06

INFORMATION AND DOCUMENTATION PERSPECTIVE ON INTANGIBLE CULTURAL HERITAGE IN THE CONTEXT OF DIGITAL HUMANITIES RESEARCH

DOI: 10.17234/SEC.35.8
Original scientific paper
Received:
29th August 2023
Accepted:
25th October 2023

GORAN ZLODI
Sveučilište u Zagrebu, Filozofski fakultet /
University of Zagreb,
Faculty of Humanities and Social Sciences
Ivana Lučića 3,
HR–10000 Zagreb, Hrvatska / Croatia
gzlodi@unizg.ffzg.hr
orcid.org/0000-0003-0860-4469

This paper presents the critical role of digital humanities methods in safeguarding, research, and revitalisation of intangible cultural heritage. The study explores digitisation methods through a documentation and data management perspective, advocating for data standards that facilitate data interoperability at structure, content, value, and encoding levels. This paper critically analyses existing data standards and conceptual reference models, considering their potential and limitations for documenting and representing intangible cultural heritage. The typology of entities and processes related to intangible cultural heritage is presented and elaborated. Mapping identified entities and processes to individual classes in the CIDOC-CRM model is discussed to support the publication of data and semantic interoperability in the Semantic Web context.

Keywords: *intangible cultural heritage, digital humanities, data standards, conceptual reference models, linked open data, documentation, information*

INTRODUCTION

The importance of intangible cultural heritage is indisputable, both for local communities (it fosters social cohesion and strengthens bonds within communities) and involved individuals (helps community members understand their history and culture, express and preserve their cultural identities), and for all of humanity (it promotes diversity, intercultural dialogue, understanding and mutual respect). As intangible cultural heritage is often transmitted orally or through diverse formal and informal practices of its living

bearers, it is fragile and vulnerable and needs to be safeguarded so it can be passed on to future generations.

The application of information and communication technologies (ICT) and digital humanities methods on intangible cultural heritage documentation, digitisation, archiving, dissemination, and communication can play an essential role in safeguarding, research, and revitalisation. Therefore, digital humanities methods relevant to intangible cultural heritage are presented here through the prism of documentation and data management workflow. However, to enable research-driven cultural heritage datasets to become accessible online as findable, accessible, interoperable, and reusable (FAIR principles), data management plans, which are currently high-level descriptive, are not enough. Data standards that can, more granularly, at a low level, describe data structure, content, value, and encoding are needed.

Nevertheless, existing digitisation and data standards for documenting cultural heritage focus on tangible assets. Hence, they need to be further developed to adequately cover intangible aspects of heritage and culture. Therefore, this paper intends to critically analyse existing data standards and conceptual reference models, considering their potential and limitations for documenting and representing intangible cultural heritage.

Firstly, the theoretical approach by Michael Buckland will be presented to provide insight into the fundamental concepts of information sciences like knowledge, information, and material manifestations such as data and documents (Buckland 1991). Since Buckland's theory addresses the highlighted concepts while focusing on their tangible and intangible facets, it is highly relevant for exploring and conceptualising data approaches to safeguarding intangible cultural heritage.

Analysis of existing relevant data standards will be presented, and the role of reference models as conceptual layers that can enable the integration and sharing of various data from heterogeneous sources will be discussed.

After analysis of existing relevant standards and conceptual models, the working typology of entities and processes related to intangible cultural heritage will be presented and further discussed with regards to abilities and challenges for their recording, archiving, processing, and dissemination in a global information environment. At the same time, mapping identified entities and processes to individual classes in the CIDOC-CRM model will be proposed.

INTANGIBLE CULTURAL HERITAGE

Some works discuss the concept and terminology related to *intangible cultural heritage* (cf. Nikočević et al. 2012) or question UNESCO's initiatives (cf. Mountcastle 2010), and such critical questioning will undoubtedly be necessary for further interdisciplinary research on modelling intangible cultural heritage. Nonetheless, this paper will not further delve into the appropriateness of the term intangible cultural heritage itself; instead, it will use UNESCO's 2003 Convention for the Safeguarding of the Intangible Cultural Heritage (UNESCO 2003) as a starting reference for intangible cultural heritage. The definitions provided below will serve as the foundation for this discussion.

According to UNESCO's 2003 Convention definition, intangible cultural heritage includes:

> "the practices, representations, expressions, knowledge, skills – as well as the instruments, objects, artefacts and cultural spaces associated therewith – that communities, groups and, in some cases, individuals recognise as part of their cultural heritage." (UNESCO 2003:2)

Intangible cultural heritage, according to UNESCO convention, is manifested in the following domains:

(a) oral traditions and expressions, including language as a vehicle of the intangible cultural heritage;
(b) performing arts;
(c) social practices, rituals and festive events;
(d) knowledge and practices concerning nature and the universe;
(e) traditional craftsmanship. (UNESCO 2003:2)

When we consider the following UNESCO definition of *safeguarding*, it is evident at first glance that digital methods could help in the implementation of the mentioned measures:

> "'Safeguarding' means measures aimed at ensuring the viability of the intangible cultural heritage, including the identification, documentation, research, preservation, protection, promotion, enhancement, and transmission, particularly through formal and nonformal education, as well as the revitalisation of the various aspects of such heritage." (UNESCO 2003:3)

Digital humanities, an interdisciplinary field that applies computational methods to the study of the humanities, offer valuable tools, approaches and methods that can significantly aid in research, preservation, and safeguarding of intangible cultural heritage. By applying the methods of digital humanities, large amounts of data are created: different digital representations of material and immaterial culture in the form of media files, metadata created for different purposes (descriptive, structural, technical, administrative) and other research data that are created in the processes of transformation, analysis, and visualisation.

Therefore, we need data management and data management plans (DMP) to ensure this vast amount of digital information is organised, preserved, easily retrievable, useful, and accessible for current and future researchers, scholars, and the broader community. Nonetheless, as is evident from the definition of data management plans: "DMPs are generally short, *high-level* descriptive plans that prescribe the data to be generated by a research project, how that data will be stored (securely, as required), who will have access, what documentation and metadata will be created with the data, and preservation intentions if the data are to be preserved long-term" (Burnette et al. 2016:2), DPMs are high-level descriptive plans and, therefore, insufficient to enable semantics and data interoperability in a wider, global information environment. It is necessary to apply several standards that will define the application of data value and data encoding standards at a low level, as well as mapping to conceptual models to define the semantics of data.

Such an upgraded data management approach aims to produce sustainable data of high value and long-term validity. It needs to be included in various research processes, from resources and data collection, documenting and curating resources, to analyse and explore data quantitatively or qualitatively, and integrate it into a global semantic network – a knowledge graph based on conceptual reference models.

Workflow of data management activities

Fafalios et al. (2023) propose a workflow model that is provenance-aware (provenance information is intended to be provided at the data element level, which means that each individual piece of data has its source as a reference!), highly recursive and focuses on semantic interoperability, and identify the following common data management activities in archival research:

- digitisation/transcription of archival documents (scanning of documents,

text recognition, manual transcription)

- documentation/metadata recording (what is the origin of a document, what is the document about, who makes the transcription, etc.)
- data curation / preparing the data for statistical analysis (correction or normalisation of data values, instance matching, term alignment, etc.)
- data integration under a common representation language (ontology-based modelling, creation of mappings, data transformation)
- data publication (e.g. as Linked Data)
- data analysis and exploration (qualitative and/or quantitative analysis, query building, data visualisation, etc.)

In this article, the above workflow will be adapted and supplemented (e.g. digitisation will be expanded with different recording approaches, and data publication will include media dissemination) to use it to consider the key digital humanities approaches in the context of intangible cultural heritage. Workflow will be supplemented with another level: data interpretation, enabling interpretation based on data and sources and supporting creative approaches in living traditions.

1. Digitization

Contemporary digitisation methods facilitate the capturing, scanning, recording, and archiving of intangible cultural heritage elements, such as oral histories, traditional performances, rituals, and more. As Zabulis et al. (2022) advise, it is important that digitisation also includes contextualisation items, literature, and other information-carrying media, e.g., documents, technical drawings, images, and audio recordings.

Recording intangible cultural heritage requires a combination of complementary approaches to adequately capture the full scope, complexity and richness of knowledge, practices, traditions, expressions, and skills. Zabulis et al. (2022) divide tasks concerning the recording into 1) recordings of objects (endurants) and 2) recordings of events (perdurants).

