TU Delft

Delft University of Technology

Document Version
Final published version

Licence
CC BY

Citation (APA)
Ramírez Eudave, R., Ferreira, T. M., & Giardina, G. (2026). Communities co-creating metadata: a new paradigm towards FAIR everyday heritage. *npj Heritage Science*, *14*(1), Article 501. https://doi.org/10.1038/s40494-026-02706-1

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

npj | heritage science

Article

https://doi.org/10.1038/s40494-026-02706-1

Communities co-creating metadata: a new paradigm towards FAIR everyday heritage

Check for updates

Rafael Ramírez Eudave[1] ✉, Tiago Miguel Ferreira[2] & Giorgia Giardina[1]

Efforts to make cultural heritage data FAIR have largely focused on expert-driven metadata within formal institutions, often marginalising community-valued heritage. This paper proposes a process-oriented framework for community-driven metadata co-creation aligned with FAIR principles while remaining sensitive to everyday heritage and community expression. The framework introduces a five-stage lifecycle (recording, submission, curation, release, and use/reuse) and repositions heritage institutions as co-authors rather than sole authorities. It is formalised through a metadata model compatible with the Europeana Data Model, supporting multilingual expression, coexistence of community folksonomies with institutional taxonomies, and semantic enrichment. A proof-of-concept survey explores its recording stage, showing that community contributors generate rich, narrative descriptions, while also revealing limitations in metadata completeness. This study provides empirical insight into initial conditions for community-authored metadata, thus contributing to a conceptual foundation for more inclusive participation in the digitisation of cultural heritage and its valorisation outside the scope of formal memory institutions.

Spanish writer Ramón del Valle-Inclán (1866–1936) once wrote, «*Things are not as we see them, but as we remember them*», highlighting the importance of memory in shaping our understanding of ourselves and the world. As a global society, we seek to preserve our collective memory in the form of data, and we are generating more data than at any other point in history. This phenomenon is evident across nearly every domain, including cultural heritage. However, the sheer volume and diversity of data pose significant challenges for its long-term preservation, accessibility, transparency, and capacity to support responsible innovation[1].

In response, experts and institutions are advocating for approaches that make data more understandable and reusable, thereby facilitating cooperation and partnership. A key development in this effort is the establishment of the FAIR principles, designed to make data Findable, Accessible, Interoperable, and Reusable. The operationalisation of these principles is often achieved by creating commonly understandable, explanatory data about real or digital objects. In other words, metadata[2].

Metadata consists of a set of structured elements, typically in textual form, that describe and explain a resource to facilitate its retrieval and use. Metadata schemas are generally categorised as either descriptive (detailing an asset's content and characteristics) or administrative (containing data related to its management). In the field of cultural heritage, descriptive metadata facilitates the association between cultural heritage manifestations in the real world and their digital representations[3].

Nonetheless, the creation and management of metadata remain largely under the control of experts and specialised institutions, with occasional community participation[4]. This risks biasing what is represented and described, particularly when approaching intangible heritage[5]. Similarly, heritage assets not recognised by or held in the custody of formal memory institutions (usually collectively named GLAM: Galleries, Libraries, Archives, and Museums) often fall outside efforts to make cultural knowledge FAIR[6]. Furthermore, the pursuit of standardised frameworks to facilitate metadata exchange can inherently conflict with a core value of cultural heritage: its role as a meeting place for human diversity, dialogue, and multiple perspectives[7].

Current efforts to make cultural heritage data FAIR are predominantly carried out by and for experts, with a primary focus on assets held by memory institutions[8]. This prevailing focus inevitably shapes the definition of cultural value. Consequently, there is a significant risk of: (1) overlooking the cultural assets that individuals and communities themselves value, and (2) transmitting to the future a hegemonic, expert-led perspective that fails to incorporate the opinions and lived experiences of minority groups.

While the value of communities is widely recognised, effective mechanisms to position them as primary promoters of metadata creation are lacking. Even when communities participate, they are often relegated to the role of guest contributors. According to Barabucci et al.[9], a persistent lack of trust in non-expert contributions is evident in participatory models that restrict communities' ability to shape their input according to their own expressive resources and needs. Indeed, the design of crowdsourcing platforms is often constrained by the ontological assumptions and beliefs of experts. Consequently, even when communities are granted access to

[1]Department of Geoscience and Engineering, Delft University of Technology, Delft, The Netherlands. [2]CERIS, Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal. ✉e-mail: R.RamirezEudave@tudelft.nl

metadata enrichment, their participation is typically relegated to microtasks and formulaic approaches, framed entirely by the taxonomic language and terminology set by institutions[7].

This tension between institutional language and community expression reflects a core divergence between standardisation and diversity. Although a great variety of templates, formats, and vocabularies have been developed for metadata creation, they are primarily intended to facilitate transactions amongst experts and specialised stakeholders[7]. Furthermore, most metadata models embed a series of assumptions that can clash with cultural diversity. A clear example is the unquestioned adoption of English as a *lingua franca*, which subsequently shapes and constrains underlying vocabularies[10]. This authoritative approach reflects an institutional reluctance to move beyond standardised terms. In practice, however, semantic diversity is unavoidable, even when using common dictionaries and terminology within the same field[2]. Thus, although metadata has been central to implementing FAIR principles in cultural heritage, it often remains neither useful nor understandable for non-expert researchers and citizens[9].

Communities have great potential for contributing to the metadata related to cultural heritage, as it is more than a technical tool but also a political and ethical instrument that can give voice to communities. An outstanding example in this sense is presented by Kim et al. when documenting archives from the perspective of their holding communities to challenge normative archival frameworks[11]. However, current initiatives often position community contributions as inputs that supplement already existing, expert-driven records. Community contributions follow a model based on handing over knowledge to enrich or supplement already existing, expert-driven records, without retaining control over knowledge and materials. This lack of control prevents continuous engagement and long-term continuity, thus hindering the development of a living archive, according to Allard et al.[12]. Knowledge sharing between communities and archivists, integral to cataloguing and description practices, should challenge mainstream notions of «expertise», as observed by Stevens et al.[13].

There is a lack of literature addressing the processes of participatory and crowdsourced metadata creation, despite their potential to serve as windows into collective memory. However, experiences such as that of Roeschley et al. enabled the creation of metadata from free-text contributions about archival objects, adding nuanced contextual information and knowledge. This work advocates for further research into the relationship between controlled-vocabulary metadata and community contributions, as well as into best practices for metadata creation in participatory archives[14].

Whilst community-generated metadata can be central to ensuring that users are able to find what they actually need (thanks to the use of common words and expressions), it remains challenging to make it interoperable and useful within the context of controlled taxonomies. In parallel, crowdsourced metadata creation has significant potential for uncatalogued or inadequately described materials, as illustrated by the work of Barber et al.[15]. This potential has been confirmed by experiences from several cataloguing projects discussed in this research, suggesting that crowdsourced metadata also offers a way to move beyond the limited capacities of professionals working in the documentation of archives and collections.

