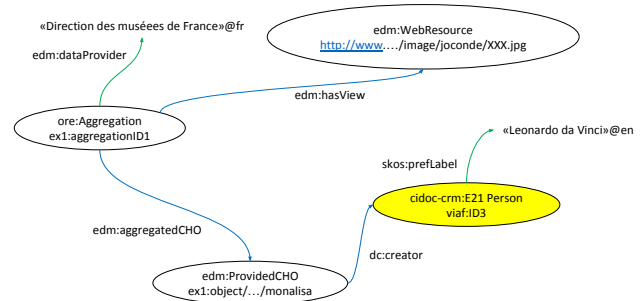


Figure 3: An example of Cultural Heritage domain and descriptive metadata on EDM.

The personalization of metadata here faces at least two challenges. The first is the individual contribution of the institutions that must be preserved while implementing interoperability; the second is the convergence of individual scientific communities on a single cultural heritage object. The latter has been somewhat neglected. The challenge is that, though some individual physical object (e.g., the Kylix) is shared by several scientists, the metadata set is different for each discipline; however, the trans-disciplinary account of some knowledge issue requires the access on request of other metadata, in related disciplinary accounts. So, it would be of paramount importance to provide selected access to metadata and to provide coordination between the several disciplines through some smart interface. This is not achieved in current digital curation projects [8]. In fact, some studies [16, 21] underline researchers' recalcitrance in recording even basic metadata which benefits not only a controlled vocabulary, but also information on related objects, on intellectual property, on user information, on versioning, on integrity checks, and on preservation. In addition, data processing and modelling can show multiple views on an object, including

⁷ <http://www.dublincore.org/>, (last visited on 10 April 2020)
⁸ <https://pro.europeana.eu/page/edm-documentation>, (last visited on 9 April 2020)
⁹ <https://pro.europeana.eu/about-us/mission> (last visited on 1 April 2020)
¹⁰ <http://www.openarchives.org/ore/> (last visited on 1 April 2020)
¹¹ <https://www.dublincore.org/specifications/dublin-core/demi-terms/> (last visited on 10 April 2020)
¹² <https://www.w3.org/TR/skos-reference/> (last visited on 10 April 2020)
¹³ www.cidoc-crm.org/docs/frbr_oo/frbr_docs/FRBROo_V2.4.pdf (last visited on 2 April 2020)

information on the physical and digitised representation – distinct, yet together.

3.0.3 Data Interpretation. Data interpretation is the process of making sense of numerical data that has been collected, analyzed, and presented. As a part of the digital curation process, interpretation allows easy reflection and dialogue. It also allows all members of the project access to a holistic overview of the data, so that interpretations about individual item, item sets, or higher-order categories are possible. For example, in an archaeological excavation, an important item is the stratigraphic unit, some layer in a trench that is distinguishable from the others, and characterized by some formation process. A stratigraphic unit may contain many archaeological findings that potentially feature the same dating. Both these categories, archaeological finding and stratigraphic unit, must be represented, interpreted, related to each other.

In contemporary digitally-enabled excavations, such as Çatalhöyük, researchers engage in a range of long-term activities that encompass the capture, description, annotation, classification, interpretation, knowledge enrichment and dissemination of documents. The personalised databases were created by diverse CH specialists who work in different sites and fields during the different time periods alongside with the different modalities of media [32].

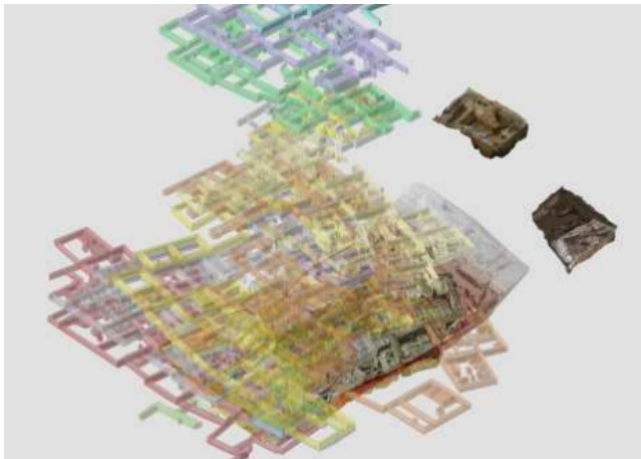


Figure 4: 3D GIS-based visualizations generated by IBM for Çatalhöyük Project[26].