For recordings of *endurants* (materials, tools, machines, products, information carriers, and sites that are related to the intangible cultural heritage practices and their context), the following methods and techniques are dominantly used:

- Photography
- Written documentation
- Drawings

- 3D models of objects and spaces
- Maps and GIS data

Photography plays an essential role in the visual documentation of cultural heritage. It is still the dominant medium in documenting and digitising material heritage, and today, it is being extended by applying 3D models of objects and spaces. Various forms of written documentation remain essential in preserving intangible heritage, from transcribing oral traditions and creating textual accounts of practices, to writing detailed descriptive records of observed events (i.e. rituals and ceremonies). Ethnologist and anthropologists constantly develop their methods of taking detailed field notes.

Maps can visually represent and spatially analyse cultural practices, traditions, and their relationships to specific geographic areas, aiding in preservation, comprehension, and interpretation. Additionally, contemporary GIS technologies are crucial for intangible cultural heritage as they facilitate data-driven documentation and spatial understanding, linking cultures to their geographical origins and ensuring a comprehensive approach to management and preservation.

For recordings of *perdurants* (practitioner voices, body actions, postures and gestures during dance, crafts, and rituals), the following methods and techniques are dominantly used:

- Audio Recordings
- Film and video recordings
- 4D reconstructions
- Motion capture (MoCap)

Audio recording enables direct recording of narratives, experiences, stories, and teachings from community members and practitioners. Sound recording is pivotal in immortalising narratives, experiences, stories, and teachings directly from community members and practitioners. Beyond just words, it is essential for preserving musical expressions, be it vocal and/or instrumental performances. Furthermore, audio technologies serve as invaluable tools for undertaking interviews, ensuring that the nuances of oral histories, with their intonations and emotions, are captured in their entirety for future generations.

Audiovisual recordings are vital for documenting dances, rituals, performances, and craft-making processes. By capturing both auditory and visual aspects over time, they provide a rich and comprehensive representation of the complexities, nuances, and temporal dynamics inherent to these activities.

In the realms of computer graphics and computer vision, 4D reconstruction involves recording processes of capturing the shape and appearance of real objects along a temporal dimension (Mustafa et al. 2016:4661), which can further support the

recording of dynamic aspects of intangible cultural heritage.

Motion capture (MoCap) is the process of recording the movements of objects or people. Hou et al. (2022) list several applications of MoCap that show how those technologies have gained increasing popularity in intangible cultural heritage archiving as a tool for recording movements or manoeuvres of performative cultures: Cypriot folk dances, Southeast Asian traditional dances, martial arts, combative sports, as well as reference gestures of traditional crafts. Motion capture primarily captures the movement of subjects (typically humans) and is widely used to study biomechanics and human movement. In contrast, 4D reconstruction captures an object's evolving shape and appearance over time, typically through multi-view stereo cameras, making it ideal for studying temporal changes in objects or shapes.

Hou et al. (2022:4) highlight that *motion capture* technologies, assisted with machine learning and modelling methods, empower multifaceted understandings of human motion: "...and as such, it lays down the computational basis for tangible visualisation, immersive engagement, and active narratives to communicate knowledge through a magnified sense of embodiment", but at the same time warn us that it is necessary to stress human-centred modelling in combination with explainable artificial intelligence in the computational design for intangible cultural heritage, due to the process of determination used in machine learning, especially *Deep learning models,* which tend to operate as a "black box".

Photography plays a key role in documenting material heritage, and besides, despite its static material nature, it is also a compelling and vital medium for documenting intangible aspects of culture. However, audio-visual media, with audio, visual and time-based recording functionalities, are indispensable for documenting dynamic expressions of intangible culture (oral histories, dances, rituals, and other performances).

2. Documentation

Documentation refers to the procedures of systematically collecting, organising, and recording information or evidence about processes, events, items, or any other subjects of interest. Documentation involves the process of creating records or specifying metadata. It can either be descriptive, in which case it records the attributes of an entity (e.g. an artefact), or procedural, as it outlines the steps to execute a specific task (e.g. transfer of ownership). "Archiving" and "documentation" are both essential processes in preserving and disseminating information, but they serve different functions and purposes. The primary purpose of archiving is the long-term preservation of archival materials or information, ensuring they remain accessible and intact (to maintain the integrity and authenticity of the resource) for future reference.

3. Data curation

Data curation emerges as a precise, dynamic, and evolving practice within the data management sphere and can be further described as "the active and ongoing management of data through its life cycle of interest and usefulness to scholarship, science, and education, which includes appraisal and selection, representation and organisation of these data for access and use over time" (Shreeves and Cragin 2008:2). Furthermore, data curation includes "all the processes needed for principled and controlled data creation, maintenance, and management, together with the capacity to add value to data" (Gopal 2016:72).

This encompasses activities and techniques like annotation, cleaning, validation, linking, and semantical enrichment. Particularly interesting approaches are those related to semantic enrichment, as processes of enhancing digital content (text, datasets) with additional, structured, and both human and computer-meaningful information, often making it more discoverable and interconnected. For these purposes, language technologies are essential, often referred to as *Natural language processing* (NLP) or *Computational linguistics*. They encompass a range of computational tools and methods used to handle, analyse, generate, and understand human language meaningfully. Since language is a critical component of intangible cultural heritage, these technologies can assist in documenting languages (speech recognition, transcription) and semantical enrichment (via *Named Entity Recognition* methods that can help in identifying and classifying names of persons, organisations, and locations, and then by connecting with existing controlled vocabularies).

4. Data integration

Data integration is not a structural nor a syntactical merging of data from different sources; semantic data integration is the process of combining data to provide a unified, coherent understanding based on shared meanings and context. Data mapping and integration should be carried out within the selected conceptual reference model (CIDOC-CRM, LRM or RiC), which serves as a common representation language that provides semantics for each data element involved in the research process. These processes enable data interoperability from heterogeneous sources in the global information context and will be described in more detail in the chapter on Semantic Web standards.

5. Data publication

Data publication that is standardised and interoperable in a global information context is crucial for safeguarding and maintaining intangible culture as a living practice. Publishing data makes it accessible to a broader audience, including researchers, educators, students, and community members. In a research context, data publication allows scholars and researchers from various fields and geographies to access, analyse, and collaborate on studies related to intangible cultural heritage. Moreover, when data related to intangible cultural heritage is published alongside details of its provenance, it provides an essential layer of authenticity and trustworthiness, facilitating the process of its review and verification. In this way, documenting and publishing data can serve as a record of origin, protecting intangible cultural heritage elements from misappropriation or misuse. Before publishing, the data should be adequately curated – prepared in standard formats, annotated and licensed to serve multiple disciplines, fostering cross-disciplinary perspectives and interdisciplinary insights, enhancing collaboration and knowledge discovery and exchange across various domains.

Data publication for intangible cultural heritage preserves valuable cultural knowledge and promotes understanding, appreciation, and continued research in the field, leading to richer and more diverse insights. Through data publication, underrepresented or lesser-known cultural practices can gain visibility and recognition on a global stage. Through publishing data and digital assets, communities can share their culture with others, leading to substantial and more reliable intercultural understanding and exchange.

Online repositories, data archives, and digital libraries ensure that anyone can access and learn about intangible cultural heritage anywhere. It is essential that data is open and connectable (e.g. as Linked Open Data) and that digital assets are available in standard and non-proprietary formats in interoperable methods (e.g. IIIF standards for image interoperability).

As an educational resource, published data can be woven into educational and academic syllabi, enlightening upcoming generations about the richness of varied intangible cultural heritages and their profound importance. As UNESCO emphasises the role of *formal and nonformal education,* the educational resources created must be open (e.g., available as *Open Educational Resources (OER)*) and based on sources. It is undoubtedly essential for all communication and dissemination methods of intangible cultural heritage to be interoperable and based on data, sources, documentation, and archived media.

6. Data analysis and exploration

In addition to the common methods within qualitative, quantitative, and mixed methodologies, digital humanities methods provide advanced text mining methods, pattern recognition and computer vision that can help researchers identify and analyse themes, influences, and evolution of various phenomena of intangible cultural heritage.

Digital tools enable innovative visualisation of tangible and intangible elements, such as mapping oral storytelling routes, tracking the evolution of a dance form, or visualising the relationships between different intangible cultural heritage elements.

Network analysis can show how various intangible cultural heritage elements relate to tangible elements of heritage, societal changes, or other cultural phenomena.