Whilst innovative, these and other similar initiatives still perpetuate the long-standing paradigm of focusing on collections and objects that already fall within the scope of formal memory institutions, without discussing how the benefits of community-contributed information could support other forms of heritage beyond these institutional circuits. This issue is particularly significant because individuals establish connections with their communities through personal objects and stories[16]. In this sense, envisioning a framework in which community knowledge forms the foundation of metadata, whilst expert knowledge serves to inform and support it, represents a disruptive proposition—one that is essential for the valorisation of a broader and more diverse spectrum of cultural heritage.

In summary, current practices are frequently unable or unwilling to recognise and document contributions from outside scientific and cultural institutions. Consequently, they fail to represent diverse views and interpretations of cultural heritage authentically, which hinders the potential benefits of FAIR knowledge in non-specialised contexts. Ultimately, under the current paradigm, only cultural heritage under the umbrella of formal institutions can truly aspire to be FAIR. Overcoming this paradigm would contribute to digital equity and social justice through the exercise of culture rights.

Creating community-centred FAIR metadata for the rich spectrum of cultural heritage, including that held outside the custody of formal memory institutions, remains a timely and urgent challenge. We address this by proposing the following research question: How can a co-creation framework for building FAIR-aligned metadata be conceptualised and designed to represent diverse perspectives on heritage held outside memory institutions?

In response, this project will develop a process-oriented model for co-creating cultural heritage metadata. This approach fundamentally reimagines the roles of heritage professionals and community members, positioning experts as co-authors and stewards of collective knowledge rather than sole arbiters of authoritative metadata[17]. By centring communities as key actors in the creation of FAIR metadata for their everyday heritage (that is, the places, objects, and memories they value) this framework aims to:

- Allow communities to initiate metadata creation, enabling heritage documentation beyond formal memory institutions.
- Enable the coexistence of standardised controlled vocabularies alongside contributions expressed using non-expert resources and language.
- Acknowledge diversity, enabling the documentation of a wide array of perspectives on heritage, both within and beyond traditional memory institutions.
- Deepen participation, moving beyond formulaic or micro-tasked contributions to foster continuous, substantive and meaningful public involvement in knowledge creation.
- Ensure community benefit, making communities direct beneficiaries of the FAIR metadata they co-create, ensuring it is both accessible and meaningful to them.

Methods

To address the research question, we must first establish a robust conceptual framework that characterises: (1) who is involved in the process of creating community-driven metadata; (2) how FAIR-compliant metadata about cultural heritage is integrated; and (3) what constitutes everyday heritage. These three pillars form the foundation for creating a metadata creation framework that is sensitive to the diverse actors within a community and their expressive needs, aligned with FAIR metadata best practices, and applicable to a wide range of heritage manifestations. Finally, the framework is explored through an experiment focusing on the Recording stage in which participants use it to document tangible or intangible manifestations of community-valued cultural heritage. Designed to illustrate the diversity of perspectives, this experiment tests the practical challenges of implementing this approach in real life.

Who: communities

This work adopts the concept of «community» as defined by the Council of Europe's Faro Convention (2005), which understands «heritage community» as people who value specific aspects of cultural heritage and who seek, within the framework of public action, to sustain and transmit them to future generations[18]. Communities are the primary agents of action, as the significance of cultural heritage is socially constructed through collective processes shaped by groups that share an interest in engaging with and interpreting the inherited past.

Heritage communities are therefore open and transversal groups, formed more or less spontaneously around a shared concern for heritage. Their members may operate across different territorial scales and may organise themselves informally or through more permanent associative structures. The Faro Convention does not prescribe a fixed model for heritage communities; rather, what defines them is their intention to act in

relation to the past and their shared commitment to cultural heritage. Such engagement may be driven by aesthetic, identity-related, social, or economic interests[19,20].

Understanding the various actors who could play a role in cultural heritage metadata is key to framing their respective rights, responsibilities, and restrictions. Current research often distinguishes between three groups: communities or citizens, experts from memory institutions, and the developers who create the digital environments for data management. While such an approach is often sufficient to describe the majority of current practices, in which communities are considered enrichers of expert-created metadata, it may not provide an adequate foundation for a co-creation approach[21,22]. For example, expert-guided processes are inherently unable to accommodate everyday heritage, as it is frequently out of the custody of formal institutions.

This points to a second limitation: the rigid barrier that often divides the custody and management of heritage manifestations in reality from their digital representations. In the context of everyday heritage, communities are situated on both sides of this boundary, as they are simultaneously the holders of tangible and intangible cultural heritage manifestations and the producers of digital resources about them (e.g., photographs, videos, texts). It is therefore necessary to transition towards a more horizontal and integrative model of interaction—one in which communities are more than just end users or contributors at the end of a pipeline.

A useful definition of roles for creating metadata about everyday heritage requires circularity rather than linearity and shared participation over both real-life and digital assets. Such a conceptual need resonates with the concept of «citizen curation», which has emerged to suggest a more horizontal interaction between actors involved in cultural heritage. Citizen curation is the practice of «*citizens applying curatorial methods to archival materials [...] to develop their own interpretations, share their own perspectives, and appreciate the perspectives of others*»[23]. Crucially, this practice requires an egalitarian infrastructure in which memory institutions, citizens, scholars, and other stakeholders share control over the assets and metadata they collectively produce.

Citizen curation partly aligns with the requirements we have identified in the process of creating metadata for cultural heritage not held in memory institutions. A central limitation, however, is that citizen curation is primarily linear and centred on assets within the custody of formal institutions. It is therefore possible to reformulate an alternative concept for «*communities applying curatorial methods to create metadata about their own cultural heritage, embedding their own interpretations and perspectives, and appreciating the collective co-created archive of metadata*». In practice, citizen curation should be operationalised through roles and role-specific actions. Daga et al.[23] comprehensively defined four roles and 37 role-specific actions in the context of general digital asset management. We have built upon these roles to develop our own set of roles and actions focused on metadata creation within a cyclic, rather than linear, process.

How: metadata

Metadata serves as a resource for explanation, designed to be interpretable and comprehensible to others. To fulfil this role, it must be produced according to a defined scheme or model. In the cultural heritage sector, the adoption of appropriate metadata standards is essential to ensuring that knowledge is FAIR[24]. Several metadata models exist within this domain, as many institutions have developed their own, often in response to collection-specific needs and established archival practices[25]. Increasingly, however, institutions have converged around the adoption of major supranational metadata schemes. This trend reflects a growing recognition of the need and value of making data accessible and shareable across boundaries.

Community-led documentation of everyday heritage calls for flexibility and adaptability. In other words, a metadata model capable of supporting diverse forms of expression, multiple languages, and varied manifestations of cultural heritage. It is also advantageous to adopt a model that is scalable and interoperable with large-scale repositories, thereby increasing the visibility of cultural assets and maximising the long-term benefits of FAIR data. We have analysed some of the most widely used models in this context to identify those that are better aligned with the expectations of community-driven metadata about everyday heritage.