In particular, interpretation addresses digital sheet documentation, diaries and audio memos, photography, reflexive video recording, sketching and drawing, stratigraphic sequence matrix drawings, 3D models, GIS-based graphic representations and hypertextual, diagrammatic and narrative interpretations of particular aspects of the archaeological record [40]. Please, consider that all such files were built by many different people at different times. Possibly, the high-resolution geo-referenced models are finally imported and visualized in spatial relation with files created during the investigation. For example, the 3D GIS models in Fig. 4 show the continuity of occupation sequences at Çatalhöyük, based on data collected in the 1950-60s up to the most recent excavations. Different colours represent different phases of occupation; the reconstructed buildings performed in the field and allowed combining - in the same

virtual space - data come from different analyses [17]. Here, singular perspectives concern the interpretations, expressed as narratives and models, as required by different scientific communities (such as, archaeologists, historians, chemists) and general audiences. The record of the archaeological site of Çatalhöyük was constructed dynamically as the excavation progressed more than twenty years, through a combination of single-context database recording of a multiple cultural heritage specialists. Therefore the practice of digital curation through the years was constantly subjected to adjustment the application of new technologies and the modification of interpretations from multiple viewpoints of interdisciplinary teams [5].

3.0.4 Data Curation and Preservation. Data Curation and Preservation records all the issues that emerge from the data creation to the conceptual model, namely the three phases previously described. It is the active management of data, in the effort to reduce threats to their long-term value and mitigating digital obsolescence [14]. This phase involves a serious reflection on the digital materials and builds up on open and sustainable formats, semantic relations between artefacts and their constituent parts, as well as aspects regarding authorisation, persistent identification, data curation and long-term archiving. In terms of curation and preservation, the data scientists also undertake all the management and administrative actions. For this purpose, data from a relational database are usually converted into RDFS/OWL graph data, allowing the representation, retrieval and traversal of data according to semantic relationships between things and people [13].

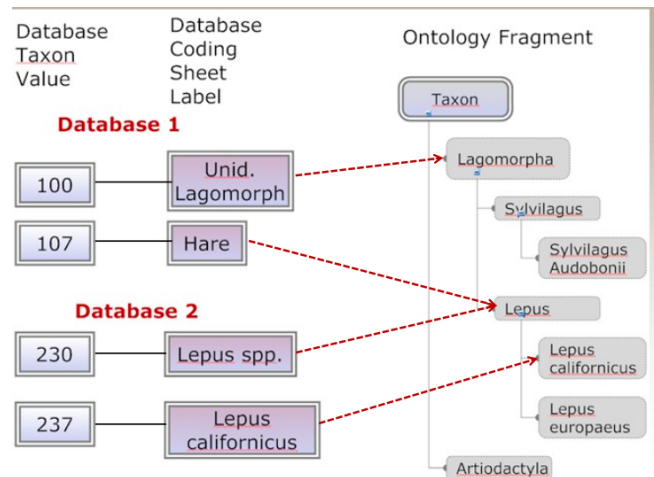


Figure 5: Data integration process of tDar with a fragment of ontology¹⁴

For example, the Digital Archaeological Record (tDAR)¹⁵ is an international digital repository for the digital records of archaeological investigations. tDAR's use, development, and maintenance are governed by Digital Antiquity, an organization dedicated to

¹⁴Image Source: <https://slideplayer.com/slide/7914481/> (last visited on 13 April 2020)

¹⁵<https://www.digitalantiquity.org/wp-uploads/2011/07/20110930-Building-tDAR.pdf> (last visited on 12 April 2020)

ensuring the long-term preservation of irreplaceable archaeological data in order to broaden the access to these data. The goal of this research was to develop tools for synthetic and comparative research based on novel, on-the-fly, ontology-based data integration to be deployed and tested in the context of the prototype infrastructure. In the semantic encoding of tDar, as can be seen in Fig. 5, ontologies are ordinarily hierarchical (tree-like) and represent an arbitrary number of levels of class-subclass relationships. tDar uses query-driven, ad-hoc data integration based on relevant taxonomies that reconciles the semantic demands of a query with the semantic content of the available datasets. There are many perspectives on the creation of the databases through these taxonomies. The collaboration issue of the project, related to the personalised data for diverse agents, such as the data curator, the curation professional, or the end user (including algorithms), requires the representation of different granularities for taxonomies, data sets, data collection sheets, and interdisciplinary repositories.