Machine learning (ML) and *artificial intelligence* (AI) can be utilised to analyse vast amounts of intangible cultural heritage data, recognise patterns, predict trends (i.e. which intangible cultural heritage elements are most vulnerable) and suggest measures for their preservation. However, it is crucial to approach AI analyses and outcomes cautiously, addressing ethical concerns, transparency challenges related to "black box" algorithms, and other potential unintended consequences of AI applications in cultural heritage contexts.

7. Data interpretation

Data are "reinterpretable representation of information in a formalised manner suitable for communication, interpretation, or processing" *(Reference Model for an Open Archival Information System (OAIS)* 2012:10) and as such, they are fundamental to different analyses and presentations performed by humans or machines. Data interpretation helps decipher cultural, historical, and societal contexts in which traditions, expressions, and knowledge are deeply rooted, ensuring that the essence of the intangible cultural heritage is understood beyond mere data processing. Without proper interpretation, there is a risk of misrepresenting or oversimplifying complex cultural practices.

Data interpretation is crucial for educational purposes where teaching intangible cultural heritage is about presenting facts and conveying the more profound significance, values, and meanings embedded within.

Combining longstanding tradition and the power of storytelling with digital multimedia elements like text, video, audio, and interactive features provides richer, layered narratives that can help represent the depth and complexity of intangible cultural heritage.

For measures of intangible cultural heritage safeguarding related to communication, interpretation and promotion, digital methods can also play a significant role. For example,

promotion (virtual exhibitions, digital storytelling) and *enhancement* technologies such as Virtual reality (VR), Augmented reality (AR) and Mixed reality (MR) can offer immersive experiences, allowing users to engage with intangible cultural heritage in novel ways, such as virtually participating in simulations of traditional rituals, crafts, or performances.

Intangible Cultural Heritage Data as Humanities Research Data

By leveraging these digital methods, stakeholders can ensure a more holistic, participative, and practical approach to safeguarding the wide variety of intangible cultural heritage. By implementing digital humanities infrastructures and methods, cultural heritage and higher education institutions must offer collaborative platforms where individuals and communities can collaboratively document, explore, discuss, and share their intangible heritage, ensuring that the recording and documentation process is participative and democratic. Interactivity should be one of the key features of websites, applications, or virtual environments where researchers and end-users can explore, engage with, and learn about intangible cultural heritage.

Tasovac, Chambers, and Tóth-Czifra (2020:1) coined the proclamation "Cultural Heritage Data as Humanities Research Data!" to emphasise the potential of cultural heritage collections which, through the dynamic extraction and curation of research-driven digital cultural heritage datasets, must become accessible online as Findable, Accessible, Interoperable and Reusable (FAIR) humanities research datasets.

FAIR principles are a set of essential and valuable guidelines for making research data more findable, accessible, interoperable, and reusable. They were developed by a group of scientists and organisations in 2016 and published in the journal Scientific Data, in the paper "The FAIR Guiding Principles for Scientific Data Management and Stewardship" (Wilkinson et al. 2016:4):

- Findable: Data should be easy to find by humans and machines alike. This means that data should be recorded consistently, according to relevant data standards.
- Accessible: Everyone interested should have access to the data. This means that data should be released under an open license and made available in a format that is easy to use.
- Interoperable: Data should be able to be used with other data. This means that data should be encoded in a standard format and be accompanied by metadata (metadata schema) that describes its structure and content.
- Reusable: Data should be reusable for different purposes. This means

that data should be well-documented and accompanied by clear usage instructions.

There are many approaches and methods to implement the FAIR principles. Some common practices include using clear and consistent naming conventions for data files and metadata; publishing data under an open license, using standard data formats and encodings, and providing detailed documentation about data and metadata *(paradata).* Paradata is, in general, formalised data on methodologies, processes, and quality associated with the production and assembly of statistical data (Karr 2010) or, in the more specific context of data set or survey, data about the process by which the data were collected.

The examples provided in this chapter demonstrate that digital humanities present a variety of approaches, technologies and tools that can facilitate recording and documenting intangible cultural heritage. However, we encounter numerous data interoperability issues when connecting or integrating data from diverse sources. FAIR principles are important reminders and guidelines, but we need systematic approaches to implement them in heterogeneous cultural contexts. Namely, different communities (e.g. GLAM) use different standards; data outcomes of projects are frequently tailored to specific project needs and aligned with limited project scope and objectives. Conceptual reference models such as CIDOC-CRM, LRM and RiC provide promising mechanisms for establishing interoperability; however, their acceptance is relatively slow and insufficiently connected with approaches and standards according to which heritage institutions document cultural heritage today. Therefore, it is necessary to further analyse how we can model intangible cultural heritage to correctly represent all its aspects and streamline its documentation in a standardised way.

INTANGIBLE CULTURAL HERITAGE: AN INFORMATION PERSPECTIVE

When analysing intangible cultural heritage in the context of heritage documentation and knowledge organisation, we recognise concepts such as *information, data, knowledge, events, documents,* and *evidence* - concepts that Michael Buckland theoretically defined in the 1990s in the context of information sciences. Given that Buckland's theory addresses the highlighted concepts with a focus on their tangible and intangible facets, it becomes highly relevant when exploring and conceptualising the safeguarding of intangible cultural heritage within the contemporary context of conceptual modelling knowledge.

Buckland introduced the approach where he identifies three principal uses of the word information: *information-as-knowledge, information-as-process,* and *information-as-thing* (Buckland 1991:351). One segment of intangible cultural heritage, which as an intangible phenomenon includes knowledge, beliefs, skills, etc., can be considered as *information-as-knowledge.* Another segment of intangible cultural heritage is when one or more community members perform some skill or/and communicate knowledge, beliefs, etc. At the same time, the other members can receive, learn, and adopt the common heritage, which can be considered as *information-as-process*. We can also agree that those transmissions of practices are intangible. However, at this point, we can measure, record and document some of them (audio-visual documentation of oral tradition, dances, working processes, etc.).

During this process, some of the physically performed *information-as-knowledge* segments would result in the *information-as-thing,* as a material object (as a product of process), sign, note or document. At this level, we can recognise, on the one hand, material manifestations of intangible cultural heritage such as different artefacts and, on the other hand, documents as results of the documentation process.

In Buckland's perspective on information, he introduces a distinct concept of *information processing,* which is especially important considering its application in various information systems (see Table 1).

	INTANGIBLE	**TANGIBLE**
ENTITY	*information-as-knowledge* (knowledge)	*information-as-thing* (data, document)
PROCESS	*information-as-process* (becoming informed)	*information processing* (data processing)

Table 1: Four aspects of information are based on the three uses of the word 'information' (Buckland 1991:352)

Modelling events, seen as spatio-temporal categories, is essential when representing cultural phenomena. They are intangible per se, so their documentation processing can be challenging. Therefore, the information and documentation processing are managed through documents related to these events.

Buckland identifies three different kinds of evidence of events seen in practice (Buckland 1991:356):

1. Objects, which can be collected or represented objects, can exist as *evidence associated with* the events;

2. Representations of the events themselves: photographs, news reports, memoirs;
3. Events can, to some extent, be *created* or *recreated.*

The concept of evidence is beneficial for further research on the provenance of knowledge, especially regarding intangible cultural heritage. The first kind of evidence corresponds with cultural heritage objects, and the second with documentation. Buckland sets the third type of evidence of events in the context of science, in which experiment is an essential method of verifying scientific hypotheses. Still, this interpretation is comparable to experimental archaeology or experimental anthropology methods, events such as living history, workshops on traditional skills and trades, and similar types of reviving heritage. More precisely, these different forms of manifesting cultural heritage – such as dance, rituals, festivities, oral traditions, and traditional crafts, seen in their original context are the types of events on which the acts of transferring, communicating, and preserving intangible cultural heritage are based.

We can observe Buckland's approach to information on two examples: *Lepoglava lacemaking,* an example of the intangible cultural heritage phenomenon where we have the results of the production process, and *Moreška* (ital. moresca), a combat dance with swords characteristic of the Mediterranean, with related physical objects such as costumes and swords. Today, *moreška* is still performed in Croatia only in Korčula. Lepoglava lacemaking is an example of living history practice and an outstanding example of intangible cultural heritage safeguarding, realised through the ongoing

Information-as-knowledge	**Information-as-process**	**Information-as-thing**		
→	→	→		
ideas, knowledge skills	production, performance, communication	object, sign, data		
		object as product	object required for process	documentation
"Lepoglava lacemaking"	*production of lace*	*Lepoglava lace*	*tool "batići"*	*audio-visual documentation, metadata...*
"Moreška" (traditional dancing)	*dance performance*		*custumes, swords*	*audio-visual documentation, metadata...*

Table 2: Buckland's approach to information on two examples: Lepoglava lacemaking and Moreška combat dance

collaboration of holders, the local community, and relevant institutions. *Lepoglava lacemaking* is undoubtedly the most known and important cultural phenomenon for the local community, especially after it was inscribed on UNESCO's Representative List of the Intangible Cultural Heritage of Humanity in 2009.