The CIDOC Conceptual Reference Model (CIDOC CRM) is intended to support the integration and mediation of heterogeneous cultural heritage knowledge by offering an event-based framework in which assets are described as the outcomes of historic processes[26]. However, the model depends on the use of specific taxonomies and enforces strict control over terminology and interconnected assets. The LIDO (Lightweight Information Describing Objects) model is more descriptive and flexible, having been developed specifically for contributing museum object data to aggregation portals[27]. Both CIDOC CRM and LIDO operate on the assumption of a single, authoritative source of information originating from one institution, which poses challenges for the convergence of multiple metadata records relating to the same heritage manifestation.

Another key metadata scheme is the Europeana Data Model (EDM). Developed for the Europeana digital library, EDM is a multi-disciplinary standard built on existing frameworks such as Dublin Core and SKOS, and is aligned with CIDOC CRM. The model supports the representation of diverse digital cultural heritage objects from libraries, archives, and museums by distinguishing between physical objects and their digital representations, while promoting the principles of linked open data.

The EDM, unlike the CIDOC CRM and the LIDO, acknowledges that data about a single cultural object can be scattered across various institutions, websites, and community sources. It breaks the «one-record-per-item» paradigm by clearly separating the digital representation of an object (a scan, a photo, a video) from the conceptual object itself. This allows EDM to act as a unifying layer, aggregating and connecting diverse perspectives (e.g., a professional museum description in LIDO, a community-generated tag, or a contextual link to a Wikipedia entry) without forcing them into a single, homogenised record[28].

In the context of creating community-driven metadata for everyday heritage, EDM offers the flexibility required to represent heritage not only from memory institutions but also from communities themselves. It recognises that a community's knowledge may differ from that of a museum, and that the metadata communities generate is as valuable as any institutional catalogue record. This supports an egalitarian infrastructure in which diverse actors can contribute their own perspectives while remaining interoperable within a broader network. In other words, it enables a shift away from the paradigm in which communities merely enrich an expert-driven core. EDM is underpinned by several core principles: (i) it maintains a clear distinction between the description of the cultural object and that of its digital representation, (ii) it allows for multiple records about the same object, accommodating potentially contradictory statements, (iii) it supports objects that are composed of other objects, and (iv) it enables the inclusion of contextual resources, such as concepts from controlled vocabularies.

Version 2.4 (2017) of the EDM is structured around three core classes: the cultural heritage object (`edm:ProvidedCHO`), its digital representation (`edm:WebResource`), and an aggregation that binds them (`ore:Aggregation`). These are supplemented by four contextual classes for persons and/or organisations (`edm:Agent`), place and time (`edm:Place`, `edm:TimeSpan`), and knowledge documented in formal external sources (`skos:Concept`), as shown in Fig. 1. In addition, there is a class for licensing information (`cc:License`) for the whole metadata file.

A metadata model designed to support communities and everyday heritage must be sufficiently flexible to accommodate both minimal, low-barrier records and detailed, granular descriptions. EDM meets this requirement by allowing the creation of compliant metadata files using as few as ten of its 133 defined properties. Moreover, the official Mapping Guidelines recommend fulfilling a total of 31 specific properties to achieve a good quality metadata record.

Sometimes, the same property can be used throughout different classes; for instance, `dc:creator` can describe the creator of either a physical sculpture (the `edm:ProvidedCHO`) or a digital photograph of it (the

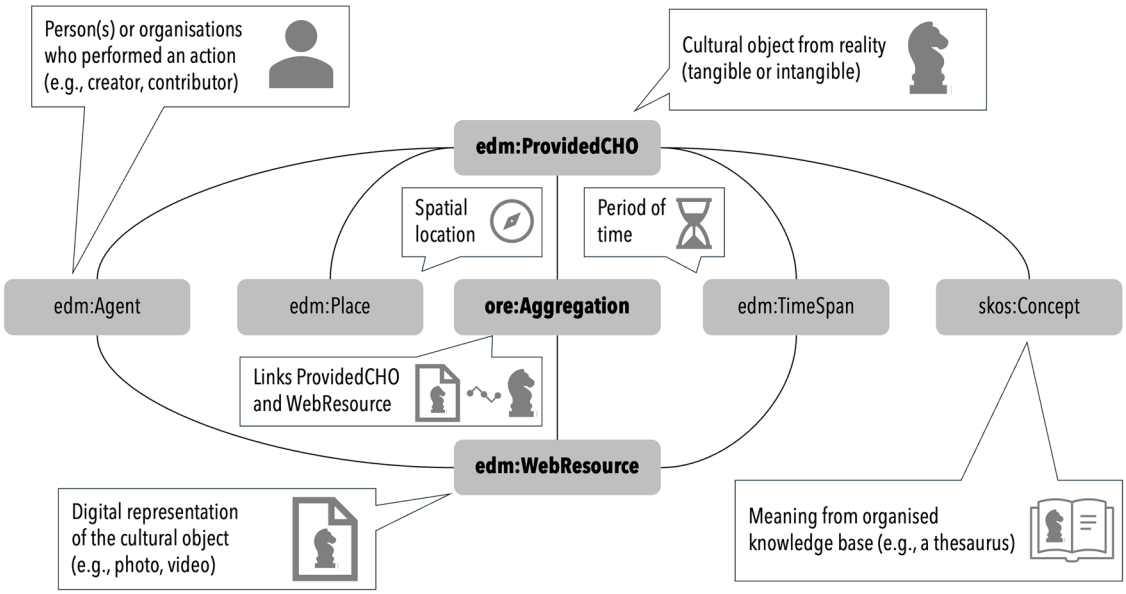

Fig. 1 | Schematic structure of the EDM based on its classes. Grey boxes represent the EDM classes, with those in bold being the core ones. White speech rectangles express and illustrate what each class refers to.

`edm:WebResource`). The full set of classes and properties used in this study are detailed in Table 1, with comprehensive descriptions available in the official Mapping Guidelines[29].

What: everyday heritage

Significant places, objects, and practices of cultural importance often fall outside the scope of memory institutions. Nevertheless, these components of cultural heritage are integral to collective memory and identity. These include everyday places and objects that gain cultural relevance over time or through their situation within significant spatio-temporal settings[30]. Such heritage is often valorised by minoritised or systematically excluded communities, which can hinder its recognition and protection.

In many instances, cultural significance is embedded within places rather than discrete objects. Places function as continuous happenings where tangible and intangible aspects converge within a unique setting[31]. The practice of registering a place's significance is itself a long-standing cultural tradition, often manifested physically through plaques commemorating historical events. It can be argued that these physical markers constitute a form of analogue metadata, as they annotate the physical space with contextual information. These vehicles of collective memory have a direct parallel in the digital realm, where metadata serves a similar function by conveying explanations about a given entity.