3.0.5 Data Documentation and Archiving. The Data Documentation and Archiving process manages the metadata or information about some data product (e.g., data table, database) that enables one to understand and use the data. For example, a database can be classified by the type of content included in it (e.g., bibliographic, statistical, document-text) or by its application area (e.g., Biological, Geological, etc). In general, such information may include the scientific context underlying the data as well as who collected the data, why the data were collected, and where, when, and how the data were collected.

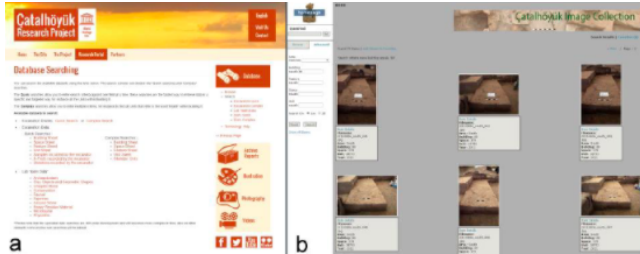


Figure 6: View of (a) the open-access online data curation platform Çatalhöyük Database and (b) the Çatalhöyük image Collection Database[27].

Digital curation activities not only allow the researchers to use appropriate standards and tools but also help the development and evolution of the database. For instance, the Çatalhöyük Database (as can be seen in Fig. 6a) and the Çatalhöyük Image Collection Database¹⁶ are two fundamental assets (see Fig. 6b) in the documentation of Çatalhöyük excavation site. These custom platforms are searchable data management systems that archive images and archaeological datasets (e.g., excavation areas, buildings, spaces, and numerous categories of finds). During every field season, new data are uploaded into these systems locally by the excavators and specialists doing research on-site. In the weeks following each field-work season, all the new data stored locally on the Çatalhöyük Database and on the Çatalhöyük Image Collection Database are

¹⁶<http://www.catalhoyuk.com/research/database> (last visited on 4 April 2020)

made available to the public via a web-based SQL (Structured Query Language) public repository linked to the Çatalhöyük's official website. For example, unit sheets and plans, once completed and checked by the area supervisor, are immediately digitized into the site database and become available to all on-site researchers. Since the project was created by diverse CH specialists who work in different sites and fields during the different time periods alongside with the different modalities of media [32], it is important to note that database entries are checked in order to avoid data-entry errors; the unit plans are digitized by GIS(Geographic Information System) specialists on-site that also integrate them into the larger site map. Altogether, open access to the database, routines that encourage interaction between team members as well as alternate mediums of data narration and recording, aim to inject more voices, as well as self-critical awareness into a relatively rigid recording system. Additionally, a number of these distinct practices have been implemented with the specific objective of softening the strict recording system, as well as, creating a structure that allows for the free dissemination of information-including site data-to research teams working at Çatalhöyük, as well as external researchers and the general public.

3.0.6 Data Publication and Dissemination. Data dissemination is the distribution or transmitting of statistical, or other, data to end-users in order to make data available from one or multiple sources. There are many ways organisations can release data to the public, i.e. electronic format or paper publications such as PDF files based on aggregated data as well as the exhibitions and websites of the collection of the cultural heritage organisations. As digital professionals lead the strategies of cultural heritage institutions through this time of physical distancing and online guided tours during the Covid-19 outbreak, increasing digitization and the emergence of new data sharing practices are also growing rapidly. The curation of data is extremely important in relation to the consequences of the interfaces that curate cultural heritage assets, especially in relation to these new changes.