Before we continue exploring intangible cultural heritage in the context of documentation and data standards, it is helpful to consider the definition of knowledge concerning data and information provided by Michael H. Zack (1999:45–46):

> "Data represent observations or facts out of context, and therefore not directly meaningful. Information results from placing data within some meaningful context, often in the form of a message. Knowledge is that which we come to believe and value based on the meaningfully organised accumulation of information (messages) through experience, communication or inference."

DOCUMENTATION STANDARDS IN THE CULTURAL HERITAGE DOMAIN

Contemporary documentation of cultural heritage involves predominantly digital methods. The processing of this documentation takes place in information systems, where metadata records, media reproductions and surrogates represent physical and digital phenomena and resources. These representations act as integrating bonds, connecting intangible elements of culture to tangible or digitally accessible and interoperable formats.

Adherence to data and documentation standards is essential to ensure that documentation and associated data are precise, unambiguous, shareable, interoperable, and open. Several typologies of metadata standards are applied in cultural heritage and the broader information environment. In the context of the standardisation of documentation of intangible cultural heritage, guidelines, standards, and norms can be classified into the following categories, all needed for ensuring consistency, comprehension, and interoperability of documentation:

1. Data structure standards

Standardising the data structure entails defining data categories (names of database fields or elements in metadata schemata). This process also establishes the semantics of elements and relationships among various entities and defines characteristics of specific

data categories, like their repeatability and mandatory nature.

Current data structure standards in the cultural heritage domain are oriented to a description of material objects *(SPECTRUM, Categories for the Description of Works of Art (CDWA))* or a description of material objects and visual surrogates *(Visual Resources Association Core Categories). Dublin Core* is not expressive enough to describe complex cultural heritage phenomena despite its flexibility by the 1:1 principle that enables describing and linking various entities. Nowadays, its primary function is to achieve minimum interoperability amongst different systems and communities.

The SPECTRUM collection management documentation standard (McKenna and Patsatzi 2009) is the most widespread internationally. It is oriented towards documenting everyday procedures in museum collection management. One primary procedure is cataloguing museum objects, but not intangible cultural assets. SPECTRUM is a well-established standard for material heritage objects. It includes information groups and units such as the O*bject production information group* and *Description information group* that encompasses processes, methods, techniques, or tools used to fabricate or decorate an object, etc. Nevertheless, this paper looks at how it enables establishing links between heritage objects and non-material aspects of cultural heritage.

The following units of information identified in *Object history and association information* group could help establish relationships among CHO and intangible cultural heritage: *Associated concept* (to refer to specific intangible cultural heritage phenomenon, e.g. "Lepoglava lacemaking"), *Associated event* (to refer to the performance of intangible cultural heritage phenomenon, e.g. "International lace festival in Lepoglava, 1997."), *Associated object* (to refer to an object associated with an object or group of objects, e.g. the specific tool "batići" needed for "Lepoglava lacemaking"), *Associated person, people or organisation* (e.g. Zlata Šufflay, famous lace maker). For all these units of information, terminology control needs to be applied, and consequently, every aspect of intangible cultural heritage can be uniquely identified at the vocabulary level.

2. Data value standards

Data value standards encompass controlled vocabularies, thesauri, and authority files to ensure data consistency and provide disambiguation. They aim to standardise the values inputted into specific data fields, ensure consistency across data records, and support search, retrieval, integration, and reuse.

A thesaurus can be defined based on its function and structure. In terms of function, a thesaurus is a tool for terminology control when translating from the natural language of a document to a more restricted documentation language. In terms of structure, a

thesaurus is a supervised and dynamic dictionary of semantically and generically linked terms that cover a specific knowledge area (ISO 2788).

Several important controlled terminological resources for cultural heritage were created and are maintained by The Getty Vocabulary Program at The J. Paul Getty Trust and are publicly available at https://www.getty.edu/research/tools/vocabularies/. The most important among them is the Art & Architecture Thesaurus (AAT), containing concepts and terms, hierarchical, equivalent, and associative relations between concepts, scope notes and sources for each statement (about terms, relations, notes) on objects, materials, techniques, styles, periods, and other concepts related to art, architecture, and the broader field of cultural heritage. The Getty Vocabulary Program is particularly important in the contemporary data environment of the Semantic Web since all vocabularies are also publicly available and accessible for machines under the *Open Data Commons Attribution License* (ODC-By) 1.0. as Linked Open Data (LOD) on the SPARQL endpoint at vocab.getty.edu.

As a multilingual thesaurus, AAT is the most used vocabulary control resource for museums and broader cultural heritage domains (for example, Europeana uses AAT for data enrichment). However, it is not intended to record local names in different dialects, so it is necessary to establish and develop local thesauri at the national, regional, or other micro levels.

The next step towards the representation of structured controlled vocabulary for the Semantic Web environment is the *Simple Knowledge Organization System* (SKOS). SKOS vocabularies should overcome some technical limitations of thesauri (thesauri are not explicitly intended for the Semantic Web), but also conceptual limitations: "The fundamental problem with traditional thesauri is that its semantic relations have been constructed mainly to help the indexer in finding indexing terms, and understanding the relations needs implicit human knowledge." (Hyvönen 2009:759). Furthermore, Hyvönen states that many cultural thesauri have been transformed into SKOS format; however, although a syntactic transformation into SKOS is beneficial, it is not always enough. He claims that the solution to these fundamental problems is to refine and reorganise a thesaurus's semantic structures into a light-weight ontology, pointing to the Swedish national semantic web infrastructure as a successful example.

3. Data content standards

Standardisation of data content prescribes how data should be entered into specific data categories (in what order, with which syntax) to ensure data consistency. They also include cataloguing rules in the form of different guidelines that help to improve

the quality of data and make data more accessible and reusable so that they can be shared and integrated from diverse sources.

Cataloguing Cultural Objects: a Guide to Describing Cultural Works and Their Images (CCO) is a data content standard, and it provides valuable cataloguing rules for improving the quality of cataloguing procedures for cultural objects, and consequently for improving the quality of data. However, CCO explicitly leaves out intangible cultural heritage from its scope: "Excluded are literary works, music, performing arts, language arts, culinary arts, science, religion, philosophy, and other intangible culture." (Baca et al. 2006:5).

4. Semantic web standards: conceptual reference models and linked open data

In recent years, digital methods have been progressively used to safeguard and share intangible cultural heritage assets in institutional (heritage and high education) or cross-institutional settings, but with very uneven approaches to data management. As stated, archives, libraries, and museums use different data structure standards and cataloguing rules and rarely share controlled vocabularies, even with regards to material and digital resources they usually manage. When it comes to intangible cultural heritage, approaches are even less standardised. Moreover, the application of digital methods in intangible cultural heritage projects and research showed very heterogeneous approaches considering descriptive metadata about intangible cultural heritage phenomena, developed separately within individual research or heritage projects, without standardised data structures, established documentation protocols, cataloguing rules, or data value standards. Furthermore, as Hou et al. (2022:6) point out: "Digital curatorial conventions of intangible cultural heritage are becoming increasingly interdisciplinary, data-driven, and multimodal." Expecting one standard to replace the existing ones in individual professional communities is unrealistic. However, there is an opportunity to apply a conceptual layer, which can, as a reference model, at a higher level of abstraction, enable integration and sharing of data. Such an approach is made possible by conceptual reference models, i.e. formal ontologies, that provide a formal description of the conceptualisation of a specific domain by specifying necessary concepts and relationships among those concepts. They serve as a reference point for various information systems, software agents, and services for exchanging information and knowledge.

A review of the literature showed that two families of conceptual models are the most represented in attempts to model intangible cultural heritage. Therefore, these models were analysed in more detail with regards to their application in Semantic Web's

broader, shared information environment.