Everyday heritage can also reside in simple objects that connect places, moments, and people. An example is presented in Fig. 2, which depicts a small sign affixed to a shop door in the city centre of Delft, The Netherlands. The sign, requesting customers to maintain a 1.5-metre social distance, documents a common public health measure from the COVID-19 pandemic (2020–2023). This artefact not only reflects social practices from a specific historical moment but also incorporates traditionally recognised cultural elements, such as the motif of the kissing couple, a common theme in local handicraft. Overall, it represents a culturally significant creation responding to a traumatic moment in history that exists entirely outside formal memory institutions.

Documenting everyday heritage can also serve as a means of addressing traumatic memories associated with displaced objects, such as those removed from their original communities[32]. In this context, preserving memory and fostering a digital connection to a site has been explored as a strategy for restitution, exemplified by the UNESCO Virtual Museum of Stolen Cultural Objects[33], launched in September 2025. Furthermore, producing metadata about lost or destroyed cultural heritage could mitigate damage to the social fabric in the wake of war or extreme natural events[34,35]. Using metadata to convey the knowledge associated with cultural heritage may be a suitable way to ensure its preservation when physical safeguarding from decay or loss is, for various reasons, unattainable for the community that values it. This could involve annotating digital representations of lost environments to create spaces for sharing memories through metadata.

However, the uncontrolled production of digital data about common spaces can be counterproductive. As argued by Regazzoni et al.[36], it risks facilitating the commodification of cultural knowledge when dominated by commercial or political biases. Furthermore, such excessive production can lead to digital disorientation, which undermines the development of an open and inclusive digital sense of place. Ultimately, this can suppress difference and diversity rather than acknowledging them[37].

Consequently, it is relevant to make explicit decisions about responsibilities, roles, and participatory frameworks. According to Szabo et al.[38], effective digital approaches can emerge from a synergy between top-down media produced by experts and bottom-up approaches that incorporate local experiences and values. Experiences in citizen participation within cultural heritage suggest that lowering technological barriers enables communities to have a direct input into which heritage is valorised, documented, and preserved[39]. Therefore, participatory approaches aimed at representing community diversity should ideally be designed without prior assumptions about technologies or participants' digital literacy[40].

In this context, providing a comprehensive framework, rather than a specific technical tool, is critical. This ensures that potential implementations are built to satisfy requirements that prioritise communities, instead of assuming that universally applicable workflows and environments can fit any community, tradition, or purpose[41].

Framework for metadata creation: contents, roles, and processes

We propose a process-oriented framework for constructing community-led metadata that directly integrates the requirements of citizen curation ecosystems[23] with the structural logic of the Europeana Data Model. This approach is designed to be enabled by a digital environment, which provides support throughout the entire process and is defined by its own set of requirements.

We first address a recurrent challenge in integrating community contributions: the diversity of terminology and expression. We focus specifically on accommodating divergent terminology, language, and

Table 1 | Properties for each class, indicating if they are mandatory (⋆), recommended (•), or optional (◦)

Property	edm:ProvidedCHO	ore:Aggregation	edm:WebResource	edm:Agent	edm:Place	edm:TimeSpan	skos:concept	cc:License
dc:contributor	•	-	-	-	-	-	-	-
dc:coverage	•	-	-	-	-	-	-	-
dc:creator	◦	-	◦	-	-	-	-	-
dc:date	•	-	-	◦	-	-	-	-
dc:description	1	-	◦	-	-	-	-	-
dc:format	◦	-	◦	-	-	-	-	-
dc:identifier	•	-	-	◦	-	-	-	-
dc:language	⋆	-	-	-	-	-	-	-
dc:publisher	•	-	-	-	-	-	-	-
dc:relation	◦	-	-	-	-	-	-	-
dc:rights	◦	◦	◦	-	-	-	-	-
dc:source	•	-	◦	-	-	-	-	-
dc:subject	2	-	-	-	-	-	-	-
dc:title	1	-	-	-	-	-	-	-
dc:type	2	-	-	-	-	-	-	-
dcterms:alternative	•	-	-	-	-	-	-	-
dcterms:conformsTo	◦	-	◦	-	-	-	-	-
dcterms:created	•	-	◦	-	-	-	-	-
dcterms:extent	◦	-	◦	-	-	-	-	-
dcterms:hasFormat	◦	-	-	-	-	-	-	-
dcterms:hasPart	◦	-	◦	◦	◦	◦	-	-
dcterms:hasVersion	◦	-	-	-	-	-	-	-
dcterms:isFormatOf	◦	-	◦	-	-	-	-	-
dcterms:isPartOf	•	-	◦	◦	◦	◦	-	-
dcterms:isReferencedBy	◦	-	◦	-	-	-	-	-
dcterms:isReplacedBy	◦	-	-	-	-	-	-	-
dcterms:isRequiredBy	◦	-	-	-	-	-	-	-
dcterms:issued	•	-	◦	-	-	-	-	-
dcterms:isVersionOf	◦	-	-	-	-	-	-	-
dcterms:medium	◦	-	-	-	-	-	-	-
dcterms:provenance	◦	-	-	-	-	-	-	-
dcterms:references	◦	-	-	-	-	-	-	-
dcterms:replaces	◦	-	-	-	-	-	-	-
dcterms:requires	◦	-	-	-	-	-	-	-
dcterms:spatial	2	-	-	-	-	-	-	-
dcterms:tableOfContents	◦	-	-	-	-	-	-	-
dcterms:temporal	2	-	-	-	-	-	-	-
edm:aggregatedCHO	-	⋆	-	-	-	-	-	-

Table 1 (continued) | Properties for each class, indicating if they are mandatory (⋆), recommended (•), or optional (◦)