For example, the Content Management System Omeka provides several user interface metaphors that are well suited to presenting and interacting with rich media content. A collection is a full archive of material, related by event. As can be seen in Fig.7, an exhibit is a selected set of materials from a collection, related by theme, topic, or other curatorial decision. With the use of Omeka software, many cultural institutions have been disseminating their collections and curated content over the years¹⁷. Participatory practices and tools changed the notion of archiving and practices in cultural heritage institutions due to the character of the audience involved and the established structures between and within institutions through social media, linked open data, and other participatory practices. However, Omeka platform opens up new opportunities and practices yet to be explored in order to disseminate the collection and curate content for the virtual exhibitions.

¹⁷Image source: <https://bearchaeo.di.unito.it/omeka-s/s/bearchaeo-resources-site/page/welcome> (last visited on 20 May 2020)

¹⁸<https://omeka.org/classic/directory/> (last visited on 5 April 2020)

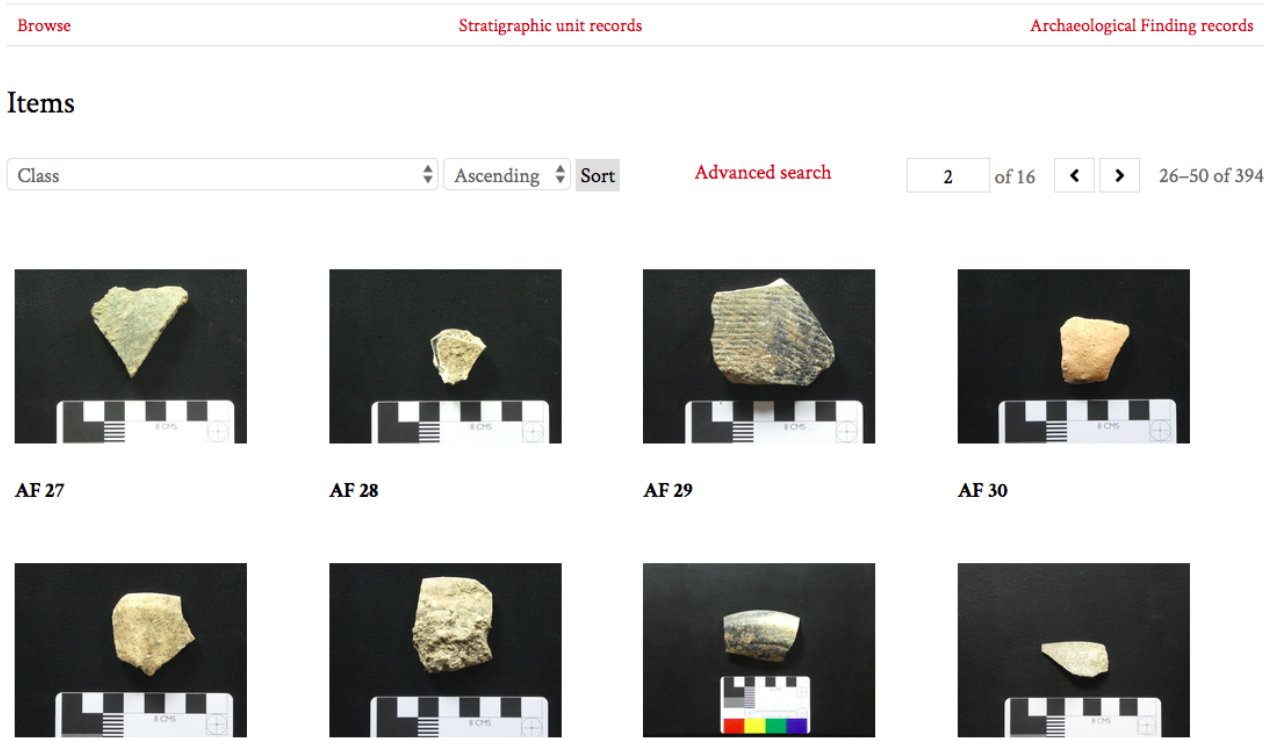


Figure 7: View of an example of Omeka-built website's interface ¹⁷

4 DISCUSSION AND CONCLUSIONS

We examined the contribution of digital curation in the cultural heritage field through reviewing the tasks and some tools implemented throughout the digital curation workflow. The goal has been to see how scholars have tried to systematize the digital curation experience and provide a basis for sharing the best practices. The tasks and tools analyzed for the purposes of this paper reflected the digital curation activities in the literature. Although we intentionally examined a sample of cultural heritage projects that represent a diversity along many relevant aspects of cultural heritage field, the digital curation is still relatively new in the field.