The first conceptual model considered is *Functional Requirements for Bibliographic Records* (FRBR), which established essential foundations for modelling different aspects of resources, some of which relate to material properties and others to intellectual or artistic content or how content is expressed. Concerning intangible cultural heritage modelling, the most interesting was the first entity defined in the FRBR model - Work. Work is defined as an abstract entity, "a distinct intellectual or artistic creation", and the work is recognised "…through individual realisations or expressions of the work" (IFLA Study Group on the Functional Requirements for Bibliographic Records 1998:16). Expression is the second entity defined in the model. Expression is how the work is realised – regardless of whether they are material or immaterial (FRBR examples for expression are "alpha-numeric notation, musical notation, spoken word, musical sound, cartographic image, photographic image, sculpture, dance, mime, etc." (ibid.).

The latest step in the development started by FRBR is The IFLA Library Reference Model, which aims to be a high-level conceptual reference model developed within an enhanced entity-relationship modelling framework. The model covers bibliographic data as understood in a broad, general sense (Riva et al. 2017). However, the model is also accepted outside the library community, including for intangible cultural heritage. For example, a case study (Coladangelo 2020) aims to preserve and disseminate cultural heritage information about the North American community folk dance tradition of contra dance through the development of a thesaurus of choreographic terms, *Contra Dance Thesaurus*, and a domain ontology named *Contra ontology.*

The next model to be examined in more detail is the CIDOC Conceptual Reference Model (CIDOC-CRM), a formal ontology designed to facilitate the integration, mediation, and interchange of heterogeneous cultural heritage information (CIDOC CRM Special Interest Group 2022). CIDOC-CRM is an ISO standard, "Information and documentation - A reference ontology for the interchange of cultural heritage information" (ISO 21127:2014), and it is not only acknowledged and embraced in the museums and broader heritage community, but also the evolving field of digital humanities. CIDOC-CRM has a long history and continuity of development and applications in diverse projects. It is event-oriented and aware of temporal and spatial concepts and conceptual entities, all necessary for modelling aspects of intangible cultural heritage.

It is important to note that there is a collaboration between IFLA and the CIDOC-CRM community, which began in 2003 with the work of the International Working Group on FRBR/CIDOC-CRM alignment. This cooperation led to the development of FRBRoo, which is itself an extension of CIDOC-CRM. A well-known example of the application of FRBRoo is the DOREMUS ontology, which was developed as an extension of the FRBRoo model, adapting

it to the broad domain of music, from popular songs to classical music works (Lisena et al. 2018). Harmonisation started with FRBRoo and is now continuing with the creation of the LRMoo model, providing an object-oriented version designed as an extension of CIDOC CRM. Compared to FRBRoo, LRMoo is more streamlined and operates at a higher level of generality, while retaining its full expressiveness (Riva and Žumer 2018).

On the one hand, all of the above promise mechanisms to bridge the mentioned data structure gaps and overcome the so-called *silo effect*, where different institutions or projects collect and manage their data separately, without ensuring adequate interoperability, communication, or information sharing. On the other hand, slow adoption and uneven use of conceptual models for intangible cultural heritage by institutions and projects are evident. The reason for the slow acceptance of conceptual models and their implementations for intangible cultural heritage can be explained by the fact that formal ontologies are created through a top-down approach – a small number of information and data scientists and heritage experts are involved in a process that is quite intellectually demanding, hermetic and time-consuming. The modelling is quite abstract, and it is difficult to explain the models and mapping process to a broader community of heritage experts, holders and practitioners of intangible cultural heritage. Another factor for the gradual adoption of formal ontologies is that they are not designed to define a data structure that serves to store and enter data. They aim to integrate information from different sources by providing a common and extensible semantic framework that various systems can map. Therefore, it is important to further focus on defining the mapping between different systems and conceptual reference models and developing intuitive user interfaces that will enable simple and quick documentation of all aspects of intangible heritage.

ANALYSIS OF ENTITIES AND PROCESSES RELATED TO INTANGIBLE CULTURAL HERITAGE ACCORDING TO THEIR REPRESENTATION IN INFORMATION SYSTEMS

In the chapter on digitisation, the possibilities and limitations of recording aspects of intangible culture through multiple media have already been discussed, and here, the analysis of the involved entities and processes will be presented concerning the possibility of representation in information systems at the metadata level. Focus will be placed on how existing metadata standards can describe different entities and processes and what standardised reference sources (thesauri, classifications, registries, databases) can be used for their unique identification, description, and indexing.

According to UNESCO's 2003 Convention definition, intangible cultural heritage includes various kinds of concepts: "practices" as a sort of events, "expressions" as a kind of intellectual, artistic or craftsmanship realisations, "knowledge" as an epistemological concept, as well as "the instruments, objects, artefacts" as physical things and "cultural spaces" as places. By this definition, intangible cultural heritage is not an entity per se, but a set of diverse but deeply related and interconnected entities and processes.

Determining the types of entities and processes related to intangible cultural heritage was needed to enable analysis and mapping according to conceptual models and data standards. The following working typology is based on several articles that researched data modelling of folklore and oral traditions (Nicolas 2005), archiving folk music and dance culture (Marolt et al. 2009), organising digital archives of tangible and intangible cultural heritage (Wijesundara and Sugimoto 2018), modelling a representation protocol for traditional crafts (Zabulis et al. 2022) and developing a web-based platform for traditional craft documentation (Partarakis et al. 2022).

Mapping identified entities and processes to individual classes in the CIDOC-CRM model is also discussed to clarify information modelling, structure, and integration. At the same time, data value standards or controlled vocabularies and sources used in the description and indexing of individual instances of the entity will be identified.

1. Persons, groups, and communities

The complex relationship between individual and collective memory was discussed by Halbwachs back in 1950: "If collective memory derives its force and duration from a group of individuals, these are after all individuals who remember as members of a group." (Halbwachs 1997:94–95). Therefore, data models must be able to document individuals, groups and wider communities, and conceptualise the relationships between them and the relationships between them and other entities.

In representations of processes, persons and groups represent individual practitioners and/or social groups during events and individual processes. In CIDOC-CRM, person entities extend *class E21 Person* and are used to refer to individuals. Social groups refer to organisations of individuals and extend *E74 Group* (CIDOC CRM Special Interest Group 2022).

The data value standard for authority control used in archives and beyond is the *International Standard Archival Authority Record for Corporate Bodies, Persons, and Families (ISAAR (CPF)).* It provides guidelines for establishing and maintaining authority records for these entities in the context of archival descriptions.

2. Places

Intangible cultural heritage is often deeply rooted in the places where it originated, practised, or otherwise connected. Documenting places can help us understand the physical, social, and cultural contexts in which intangible cultural heritage is transmitted and performed. Zabulis et al. (2022) state that information on places can include imaginary places, as in legends or tales, and are not associated with a geographical location.

In CIDOC-CRM, places are instances of class *E53 Place*. The instances of *E53 Place* are usually determined by reference to the position of "immobile" objects such as buildings, cities, mountains, rivers, or dedicated geodetic marks. However, they may also be determined by reference to mobile objects (CIDOC CRM Special Interest Group 2022).

Two examples of data value standards used to control values of fields containing information about places are the *Getty Thesaurus of Geographic Na*mes (TGN) and the *GeoNames* geographical database.

3. Objects

Various human-made objects (artefacts) may be used during events (i.e., tools, instruments, machines, costumes, etc.) or produced in the processes (cultural heritage objects as results of creation and production) of manifestations of intangible cultural heritage. When discussing relationships, it is essential to enable links to the *materials* from which the objects are composed. When representing the crafting process, craft products can be associated with the crafting process, its place of occurrence, and the practitioner(s) involved in the process (Zabulis et al. 2022), while in the case of dance, having physical objects is not required (neither as tools nor as products), we already have practitioner(s) involved in the process of dance performance.

CIDOC-CRM class E24 Physical Human-Made Thing comprises all persistent physical items purposely created by human activity. This class comprises objects, such as a sword, and human-made features, such as rock art (for example, a "cup and ring" carving on bedrock is regarded as an instance of E24 Physical Human-Made Thing) (CIDOC CRM Special Interest Group 2022).

The International Standard Name Identifier (ISNI) is an identifier system uniquely identifying persons and organisations involved in creative activities.