Property	edm:ProvidedCHO	ore:Aggregation	edm:WebResource	edm:Agent	edm:Place	edm:TimeSpan	skos:concept	cc:License
edm:begin	-	-	-	◦	-	◦	-	-
edm:currentLocation	◦	-	-	-	-	-	-	-
edm:dataProvider	-	⋆	-	-	-	-	-	-
edm:end	-	-	-	◦	-	◦	-	-
edm:hasMet	◦	-	-	◦	-	-	-	-
edm:hasType	◦	-	-	-	-	-	-	-
edm:hasView	-	◦	-	-	-	-	-	-
edm:incorporates	◦	-	-	-	-	-	-	-
edm:intermediateProvider	-	•	-	-	-	-	-	-
edm:isDerivativeOf	◦	-	-	-	-	-	-	-
edm:isNextInSequence	•	-	◦	-	◦	◦	-	-
edm:isRelatedTo	◦	-	-	◦	-	-	-	-
edm:isRepresentationOf	◦	-	-	-	-	-	-	-
edm:isShownAt	-	3	-	-	-	-	-	-
edm:isShownBy	-	3	-	-	-	-	-	-
edm:isSimilarTo	◦	-	-	-	-	-	-	-
edm:isSuccessorOf	◦	-	-	-	-	-	-	-
edm:object	-	•	-	-	-	-	-	-
edm:provider	-	⋆	-	-	-	-	-	-
edm:realizes	◦	-	-	-	-	-	-	-
edm:rights	-	⋆	•	-	-	-	-	-
edm:type	⋆	-	-	-	-	-	-	-
edm:ugc	-	◦	-	-	-	-	-	-
owl:sameAs	◦	-	◦	◦	◦	◦	-	-
skos:prefLabel	-	-	-	•	•	•	•	-
skos:altLabel	-	-	-	◦	◦	◦	◦	-
skos:broader	-	-	-	-	-	-	◦	-
skos:narrower	-	-	-	-	-	-	◦	-
skos:related	-	-	-	-	-	-	◦	-
skos:broadMatch	-	-	-	-	-	-	◦	-
skos:narrowMatch	-	-	-	-	-	-	◦	-
skos:relatedMatch	-	-	-	-	-	-	◦	-
skos:exactMatch	-	-	-	-	-	-	◦	-
skos:closeMatch	-	-	-	-	-	-	◦	-
skos:notation	-	-	-	-	-	-	◦	-
skos:inScheme	-	-	-	-	-	-	◦	-
skos:note	-	-	-	◦	◦	◦	◦	-
foaf:name	-	-	-	◦	-	-	-	-

Table 1 (continued) | Properties for each class, indicating if they are mandatory (⋆), recommended (•), or optional (◦)

Property	edm:ProvidedCHO	ore:Aggregation	edm:WebResource	edm:Agent	edm:Place	edm:TimeSpan	skos:concept	cc:License
rdaGr2:biographicalInformation	-	-	-	◦	-	-	-	-
rdaGr2:dateOfBirth	-	-	-	•	-	-	-	-
rdaGr2:dateOfDeath	-	-	-	•	-	-	-	-
rdaGr2:dateOfEstablishment	-	-	-	◦	-	-	-	-
rdaGr2:dateOfTermination	-	-	-	◦	-	-	-	-
rdaGr2:gender	-	-	-	◦	-	-	-	-
rdaGr2:placeOfBirth	-	-	-	◦	-	-	-	-
rdaGr2:placeOfDeath	-	-	-	◦	-	-	-	-
rdaGr2:professionOrOccupation	-	-	-	◦	-	-	-	-
wgs84_pos:lat	-	-	-	-	•	-	-	-
wgs84_pos:long	-	-	-	-	•	-	-	-
wgs84_pos:alt	-	-	-	-	◦	-	-	-
odrl:inheritFrom	-	-	-	-	-	-	-	⋆
cc:deprecatedOn	-	-	-	-	-	-	-	◦

Numbers indicate that adding at least one of the properties with the same number is mandatory, while the others remain recommended.



Fig. 2 | Official sign issued by the municipality of Delft instructing shoppers to maintain a social distance of at least 1.5 m. It depicts an example of local handcraft work: that of the kissing couple made in ceramic, with the figures separated by the message «Keep 1.5m distance» in Dutch. At the bottom it reads «Healthy together in Delft» and «delft.com».

perspectives on the same cultural heritage artefact. To this end, we categorise EDM metadata properties to differentiate between those that describe an entity and those that situate it within a broader corpus. This distinction is crucial, as it allows communities to contribute knowledge about their heritage while enabling memory institutions to retain custody of records for management and preservation. This approach serves two primary purposes: to enable custodians to maintain the integrity of community contributions without compromising the FAIR properties of the metadata and to facilitate the coexistence of institutional taxonomies and community folksonomies in an independent yet symbiotic manner.

We then define the roles of the agents involved, partly borrowing from the theoretical framework of citizen curation. We determine user requirements to the practical context of metadata creation. The entire model is synthesised into a five-stage process (recording, submission, curation, release, use/reuse). For each stage, we specify the role requirements and the corresponding metadata to be added, thereby creating a clear bridge between the actors, their tasks, and the final metadata record.

Diversity: conciliating folksonomies and taxonomies in the EDM

The use of common terminology has long been a cornerstone of data sharing and interoperability. In cultural heritage, as in many other fields, a significant gap exists between how experts and non-experts conceptualise and describe the same thing. Consequently, authoritative narratives and data remain codified in the language of specialists, leading to a loss of richness in the expression of heritage values and interpretations[42,43].

This divergence has been widely recognised in crowdsourcing initiatives enabling individuals to contribute comments and information about objects exhibited or held by museums. In contrast to the formal taxonomies that describe assets using specialised vocabularies, the term «folksonomy» was coined to name crowdsourced contributions. A folksonomy is not a top-down strategy but rather a widespread practice facilitated by websites and crowdsourcing platforms. The term, coined by Thomas Vander Wal, blends «folk» and «taxonomy» to describe the result of personal, free tagging of information for one's own retrieval[44].

Key characteristics of folksonomies, which distinguish them from formal taxonomies, include a flat, non-hierarchical structure; the use of uncontrolled, community-sourced vocabulary; greater flexibility; and a user-centric scope. This approach is inherently prone to polysemy

Table 2 | Functional requirements and roles of an ecosystem for citizen curation

Community (owner & end user)	Custodian	Builder	Ecosystem
R1 community:express	R6 custodian:manage	R18 builder:interop	R30 eco:link
R2 community:rights	R7 custodian:curate	R19 builder:request	R31 eco:consume
R3 community:negotiate	R8 custodian:publish	R20 builder:viewterms	R32 eco:collect
R4 community:know	R9 custodian:visibility	R21 builder:explain	R33 eco:identity
R5 community:claim	R10 custodian:copyright	R22 builder:standards	R34 eco:entity
R23 community:explore	R11 custodian:enable		R35 eco:agency
R24 community:share	R12 custodian:access		R36 eco:manage
R25 community:compare	R13 custodian:viewpoints		R37 eco:extend
R26 community:enhance	R14 custodian:monitor		
R27 community:privacy	R15 custodian:inappropiate		
R28 community:copyright	R16 custodian:link		
R29 community:claim	R17 custodian:usage		

Based on the roles proposed by Daga et al.[23].

(exhibiting multiple meanings) and a degree of imprecision. Consequently, freely contributed public content is often disregarded as a reliable source of metadata. However, empirical evidence from memory institutions demonstrates that a lack of synonym control, polysemy, and even misspelt entries are equally prevalent in metadata created by experts[7].

As Seth van Holland[45] warned, while folksonomies cannot replace cataloguing and indexing, they hold significant semantic value. As part of a conceptual shift that envisions metadata serving both communities and specialists, leveraging folksonomies could foster mutual understanding rather than perpetuate an *indoctrinating* approach. This entails envisioning metadata as a shared space where taxonomies and folksonomies are mutually supportive: taxonomies provide a non-local context for folksonomies, while folksonomies ground specialised language with sense and relevance within local settings. The ability of the EDM to incorporate alternative names for concepts (e.g., a title and an alternative title), alongside its use of contextual classes, provides an outstanding opportunity for this coexistence.