Although there is some confusion amongst the practitioners of the digital curation process [8], we realised that in order to enrich the collaboration between the disciplinary specialists of cultural heritage, digital curation should encompass not only the management of the information from its creation to its possible discard but also its long-term accessibility and preservation, reusability, authenticity and integrity [41]. In traversing our schema (Fig. 1), it is important to 1) conceptualise the entire process, 2) decide about the tasks, 3) define the roles and responsibilities, and 4) provide training of related tools and software applications for collaborative work.

The intent of this research has been to provide insights into the collaborative nature of digital curation practice, and to provide

evidence because the foundations for future empirical research include personalization issues. Metadata would support nearly all of the steps in the digital curation lifecycle [28, 37]. In fact, without regular, sustained processing of the data of the cultural heritage assets and its reflections on metadata, digital curation activities are only put into practice intermittently which eventually creates a gap between practices and multiple actors.

Going through the major tasks of digital curation, we also detected some limitations of the formal representations. These insights are instructive for further research. In the near future, we are going to address a deeper analysis of the currently used frameworks to propose a unified framework to encompass all the tasks and activities. The digital curation framework will be employed in an ongoing EU project named Be-Archaeo ¹⁹. A preliminary achievement of this research has been the design and implementation of a semantic database for the encoding and storing of the digital data concerning an archaeological excavation and addressing the metadata belonging to the several disciplines though concerning the same cultural heritage object (a stratigraphic unit or an archaeological finding)²⁰.

¹⁹<https://www.bearchaeo.com/> (last visited on 7 April 2020).

²⁰<https://bearchaeo.unito.it/omeka-s> (last visited on 15 April 2020).

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 754511. This research is also related to a project called "Be-Archaeo" which is funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie, Grant Agreement No 823826.