4. Knowledge and skills

Zabulis et al. (2022) discuss the concepts of knowledge and skill in the context of intangible dimensions of intangible cultural heritage. However, these two concepts needed to be further differentiated. Knowledge is a complex conceptual, mental, and cognitive entity which is not observable nor measurable in a direct way (cf. Buckland's *information-as-knowledge*). At the same time, skill is closely related to performing intangible cultural heritage (as one of the expressions of knowledge).[1]

Here, we can apply Michael Polanyi's concept of "tacit knowledge" (Polanyi 1958) that encompasses skills, ideas and experiences that people have, but could not be easily expressed and explicated verbally, in text or by plan of procedure. In *Knowledge Taxonomy* (Alavi and Leidner 2001), tacit knowledge is further divided into "cognitive tacit" (mental models of knowledge) and "technical tacit" (know-how applicable to specific work). It is also necessary to distinguish cognitive, mental, and technical knowledge (know-how) from performing the skill.

Zack (1999:46) sets "explicit knowledge" as a contrast to "tacit knowledge", which is useful in the context of formal conceptualisations and their data expression:

> "Tacit knowledge is subconsciously understood and applied, difficult to articulate, developed from direct experience and action, and usually shared through highly interactive conversation, story-telling and shared experience. Explicit knowledge, in contrast, can be more precisely and formally articulated. Therefore, although more abstract, it can be more easily codified, documented, transferred or shared."

To map knowledge and know-how aspects of intangible cultural heritage to CIDOC-CRM, we can use only one class – *E28 Conceptual Object*. CIDOC CRM Special Interest Group (2022) defines that this class comprises non-material products of our minds and other human-produced data that have become objects of a discourse about their identity, circumstances of creation or historical implication.

For modelling the safeguarding of intangible cultural heritage within the CIDOC CRM model, it is crucial to how the model conceptualises the existence of *E28 Conceptual objects:*

[1] We can imagine a situation where a person has precise knowledge of some technique, but for example, because of age or illness, is not able to perform the technique as a skill. Furthermore, we can say that person could not, or could only to some extent, teach someone (by words, or limited moves) the skill without actually performing the skill in its full extent.

> "Characteristically, instances of this class are created, invented or thought by someone, and then may be documented or communicated between persons. Instances of E28 Conceptual Object can exist on more than one particular carrier simultaneously, such as paper, electronic signals, marks, audio media, paintings, photos, human memories, etc. They cannot be destroyed. They exist as long as they can be found on at least one carrier or in at least one human memory. Their existence ends when the last carrier and the last memory are lost." (CIDOC CRM Special Interest Group 2022:77)

In CIDOC-CRM, *E28 Conceptual Object* is a superclass of two important classes:

1) *E89 Propositional Object* class "comprises immaterial items, including but not limited to stories, plots, procedural prescriptions, algorithms, laws of physics or images that are, or represent in some sense, sets of propositions about real or imaginary things and that are documented as single units or serve as a topic of discourse." (CIDOC CRM Special Interest Group 2022:106);

2) *E55 Type* class comprises concepts denoted by terms from thesauri and controlled vocabularies used to characterise and classify instances of CIDOC CRM classes. It is important to note that instances of the *E55 Type represent* concepts in contrast to instances of the *E41 Appellation*, which are used to *name* instances of CIDOC CRM classes. According to CIDOC-CRM, appellations are cultural constructs; as such, they have a context, a history, and a use in time and space by some group of individuals, which is vital for modelling and recording local terminology.

Furthermore, we must distinguish knowledge from its performance. As Buckland pointed out: "If you can touch it or measure it directly, it is not knowledge." (Buckland 1991:352). Also, it is important to emphasise that although the practice of intangible cultural heritage is not material *per se,* it still has physical manifestation that can be transmitted, measured, or recorded.[2]

[2] For example, vocal performance is not tangible, but its realisation is a physical phenomenon – sound as a wave motion, oscillation in pressure, stress and particle displacement and velocity in medium (air), can stimulate human auditory organs or convert mechanical wave energy into electrical energy (audio signals) of measuring or recording devices.

5. Events

Several researchers have confirmed that an event-oriented approach is essential for modelling intangible cultural heritage, in contrast to the object-oriented or item-centric perspective that has been dominant thus far (Wijesundara and Sugimoto 2018).

CIDOC-CRM E5 Event is a class comprising distinct, delimited, and coherent processes and interactions of a material nature in cultural, social, or physical systems. Mental processes are considered events in cases connected with the material externalisation of their results, for example, the creation of a poem or a performance (CIDOC CRM Special Interest Group 2022).

Zabulis et al. (2022:722–723) also use CIDOC-CRM for modelling events (as instances of class *E5 Event*), whereby two key events are singled out as the building blocks of contextualisation narratives and crafting processes: "Respectively, they are called 'contextual' and 'crafting' events and are used to represent changes in (a) social, technological, and economic systems and (b) materials that are transformed into craft products, respectively."

Models and information systems must foresee the creation of metadata records representing certain events. This makes it possible to create a record for a particular event when documenting it, whether it is an event currently happening or reconstructing an event from the past. All this enables us to record the history and changes of a particular phenomenon of intangible cultural heritage. The metadata record for an event is the key and central point of linking with contextual entities (place, time, participants) and objects collected or represented (cf. Buckland's *evidence associated with* the events) or representations of the events themselves (audio-visual recordings, memoirs etc.).

6. Processes

As already mentioned, events are central points in modelling intangible cultural heritage. However, for its exhaustive documentation, it is no longer enough to document only events, but the actual process as well (craftsmanship skills, dance moves, etc.). Kettula and Hyvönen (2012) developed process-centric metadata for intangible cultural heritage, whereby process descriptions can be linked to related tangible and intangible objects in collections and Linked Data repositories on the web, facilitating rich and detailed semantic recommendations for end-users (they tested and evaluated this idea by creating a metadata model for representing cultural processes, and applied it to the video documentation of traditional shoemaking). As part of the development project of the web-based platform for traditional craft documentation, Partarakis et al. (2022) have

designed process schemas that can be considered as the conceptualisation of an activity diagram that is authored based on the understanding of the crafting process.

CIDOC-CRM uses the E7 Activity class as a broader class for expressing activities for various processes and procedures. The class (subclass of *E5 Event*) is a class that comprises actions intentionally carried out by instances *E7 Activity* of *E39 Actor* that result in changes of state in the cultural, social, or physical systems documented. Here are two examples of the properties found in the *E7 Activity* class:

- P33 used specific technique (was used by): E29 Design or Procedure
- P125 used object of type (was type of object used in): E55 Type (CIDOC CRM Special Interest Group 2022:63)

The following scheme (Table 3) illustrates the basic mapping of key entities related to intangible cultural heritage according to the CIDOC-CRM conceptual reference model for the heritage domain. The scheme takes the *Lepoglava lacemaking* phenomenon as an example.

<table>
<tr><th>Entity</th><th colspan="2">Person as carrier of intangible cultular heritage</th><th>Object</th><th colspan="2">Event</th><th>Object</th></tr>
<tr><td>Type</td><td>Person</td><td>Knowledge / skill</td><td>Object required for process of intangible cultular heritage</td><td colspan="2">Production of lace</td><td>Object as product of intangible cultular heritage</td></tr>
<tr><td rowspan="2">Examples</td><td rowspan="2">Unknown woman from Lepoglava</td><td>Knowledge of Lepoglava lacemaking</td><td rowspan="2">Tools needed for Lepoglava lacemaking</td><td>Lepoglava (E53 Place)</td><td>1921 (E52 Time-Span)</td><td rowspan="2">Lepoglava lace</td></tr>
<tr><td>Skills for Lepoglava lacemaking</td><td colspan="2">Related techniques materials (E57 Material)</td></tr>
<tr><td>CIDOC CRM classes</td><td>E21 Person</td><td>E28 Conceptual Object</td><td>E24 Physical Human-Made Thing</td><td colspan="2">E7 Activity
E12 Production</td><td>E24 Physical Human-Made Thing</td></tr>
</table>

Table 3: Mapping of key intangible cultural heritage related entities to CIDOC-CRM classes, using the example of the 'Lepoglava lacemaking'

RESULTS

The previous analysis aims to identify intangible cultural heritage documentation gaps in contemporary metadata standards, registries, and vocabulary sources. While thesauri, classifications and authority files can be used efficiently for indexing contextual entities (place, time, persons, groups), they cannot adequately represent the dynamic nature of intangible cultural heritage (such as events and processes) nor tacit aspects (such as knowledge or skills). Conceptual reference models such as CIDOC-CRM can serve as a common language for information interoperability and integration. Still, extensions and application profiles need to be developed to provide streamlined processing of information related to the domain of intangible cultural heritage. The following table (Table 4) summarises the types of entities or processes, examples of their instances, types of recording and archiving media, relevant documentation standards, and individual information systems that serve as reference sources, and a lack of them.