Stock[46] suggests that establishing a symbiotic combination of taxonomies and folksonomies may represent a robust mechanism for bridging fields, knowledge domains, and communities. While many efforts to develop shared terminology take a top-down approach, with experts promoting official knowledge organisation systems, folksonomies facilitate bottom-up approaches where language develops organically and can serve more diverse purposes[47]. Rather than excluding professional involvement, this approach requires new skills for engaging both experts and communities. These competencies, founded on a willingness for mutual understanding, are essential for the effective digital curation of information. For example, De Oliveira Souza et al.[48] argues that folksonomies can form the cornerstone of local thesauri, embedding value and diversity into sources of knowledge.

The proposed framework is built on the premise that community-generated folksonomies and expert taxonomies should coexist to provide complementary or alternative interpretations. This synergy is technically enabled by EDM semantic web properties, such as `owl:sameAs`, which can formally link different conceptualisations of the same core entity such as a cultural object, a digital resource, or a concept like a place or time period. Furthermore, other EDM properties allow for varying levels of descriptive detail, where the most granular data units are often expressed by using taxonomies.

A concrete example of this is the use of the `dc:coverage` property. A community may describe a photograph's location as «the main square in our village in 1950». Later, a museum expert can add a more formal, structured description using official records, like the coordinates, the historical name of the square, and the date in EDM format (`dcterms:spatial` and `dcterms:temporal`). The expert's work makes the data easier to search and link to other records, but it doesn't erase the community's original, personal description. Both versions are kept, creating a richer, layered history.

Furthermore, properties from the `dcterms:` and `edm:` namespaces allow contributions to be situated within the archive's broader context. This supports metadata versioning, effectively creating a layered «palimpsest» of descriptions for the same object. Finally, the `skos:Concept` class acts as a bridge, connecting informal tags to formal knowledge sources. Overall, this EDM-based framework allows communities to represent their knowledge without being overridden by institutional cataloguing practices.

Roles

Successful co-creation requires the careful coordination of contributions from both communities and experts. In this context, communities must be able to express knowledge about their heritage according to their own communicative resources and needs. Meanwhile, memory institutions, acting as co-authors of collective knowledge, should be able to perform tasks such as data management and contributing within established frameworks of taxonomies and controlled vocabularies. Beyond communities and memory institutions, a central actor in this process is the environment (digital platform) in which interactions such as data capture, submission, management, and presentation are enabled, along with the individuals or institutions behind these developments.

The proposed role structure consolidates the «Owner» and «End User» categories identified by Daga et al.[23] into a single, more comprehensive «Community» one. This reflects the rationale of the participatory model, wherein communities are expected to both create new metadata (thus owning their contributions) and interact with existing records, thereby embodying both functions simultaneously. Consequently, this unified «Community» role inherited the combined requirements of the original Owner and End User roles. The requirements of this role, along with those of the Custodian, Builder, and Ecosystem roles, is detailed in Table 2.

Communities play the role of Owner when creating metadata. At this stage, owners must be able to `express` the terms, conditions, and `rights` associated with the metadata, such as its licence. Furthermore, owners should be able to `know` how their creation is used, `claim` its ownership, and `negotiate` digital rights with third parties.

The Custodian represents an agent responsible for ingesting and preserving metadata. Custodians must therefore `manage` this metadata within a digital archive. A key duty is to `curate` the content of citizen-produced metadata according to a given mission, institutional needs, and quality standards. They should integrate authoritative metadata alongside community `viewpoints` and `monitor` for potential issues related to

copyright and inappropriate content (e.g., hate speech, use of discriminatory language, etc). Ultimately, custodians could link their repositories to larger networks, thereby tracing the usage of the data. Custodians interact with third parties and publish the repository, managing its visibility, which includes the provision of open data. In this capacity, custodians declare the copyright of the metadata, enable access requests, and manage access controls.

Builders are agents who utilise metadata to develop their own applications, thus interacting with the custodian. Interoperability (interop) is key in this context, so builders must be able to request data from the custodian. Similarly, builders should favour the use of solutions based on open standards. Builders should be able to review the terms of use of the assets (viewterms) and receive an explanation (explain) when access is denied.

Finally, Communities also enact as End Users when interacting with pre-existing metadata, often with the potential to enrich it. Consequently, the same person may act as an Owner and an End User at different stages. To facilitate this, end users must be able to explore the contents of a digital archive. These users should be able to share digital materials with others, expressing their own viewpoints and opinions, compare productions from different sources, and enhance digital content through annotations and additions, such as personal perspectives. To ensure safe digital environments, End Users should be supported in recognising whether shared content could pose a threat to privacy. Furthermore, end users must be made aware of copyright terms and conditions and should be able to claim ownership and intellectual property on their contributions.

Stages of the process of community-driven metadata co-creation

The process is divided into five principal stages: the recording of a cultural asset or addition of new information to an existing metadata file, its submission to a digital environment, the curation of its content, its public release, and finally, a final stage dedicated to its use and reuse. Within this model, we identify the roles involved, the metadata-related inputs, and the operations performed on the content at each stage.

Recording

The initial stage takes place when a metadata file is created or enriched through community interactions with cultural heritage and/or existing metadata. This involves both the cultural heritage manifestation itself (the edm:ProvidedCHO) and the digital file that represents it (the edm:WebResource). These interactions include exploring the metadata archive, sharing its contents, comparing records, and asserting ownership or raising other issues. Communities may also decide on privacy and rights matters related to their contributions. At this stage, the ecosystem must be capable of ingesting external data (e.g., the files pertaining to edm:WebResource). To define the metadata properties for the Community role, we first considered the entire set of EDM properties (Table 2). A subset was then systematically refined by subtracting properties according to the following criteria:

- Properties that explicitly require controlled taxonomies, provided another property can capture the same information as a folksonomy. For instance, while dc:type admits any input, edm:hasType requires a controlled vocabulary, yet both express the same concept.
- The entire edm:TimeSpan class, as all its properties are either relational or require taxonomic terms. The original source information, however, is still linked as a folksonomy via the property dcterms:temporal.
- Properties related to data management and visibility, as these are specific to the custodian's requirements (e.g., the entire ore:Aggregation class).
- Properties used for labelling within contextual classes, which are essential for maintaining the relational integrity of the archive (e.g., owl:sameAs).
- The entire skos:Concept class, as its purpose is to link the Cultural Heritage Object (CHO) with authoritative and formal sources, a responsibility that falls to the custodian.

The resulting set of properties available for addition at this stage is presented in Table 3.

Submission

This stage entails the formal submission of the recorded information to a custodian. The digital ecosystem must facilitate the processing of the data collected during the initial recording phase. Crucially, the contributor must be able to specify the terms and conditions governing the use and management of their contribution. Furthermore, the ecosystem should support the contributor in negotiating, understanding, and, where applicable, asserting their corresponding rights. At this point, the only additional metadata properties are those relating to the rights of the edm:ProvidedCHO, the edm:WebResource, and the ore:Aggregation (see Table 4).