REFERENCES

- [1] Daisy Abbott. 2008. *What is Digital Curation?* Retrieved March 2, 2020 from <http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/what-digital-curation>
- [2] Daniel Aliaga, Elisa Bertino, and Stefano Valtolina. 2011. DECHO-A framework for the digital exploration of cultural heritage objects. *JOCCH* 3 (03 2011), 12. <https://doi.org/10.1145/1921614.1921619>
- [3] Suzie Allard. 2014. *Evaluating a complex project : DataOne*. Purdue University Press, West Lafayette, IN, 149–165.
- [4] Janna Quitney Anderson, Lee Rainie, Martha M. Batorski, and D Hadden. 2010. The impact of the Internet on institutions in the future.
- [5] Asa Berggren, Nicolò Dell'Unto, Maurizio Forte, Scott Haddow, Ian Hodder, Justine Issavi, Nicola Lercari, Camilla Mazzucato, Allison Mickel, and James Taylor. 2015. Revisiting reflexive archaeology at Catalhöyük: Integrating digital and 3D technologies at the trowel's edge. *Antiquity* 89 (04 2015), 433–448. <https://doi.org/10.15184/aqy.2014.43>
- [6] William Block, Thomas Hartmann, Bryan Fitzpatrick, Dan Gillman, Jay Greenfield, Arofan Gregory, Marcel Hebing, Larry Hoyle, Chuck Humphrey, Jon Johnson, Jenny Linnerud, Brigitte Mathiak, Steven McEachern, Barry Radler, Ørnulf Risnes, Dan Smith, Wendy Thomas, Joachim Wackerow, Dennis Wegener, and Wolfgang Zenk-Möltgen. 2012. Developing a Model-Driven DDI Specification. *DDI Working Paper Series* (01 2012).
- [7] Jan Brase and Adam Farquhar. 2011. Access to Research Data. *D-Lib Magazine* 17 (01 2011). <https://doi.org/10.1045/january2011-brase>
- [8] Christopher A Lee Andrew Rabkin Yinglong Zhang Katherine Skinner Sam Meister Colin Post, Alexandra Chassanoff. 2019. Digital Curation at Work: Modeling Workflows for Digital Archival Materials. In *Proceedings of ACM/IEEE Joint Conference on Digital Libraries*. ACM Press, 39–48.
- [9] Maria Collins. 2009. Evolving Workflows: Knowing when to Hold'em, Knowing when to Fold'em. *The Serials Librarian* 57, 3 (2009), 261–271. <https://doi.org/10.1080/03615260902877050>
- [10] Panos Constantopoulos, Costis Dallas, Ion Androutsopoulos, Stavros Angelis, Antonios Deligiannakis, Dimitris Gavrilis, Yannis Kotidis, and Christos Papatheodorou. 2009. DCC&U: An Extended Digital Curation Lifecycle Model. *International Journal of Digital Curation* 4 (06 2009). <https://doi.org/10.2218/ijdc.v4i1.76>
- [11] Kevin Crowston and Jian Qin. 2011. A Capability Maturity Model for Scientific Data Management: Evidence from the Literature. *Proceedings of the American Society for Information Science and Technology* 48 (01 2011). <https://doi.org/10.1002/meet.2011.14504801036>
- [12] Constanze Curdt, Dirk Hoffmeister, Guido Waldhoff, C. Jekel, and G. Bareth. 2012. Development of a metadata management system for an interdisciplinary research project. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* I-4, 7–12. <https://doi.org/10.5194/isprsannals-I-4-7-2012>
- [13] Costis Dallas. 2015. Curating Archaeological Knowledge in the Digital Continuum: from Practice to Infrastructure. *Open Archaeology* 1 (01 2015). <https://doi.org/10.1515/opar-2015-0011>
- [14] Martin Doerr, Christian-Emil Ore, and Stephen Stead. 2007. The CIDOC Conceptual Reference Model - A New Standard for Knowledge Sharing. 51–56. <https://doi.org/10.13140/2.1.1420.6400>
- [15] Beverly Dowdy and Rosalyn Raeford. 2014. Electronic Resources Workflow: Design, Analysis and Technologies for an Overdue Solution. *Serials Review* 40, 3 (2014), 175–187. <https://doi.org/10.1080/00987913.2014.950040>
- [16] Paul N. Edwards, Matthew S. Mayernik, Archer L. Batcheller, Geoffrey C. Bowker, and Christine L. Borgman. 2011. Science friction: Data, metadata, and collaboration. *Social Studies of Science* 41, 5 (2011), 667–690. <https://doi.org/10.1177/0306312711413314>
- [17] Maurizio Forte and Nicola Lercari. 2015. Digital Archaeology at Çatalhöyük: New Inferential Methods for the Interpretation of Neolithic Buildings.
- [18] Eben Gay, Katarina Galor, D.B. Cooper, Andrew Willis, Benjamin B. Kimia, Suman Karumuri, Gabriel Taubin, Will Doutré, Donald Sanders, and Shubao Liu. 2010. REVEAL Intermediate Report. In *Proceedings of CVPR Workshop on Applications of Computer Vision in Archaeology (ACVA)*.
- [19] Anna Gold. 2010. Data Curation and Libraries: Short-Term Developments, Long-Term Prospects. *Office of the Dean (Library)* (01 2010).