Furthermore, it is important not to consider the mentioned entities and processes separately, out of context. Documenting and understanding interconnections between people, places, objects, events, knowledge, and processes is essential. This provides a more comprehensive, contextual, and meaningful view and understanding of cultural practice and tradition. Furthermore, it is crucial to facilitate a dynamic, time-based representation of intangible cultural heritage. Documenting connections in temporally and spatially aware modes allows the tracking of changes and understanding the evolution of practices, traditions, or knowledge over time. Explicit recording and description of these interconnections and systematic documentation enable deeper insights and visualisations of how different cultural practices might influence each other, who the key practitioners are, and the spatiotemporal distribution of intangible cultural heritage manifestations.

Carboni and de Luca claim that the dichotomy between tangible and intangible is an accidental social construct and, consequently, ineffective for creating holistic documentation, and therefore suggest "placing focus on the construction of meaning, and the living relations between performances and objects, presenting a full account of the relationships between person-object-event." (Carboni and de Luca 2016:3).

Type of entity or process	Examples of types of instances	Media	Documentation standards	Type of information system (- reference sources)
Persons, groups, and communities	• individuals • groups • communities	• Photography • Audio recordings • Video recordings	Data value / authority records standards: • ISAAR (CPF) • LRM	Authority database: • ULAN • ISNI • VIAF
Places	• cultural spaces • imaginary places	• Photo • Video recordings • 3D models of space	Data value / thesaurus standards: • ISO 2788 • ISO 5964 • ISO 25964	Thesaurus: • TGN Geographical databases: • GeoNames
Objects	• artefacts • tools • instruments • books • journal articles	• Photography • 3D models • Video recordings	Data structure / collection management: • SPECTRUM, Data exchange: • LIDO	Cultural inventories: • CONA • The Register of Cultural Property of the Republic of Croatia • Collection management systems in the GLAM sector
Events	• social practices • rituals • festive events	• Photography • Audio recordings • Video recordings	Conceptual reference model: • CIDOC-CRM	
Knowledge	• knowledge	Not applicable[3]	Conceptual reference model: • CIDOC-CRM	Thesaurus: • AAT Classifications: • ICONCLASS Knowledge base: • Wikidata
Processes	• traditional craftsmanship	• Music notations • Dance notations • Motion Capture (MoCap)	Conceptual reference model: • CIDOC-CRM	

Table 4: Types of entities or processes, examples of their instances, types of recording media, relevant documentation standards and individual information systems that serve as reference sources

[3] Knowledge is a complex conceptual, mental, cognitive entity which is not observable or measurable in a direct way (cf. Buckland's *information-as-knowledge*).

CONCLUSION

In conclusion, digital humanities offer vital potential for documenting, digitising, archiving, and disseminating intangible cultural heritage, fundamentally impacting safeguarding, research, and ensuring vitality. Examining these methods through a documentation and data management perspective reveals a demanding need to shift from high-level descriptive data management plans to more precise, granular data standards specifying structure, content, value, and data encoding.

As the discussion has shown, the existing data standards are primarily object-oriented, supporting item-centred metadata that are more suitable for material culture and, therefore, cannot adequately represent the intangible aspects of culture. Further challenges lie in achieving data interoperability across diverse and heterogeneous sources related to intangible cultural heritage, especially given the differing standards among heritage institutions, research communities and projects.

As an introduction to the analysis of relevant data standards and conceptual models, Michael Buckland's theoretical approach was presented in order to provide insight into the fundamental concepts of information sciences like knowledge, information, data and documents, as it is highly relevant in addressing the highlighted concepts with a focus on their tangible and intangible facets.

After an analysis of existing relevant standards and conceptual models, the working typology of entities and processes related to intangible cultural heritage is presented and further discussed regarding the abilities and challenges for their recording, archiving, processing, and dissemination in the global information environment.

The study identified gaps in documenting intangible cultural heritage using current metadata standards, registries, and vocabularies. While existing tools can be used efficiently for indexing contextual entities such as place, time, persons, and groups, they inadequately capture dynamic intricacies and subtle facets of intangible cultural heritage, such as events processes and tacit knowledge.

Overcoming the limitations of existing data standards, as well as the possibility of integrating heterogeneous data sources, is enabled by conceptual reference models, i.e. formal ontologies, that provide a formal description of the conceptualisation of a specific domain by a specification of necessary concepts and relationships among those concepts. They are a reference point for various information systems, software agents, and services for exchanging information and knowledge. They also promise functionalities of reasoning and inference, enabling more intelligent data analysis and knowledge extraction.

The CIDOC-CRM is the most frequently used conceptual reference model in

the broader domain of heritage and digital humanities, serving as a common ground for fostering information interoperability and cohesion. Therefore, the paper discusses mapping identified entities and processes according to the mentioned model. Nevertheless, the moderate uptake of conceptual reference models was identified, attributed to their design, not primarily intended for data storage and entry, but for integrating information across diverse sources using a shared semantic framework. Future efforts should emphasise mapping between systems and reference models and creating user-friendly interfaces, all with the goal of comprehensive documentation for all entities, processes and interconnections present in the domain of cultural heritage. These steps must enable data workflows to support models of participatory cooperation between heritage experts, researchers, and community members in a joint effort that includes digitisation, documentation, research, safeguarding and revitalising intangible cultural heritage.

REFERENCES

ALAVI, Maryam and LEIDNER, Dorothy E. 2001. "Knowledge management and knowledge management systems: conceptual foundations and research issues". *MIS Quarterly,* vol. 25/1:107–136. DOI: https://doi.org/10.2307/3250961

BACA, Murtha, Patricia HARPRING, Elisa LANZI, Linda MCRAE and Ann WHITESIDE. 2006. *Cataloging cultural objects: a guide to describing cultural works and their images.* Chicago: American Library Association.

BERRYMAN, Donna R. 2019. "Ontology, epistemology, methodology, and methods: information for librarian researchers". *Medical Reference Services Quarterly,* vol. 38/3:271–279. DOI: https://doi.org/10.1080/02763869.2019.1623614

BUCKLAND, Michael K. 1991. "Information as thing". *Journal of the American Society for Information Science,* vol. 42/5:351–360. DOI: https://doi.org/10.1002/(SICI)1097-4571(199106)42:5%3C351::AID-ASI5%3E3.0.CO;2-3

BURNETTE, Margaret H., Sarah C. WILLIAMS and Heidi J. IMKER. 2016. "From plan to action: successful data management plan implementation in a multidisciplinary project". *Journal of eScience Librarianship,* vol. 5/1. DOI: https://doi.org/10.7191/jeslib.2016.1101

CARBONI, Nicola and Livio de LUCA. 2016. "Towards a conceptual foundation for documenting tangible and intangible elements of a cultural object". *Digital Applications in Archaeology and Cultural Heritage,* vol. 3/4:108–116. DOI: https://doi.org/10.1016/j.daach.2016.11.001

CIDOC CRM Special Interest Group. 2022. *Definition of the CIDOC Conceptual Reference Model.* Ver. 7.1.2. Eds. Chryssoula Bekiari, George Bruseker, Erin Canning, Martin Doerr, Philipe Michon, Christian-Emil Ore, Stephen Stead and Athanasios Velios. CIDOC CRM Special Interest Group. URL: https://cidoc-crm.org/Version/version-7.1.2 (access January 2, 2023).

COLADANGELO, L. P. 2020. "Thesaurus and ontology construction for Contra Dance: knowledge organization of a North american folk dance domain". *Knowledge Organization,* vol. 47/7:523–542. DOI: https://doi.org/10.5771/0943-7444-2020-7-523

FAFALIOS, Pavlos, Yannis MARKETAKIS, Anastasia AXARIDOU, Yannis TZITZIKAS and Martin DOERR. 2023. "A workflow model for holistic data management and semantic interoperability in quantitative archival research". *arXiv.org.* URL: http://arxiv.org/abs/2301.07676 (access February 2, 2023).

GOPAL, Krishna. 2016. "Data curation: the processing of data". *Journal of Information Management,* vol. 3/1:71–79. URL: https://splpjim.org/wp-content/uploads/2016/11/8_Gopal_Data_Curation.pdf (access January 2, 2023).

MCKENNA, Gordon and Efthymia PATSATZI, eds. 2009. *Spectrum: The UK Museum Documentation Standard.* Collections Trust.