Curation

At this stage, the custodian manages the submitted contribution within the repository. Acting as the party responsible for ensuring content adheres to quality standards and established participatory guidelines, the custodian performs a curation process. This involves integrating diverse perspectives on a single entity by linking the new submission to relevant pre-existing records. This is primarily achieved using EDM properties from the ore:Aggregation class, as well as dcterms: and edm: properties within the edm:ProvidedCHO and edm:WebResource classes.

A further critical requirement is the monitoring of potential issues relating to copyright infringement or privacy law, and the filtering of inappropriate or socially harmful content. Ethical guidelines for managing sensitive heritage data are exemplified by the Open Heritage Statement[49]. The practices of memory institutions in this area have been critically analysed by Vezina et al.[50]. Furthermore, a significant ethical framework was developed as part of the Nuohtti project to guide the accessibility of digital data concerning Indigenous cultural heritage[51].

The custodian enriches the contribution by connecting it to relevant external knowledge sources using the contextual class skos:Concept. A key principle of this process is that enrichment occurs without overwriting or erasing the original expressions and content provided during the Recording stage. In view of the above, the data to be incorporated to the metadata file at this stage is that presented in Table 5.

Release

Once the content has been curated, the custodian can make it available through a public archive. This actor is responsible for managing the visibility of the metadata and declaring the corresponding intellectual property rights. At this stage, the metadata must be associated with a licence, which involves the use of the cc:License class, as specified in Table 6. The ecosystem must also be capable of linking the archive to external resources, such as those designed to increase public accessibility.

Use and reuse

Metadata is meant to make knowledge about heritage FAIR. To leverage this, the custodian must be able to enable third parties to request access to the data. Beyond merely managing data access, the custodian should also be capable of monitoring and tracing its usage. Agents interested in using this data within their own developments (e.g., interactive apps, virtual reality environments, augmented reality interfaces, etc.) should engage in interoperable data management with the custodian, be able to request permission for its use, and review the applicable terms. Whenever access is not granted, a clear and fair explanation must be provided. In all cases, builders should be committed to design solutions based on open standards.

Table 3 | Properties for each class, indicating if they are mandatory (⋆), recommended (•), or optional (◦)

Class or property	edm:ProvidedCHO	edm:WebResource	edm:Agent	edm:Place
dc:contributor	•	-	-	-
dc:coverage (or dcterms:spatial & dcterms:temporal)	•	-	-	-
dc:creator	◦	◦	-	-
dc:date	•	-	◦	-
dc:description	1	◦	-	-
dc:format	◦	-	-	-
dc:language	⋆	-	-	-
dc:publisher	•	-	-	-
dc:source	•	◦	-	-
dc:subject	2	-	-	-
dc:title	1	-	-	-
dc:type	2	-	-	-
dcterms:created	•	◦	-	-
dcterms:extent	◦	-	-	-
dcterms:issued	•	-	-	-
dcterms:medium	◦	-	-	-
dcterms:provenance	◦	-	-	-
dcterms:spatial	2	-	-	-
dcterms:tableOfContents	◦	-	-	-
dcterms:temporal	2	-	-	-
foaf:name	-	-	◦	-
rdaGr2:biographicalInformation	-	-	◦	-
rdaGr2:dateOfBirth	-	-	•	-
rdaGr2:dateOfDeath	-	-	•	-
rdaGr2:dateOfEstablishment	-	-	◦	-
rdaGr2:dateOfTermination	-	-	◦	-
rdaGr2:gender	-	-	◦	-
rdaGr2:placeOfBirth	-	-	◦	-
rdaGr2:placeOfDeath	-	-	◦	-
rdaGr2:professionOrOccupation	-	-	◦	-
wgs84_pos:lat	-	-	-	•
wgs84_pos:long	-	-	-	•
wgs84_pos:alt	-	-	-	◦

Numbers indicate that adding at least one of the properties with the same number is mandatory, while the others remain recommended.

Table 4 | Properties for each class, which at this stage are all optional (◦)

Class or Property	edm:ProvidedCHO	ore:Aggregation	edm:WebResource
dc:rights	◦	◦	◦

Communities must be able to engage with the data in a manner that fulfils the requirements envisioned in the recording stage. It is important to note that we conceive the interaction with pre-existing data as a potentially creative act. The interpretations and experiences arising from the archive open a channel for new contributions and enrichment, fostering a continuous cycle of engagement.

We propose that any enrichment or addition to a pre-existing entry should give rise to a new version, rather than overwriting the original record. The inherent capacity of the EDM to support multiple records referring to the same entity facilitates the conception of metadata archives as manifestations of a cultural palimpsest. These archives are rooted in heritage assets yet transient in time, thereby tracking how people valorise and manage cultural environments across different periods.

Stages and sequential compliance of the FAIR principles

The FAIR Guiding Principles[52] propose the criteria that data and metadata should meet in order to be considered Findable, Accessible, Interoperable, and Reusable. Since our framework conceives the metadata as a procedural co-creation rather than an individual production, the FAIR Guiding Principles (see Table 7) can be mapped within the five stages described above.

The Recording stage primarily consists of ingesting data provided by communities, namely folksonomic content that expresses knowledge and values associated with heritage manifestations. As such, this step contributes to the Findability and Reusability of metadata, in line with the FAIR Guiding Principles F2 and R1.

Once submitted, this data is registered in a resource accessible to the Custodian. During the Curation stage, the Custodian assigns a globally unique and persistent identifier linked to the CHO (F1, F3), and adds qualified references to other metadata (I3) as well as domain-relevant community standards (R1.3). Furthermore, at this stage, metadata is enriched with detailed provenance information (R1.2). The metadata is then registered in a searchable resource (F4) that can be accessed by using its identifier (A1). Usage licenses (R1.1) are added after the Release of metadata.