- [20] Jane Greenberg. 2017. Big Metadata, Smart Metadata, and Metadata Capital: Toward Greater Synergy Between Data Science and Metadata. *Journal of Data and Information Science* 2 (2017), 19–36.
- [21] Ross Harvey. 2010. *Digital Curation: A How to Do It Manual*. Neal Schuman, New York, NY.
- [22] Sarah Higgins. 2008. The DCC curation lifecycle model. *International Journal of Digital Curation* 3, 453. <https://doi.org/10.1145/1378889.1378998>
- [23] Jukka Jokilehto. 2017. *Questions of authenticity concerning different cultures and preservation*. 27–36. https://doi.org/10.30618/978-989-658-467-2_02
- [24] Ray Joyce. 2009. Shards digital curation and the education of information professionals. *Museum Management and Curatorship* 24, 4 (2009), 357–368. <https://doi.org/10.1080/09647770903314720>
- [25] Christopher A. Lee and Helen Tibbo. 2011. Where's the archivist in digital curation? exploring the possibilities through a matrix of knowledge and skills. *Archivaria* 72 (01 2011), 123–168.
- [26] Nicola Lercari, Maurizio Forte, Nicolò Dell'Unto, and Kristina Jonsson. 2015. *Interpretation Process at Çatalhöyük using 3D*.
- [27] Nicola Lercari, Emmanuel Shiferaw, Maurizio Forte, and Regis Kopper. 2017. Immersive Visualization and Curation of Archaeological Heritage Data: Çatalhöyük and the DigIT App. *Journal of Archaeological Method and Theory* (06 2017). <https://doi.org/10.1007/s10816-017-9340-4>
- [28] Melissa Levine. 2014. *Copyright, Open Data, and the Availability-Usability Gap: Challenges, Opportunities, and Approaches for Libraries*. Purdue University Press, 129–148.
- [29] Vincenzo Lombardo, Andrea Valle, John Fitch, Kees Tazelaar, Stefan Weinzierl, and Wojciech Borczyk. 2009. A Virtual-Reality Reconstruction of Poème Électronique Based on Philological Research. *Computer Music Journal* 33 (06 2009), 24–47. <https://doi.org/10.1162/comj.2009.33.2.24>
- [30] Vincenzo Lombardo, Andrea Valle, Fabrizio Nunnari, Francesco Giordana, and Andrea Arghinetti. 2006. Archeology of Multimedia. In *Proceedings of the 14th ACM International Conference on Multimedia* (Santa Barbara, CA, USA) (MM '06). Association for Computing Machinery, New York, NY, USA, 269–278. <https://doi.org/10.1145/1180639.1180706>
- [31] Liz Lyon. 2012. The Informatics Transform: Re-Engineering Libraries for the Data Decade. *International Journal of Digital Curation* 7 (03 2012), 126–138. <https://doi.org/10.2218/ijdc.v7i1.220>
- [32] Michael Ashley López, Ruth Tringham, and Cinzia Perlingieri. 2011. Last House on the Hill: Digitally Remediating Data and Media for Preservation and Access. *Journal on Computing and Cultural Heritage (JOCCH)* 4, 109–116. <https://doi.org/10.1145/2050096.2050098>
- [33] Manjula Patel, Simon Coles, David Giaretta, Stephen Rankin, and Brain McIlwrath. 2009. The role of OAIIS representation information in the digital curation of crystallography data. <https://doi.org/10.1109/e-Science.2009.27> IEEE eScience 2009 ; Conference date: 09-12-2009 Through 11-12-2009.
- [34] Alex Poole. 2016. The Conceptual Landscape of Digital Curation. *Journal of Documentation* 72 (09 2016). <https://doi.org/10.1108/JD-10-2015-0123>
- [35] Line Pouchard. 2015. Revisiting the Data Lifecycle with Big Data Curation. *International Journal of Digital Curation* 10 (06 2015). <https://doi.org/10.2218/ijdc.v10i2.342>
- [36] Graham Pryor. 2012. *Managing Research Data*. Facet, London, Chapter Why Manage Research Data?, 1–16.
- [37] Jen Riley. 2014. *Research Data Management: Practical Strategies for Information Professionals*. Purdue University Press, West Lafayette, IN, 149–165.
- [38] Chuan Shi, Yitong Li, Jiawei Zhang, Yizhou Sun, and Philip S. Yu. 2016. A survey of heterogeneous information network analysis. *IEEE Transactions on Knowledge and Data Engineering* 29, 1 (2016), 17–37.
- [39] Helen Tibbo and Carolyn Hank. 2015. Digital Data Curation Essentials for Data Scientists and Data Curators and Librarians. 293–294. <https://doi.org/10.1145/2756406.2756928>
- [40] Burcu Tung. 2013. *Excavations in the North Area*. Retrieved March 30, 2020 from http://www.catalhoyuk.com/sites/default/files/media/pdf/Archive_Report_2013.pdf
- [41] Elizabeth Yakel, Paul Conway, Margaret Hedstrom, and David Wallace. 2011. Digital Curation for Digital Natives. *Journal of Education for Library and Information Science* 52 (01 2011), 23.