HALBWACHS, Maurice. 1997. *La Mémoire Collective.* Paris: Albin Michel.

HOU, Yumeng, Sarah KENDERDINE, Davide PICCA, Mattia EGLOFF and Alessandro ADAMOU. 2022. "Digitizing intangible cultural heritage embodied: state of the art". *Journal on Computing and Cultural Heritage* (JOCCH), vol. 15/3:1–20. DOI: https://doi.org/10.1145/3494837

HYVÖNEN, Eero. 2009. "Semantic portals for cultural heritage". In *Handbook on ontologies,* eds. Steffen Staab and Rudi Studer. Berlin – Heidelberg: Springer, 757–778. DOI: https://doi.org/10.1007/978-3-540-92673-3_34

IFLA Study Group on the Functional Requirements for Bibliographic Records. 1998. *Functional Requirements for Bibliographic Records. Final Report.* München: K. G.Saur. URL: https://repository.ifla.org/handle/123456789/830

KARR, Alan. 2010. *Metadata and paradata: information collection and potential initiatives.* Washington, DC: National Institute of Statistical Sciences. URL: https://www.niss.org/sites/default/files/research_attachments/Metadata%20vs%20Paradata-FT.pdf (access January 2, 2023).

KETTULA, Suvi and Eero HYVÖNEN. 2012. "Process-centric cataloguing of intangible cultural heritage". Paper presented at: *CIDOC2012 – Enriching Cultural Heritage (Helsinki, Finland, June 10–14, 2012).* URL: https://cidoc.mini.icom.museum/wp-content/uploads/sites/6/2018/12/kettula.pdf (access January 2, 2023).

LISENA, Pasquale, Manel ACHICHI, Pierre CHOFFÉ, Cécile CECCONI, Konstantin TODOROV, Bernard JACQUEMIN and Raphaël TRONCY. 2018. "Improving (re-) usability of musical datasets: an overview of the DOREMUS Project". *Bibliothek*

Forschung Und Praxis, vol. 42/2:194–205.
DOI: https://doi.org/10.1515/bfp-2018-0023

MAROLT, Matija, Janez Franc VRATANAR and Gregor STRLE. 2009. "Ethnomuse: archiving folk music and dance culture". In *IEEE EUROCON 2009 (St. Petersburg, Russia, 18-23 May 2009).* [Piscataway, NJ]: IEEE, 322–326.
DOI: https://doi.org/10.1109/EURCON.2009.5167650

MOUNTCASTLE, Amy. 2010. "Safeguarding intangible cultural heritage and the inevitability of loss: a Tibetan example". *Studia Ethnologica Croatica,* vol. 22:339–359.
URI: https://hrcak.srce.hr/62257

MUSTAFA, Armin, Hansung KIM, Jean-Yves GUILLEMAUT and Adrian HILTON. 2016. "Temporally coherent 4D reconstruction of complex dynamic scenes". In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (Las Vegas, NV, USA, June 27-30, 2016).* [Los Alamitos, CA]: IEEE, 4660–4669. Las Vegas, NV, USA: IEEE. DOI: https://doi.org/10.1109/CVPR.2016.504

NICOLAS, Yann. 2005. "Folklore requirements for bibliographic records: oral traditions and FRBR". *Cataloging & Classification Quarterly,* vol. 39/3–4:179–195.
DOI: https://doi.org/10.1300/J104v39n03_11

NIKOČEVIĆ, Lidija, Ljiljana GAVRILOVIĆ, Mirela HROVATIN, Daniela Angelina JELINČIĆ, Dragana Lucija RATKOVIĆ, Amy MOUNTCASTLE, Petar SIMONIČ and Tvrtko ZEBEC. 2012. "Discussion: Culture or heritage? The problem of intangibility". *Etnološka tribina,* vol. 42/35:57–112. URI: https://hrcak.srce.hr/93768

PARTARAKIS, Nikos, Voula DOULGERAKI, Effie KARUZAKI, George GALANAKIS, Xenophon ZABULIS, Carlo MEGHINI, Valentina BARTALESI and Daniele METILLI. 2022. "A web-based platform for traditional craft documentation". *Multimodal Technologies and Interaction,* vol. 6/5:37. DOI: https://doi.org/10.3390/mti6050037

POLANYI, Michael. 1958. *Personal knowledge: towards a post-critical philosophy.* Chicago: University of Chicago Press.

Reference Model for an Open Archival Information System (OAIS). 2012. Washington, DC: Council of the Consultative Committee for Space Data Systems.
URL: https://public.ccsds.org/pubs/650x0m2.pdf (access January 2, 2023).

RIVA, Pat, Patrick Le BOEUF and Maja ŽUMER. 2017. *IFLA Library Reference Model: a conceptual model for bibliographic information.* Den Hag: IFLA.
URL: https://repository.ifla.org/handle/123456789/40 (access January 2, 2023).

RIVA, Pat and Maja ŽUMER. 2018. "FRBRoo, the IFLA Library Reference Model, and Now LRMoo: A Circle of Development". Paper presented at: *IFLA WLIC 2018 – Transform Libraries, Transform Societies (Kuala Lumpur, Malaysia, August 24-30, 2018).* URL: https://library.ifla.org/id/eprint/2130/ (access January 2, 2023).

SHREEVES, Sarah L. and Melissa H. CRAGIN. 2008. "Introduction: Institutional repositories: current state and future". *Library Trends,* vol. 57/2:89–97.

TASOVAC, Toma, Sally CHAMBERS and Erzsébet TÓTH-CZIFRA. 2020. *Cultural heritage data from a humanities research perspective: a DARIAH Position Paper.* Ver. 1, October 8, 2023. URL: https://hal.science/hal-02961317 (access January 2, 2023).

UNESCO. 2003. *Convention for the safeguarding of the intangible cultural heritage.* Paris: UNESCO. URL: https://unesdoc.unesco.org/ark:/48223/pf0000132540 (access January 2, 2023).

WIJESUNDARA, Chiranthi and Shigeo SUGIMOTO. 2018. "Metadata model for organizing digital archives of tangible and intangible cultural heritage, and linking cultural heritage information in digital space". *LIBRES: Library and Information Science Research Electronic Journal,* vol. 28/2:58–80. DOI: https://doi.org/10.32655/LIBRES.2018.2.2

WILKINSON, Mark D., Michel DUMONTIER, IJsbrand Jan AALBERSBERG, Gabrielle APPLETON, Myles AXTON, Arie BAAK, Niklas BLOMBERG, et al. 2016. "The FAIR Guiding Principles for scientific data management and stewardship". *Scientific Data,* vol. 3:160018. DOI: https://doi.org/10.1038/sdata.2016.18

ZABULIS, Xenophon, Nikolaos PARTARAKIS, Carlo MEGHINI, Arnaud DUBOIS, Sotiris MANITSARIS, Hansgeorg HAUSER, Nadia Magnenat THALMANN, et al. 2022. "A representation protocol for traditional crafts". *Heritage,* vol. 5/2:716–741. DOI: https://doi.org/10.3390/heritage5020040

ZACK, Michael H. 1999. "Managing codified knowledge". *Sloan Management Review,* vol. 40/4:45–58.

Informacijska i dokumentacijska perspektiva na nematerijalnu kulturnu baštinu u kontekstu istraživanja u digitalnoj humanistici

Goran Zlodi

U ovom radu predstavljena je ključna uloga metoda digitalne humanistike u očuvanju, istraživanju i revitalizaciji nematerijalne kulturne baštine. U tekstu se sagledavaju metode digitalizacije kroz perspektive dokumentacije i upravljanja podacima, zagovarajući standarde koji omogućuju interoperabilnost podataka na razinama strukture, sadržaja, vrijednosti i kodiranja. Rad kritički analizira postojeće podatkovne standarde i konceptualne referentne modele, uzimajući u obzir njihove mogućnosti i ograničenja pri dokumentiranju i predstavljanju nematerijalne kulturne baštine. Predstavljena je i raspravljena tipologija entiteta i procesa vezanih uz nematerijalnu kulturnu baštinu. Elaborirano je i mapiranje identificiranih entiteta i procesa na pojedinačne klase u CIDOC-CRM modelu s obzirom na objavljivanje podataka i semantičku interoperabilnost u kontekstu Semantičkog weba.

Ključne riječi: *nematerijalna kulturna baština, digitalna humanistika, podatkovni standardi, konceptualni referentni modeli, povezani otvoreni podaci, dokumentacija, informacije*