Table 5 | Properties for each class, indicating if they are mandatory (⋆), recommended (•), or optional (◦)

Class or property	edm:ProvidedCHO	ore:Aggregation	edm:WebResource	edm:Agent	edm:Place	edm:TimeSpan	skos:concept
dc:format	-	-	◦	-			
dc:identifier	•	-	-	◦	-	-	-
dc:relation	◦	-	-	-	-	-	-
dcterms:extent	-	-	◦	-			
dcterms:alternative	•	-	-	-	-	-	-
dcterms:conformsTo	◦	-	◦	-	-	-	-
dcterms:hasFormat	◦	-	-	-	-	-	-
dcterms:hasPart	◦	-	◦	◦	◦	◦	-
dcterms:hasVersion	◦	-	-	-	-	-	-
dcterms:isFormatOf	◦	-	◦	-	-	-	-
dcterms:isPartOf	•	-	◦	◦	◦	◦	-
dcterms:isReferencedBy	◦	-	◦	-	-	-	-
dcterms:isReplacedBy	◦	-	-	-	-	-	-
dcterms:isRequiredBy	◦	-	-	-	-	-	-
dcterms:isVersionOf	◦	-	-	-	-	-	-
dcterms:references	◦	-	-	-	-	-	-
dcterms:replaces	◦	-	-	-	-	-	-
dcterms:requires	◦	-	-	-	-	-	-
edm:aggregatedCHO	-	⋆	-	-	-	-	-
edm:begin	-	-	-	◦	-	◦	-
edm:currentLocation	◦	-	-	-	-	-	-
edm:dataProvider	-	⋆	-	-	-	-	-
edm:end	-	-	-	◦	-	◦	-
edm:hasMet	◦	-	-	◦	-	-	-
edm:hasType	◦	-	-	-	-	-	-
edm:hasView	-	◦	-	-	-	-	-
edm:incorporates	◦	-	-	-	-	-	-
edm:intermediateProvider	-	•	-	-	-	-	-
edm:isDerivativeOf	◦	-	-	-	-	-	-
edm:isNextInSequence	•	-	◦	-	◦	◦	-
edm:isRelatedTo	◦	-	-	◦	-	-	-
edm:isRepresentationOf	◦	-	-	-	-	-	-
edm:isShownAt	-	3	-	-	-	-	-
edm:isShownBy	-	3	-	-	-	-	-
edm:isSimilarTo	◦	-	-	-	-	-	-
edm:isSuccessorOf	◦	-	-	-	-	-	-
edm:object	-	•	-	-	-	-	-
edm:provider	-	⋆	-	-	-	-	-
edm:realizes	◦	-	-	-	-	-	-
edm:rights	-	⋆	•	-	-	-	-
edm:type	⋆	-	-	-	-	-	-
edm:ugc	-	◦	-	-	-	-	-
owl:sameAs	◦	-	◦	◦	◦	◦	-
skos:prefLabel	-	-	-	•	•	•	•
skos:altLabel	-	-	-	◦	◦	◦	◦
skos:broader	-	-	-	-	-	-	◦
skos:narrower	-	-	-	-	-	-	◦
skos:related	-	-	-	-	-	-	◦
skos:broadMatch	-	-	-	-	-	-	◦
skos:narrowMatch	-	-	-	-	-	-	◦

Table 5 (continued) | Properties for each class, indicating if they are mandatory (⋆), recommended (•), or optional (◦)

Class or property	edm:ProvidedCHO	ore:Aggregation	edm:WebResource	edm:Agent	edm:Place	edm:TimeSpan	skos:concept
skos:relatedMatch	-	-	-	-	-	-	◦
skos:exactMatch	-	-	-	-	-	-	◦
skos:closeMatch	-	-	-	-	-	-	◦
skos:notation	-	-	-	-	-	-	◦
skos:inScheme	-	-	-	-	-	-	◦
skos:note	-	-	-	◦	◦	◦	◦

Numbers indicate that adding at least one of the properties with the same number is mandatory, while the others remain recommended.

Table 6 | Properties for each class, indicating if they are mandatory (⋆) or optional (◦)

Class or property	edm:WebResource	cc:License
odrl:inheritFrom		⋆
cc:deprecatedOn		◦
dcterms:issued	◦	

Accessibility is granted during the Use and Reuse stage, as metadata becomes retrievable through an open, free, and universally implementable communications protocol (A1.1, A1.2), accessible even if the CHO or its digital representations are no longer available (A2). During this stage and the Recording stage, data remains interoperable through the use of both folksonomies and taxonomies (I1, I2). This correspondence is also represented in Fig. 3

Structure of a metadata record

A metadata record of cultural heritage produced by means of this process can be approached through its classes and properties, the stages, the roles of contributors, and the nature of the data (i.e., taxonomies and folksonomies), as per Fig. 4. This diagram illustrates how a FAIR-compliant metadata record is achieved through the addition of data throughout the stages described above.

The different classes that constitute the EDM metadata model are increasingly added following a sequence that starts with the knowledge contributed by the community, relying on folksonomic content (orange) and its intellectual rights (green). This is then enriched with data based on controlled vocabularies (taxonomies, in red) and semantic links (blue) that will situate the contribution within a given corpus, thus ensuring interoperability. For example, by using the `ore:Aggregation` class to establish the links between objects, their digital representations, and the class `skos:Concept` to link the metadata with external contextual resources. Finally, the licence (green) represents the final addition prior to the use of the metadata file. Each box specifies the total number of properties it contains. A further breakdown, provided in brackets, details the quantity of mandatory, recommended, and optional EDM properties. The arrows illustrate the derivation of taxonomies from information originally expressed as folksonomies.

Example of procedure

The framework is intentionally general to support adaptation to different contexts. Its processes are organised around roles that do not presuppose specialised expertise; instead, any participant able to contribute a specific input, such as a metadata component, may assume them. This implies that the framework can be redesigned, or co-designed, to suit each community's characteristics and needs. As a guiding principle, diverse implementations are expected to produce FAIR metadata, with standardisation emerging at the end of the process rather than at the outset. The following is a hypothetical example, partly grounded in reality. The institution and its background are real, while its adoption of the framework presented here is intended to illustrate a possible implementation.

Table 7 | Mapping of the processes considered in the process of community-driven co-cration of metadata and the FAIR guiding principles[52]

	FAIR guiding principles[52]
Findability	
F1	(meta)data are assigned a globally unique and persistent identifier
F2	data are described with rich metadata (defined by R1 below)
F3	metadata clearly and explicitly include the identifier of the data it describes
F4	(meta)data are registered or indexed in a searchable resource
Accessibility	
A1	(meta)data are retrievable by their identifier using a standardized communications protocol
A1.1	the protocol is open, free, and universally implementable
A1.2	the protocol allows for an authentication and authorization procedure, where necessary)
A2	metadata are accessible, even when the data are no longer available
Interoperability	
I1	(meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation
I2	(meta)data use vocabularies that follow FAIR principles
I3	(meta)data include qualified references to other (meta)data
Reusability	
R1	meta(data) are richly described with a plurality of accurate and relevant attributes
R1.1.	(meta)data are released with a clear and accessible data usage license
R1.2.	(meta)data are associated with detailed provenance
R1.3.	(meta)data meet domain-relevant community standards)

The Museu da Luz is located in the Alentejo region of Portugal, in the village of Luz. The present-day settlement was constructed in the late 1990s and completed in 2002 to accommodate the inhabitants of the original village of Luz, which was deliberately flooded following the construction of the Alqueva Dam. The museum is an interpretative space conceived as a compensatory measure, designed to support the transition from the old to the new village and to help preserve memories and meanings, thereby providing a stronger foundation for the social reconstruction of the community.

The museum's collection is rich in everyday objects typical of the daily life of Luz's inhabitants, including tools related to agricultural activities and domestic items. The museum runs an activity entitled Object of the Month, which highlights a single item from the collection as the focus of discussions,

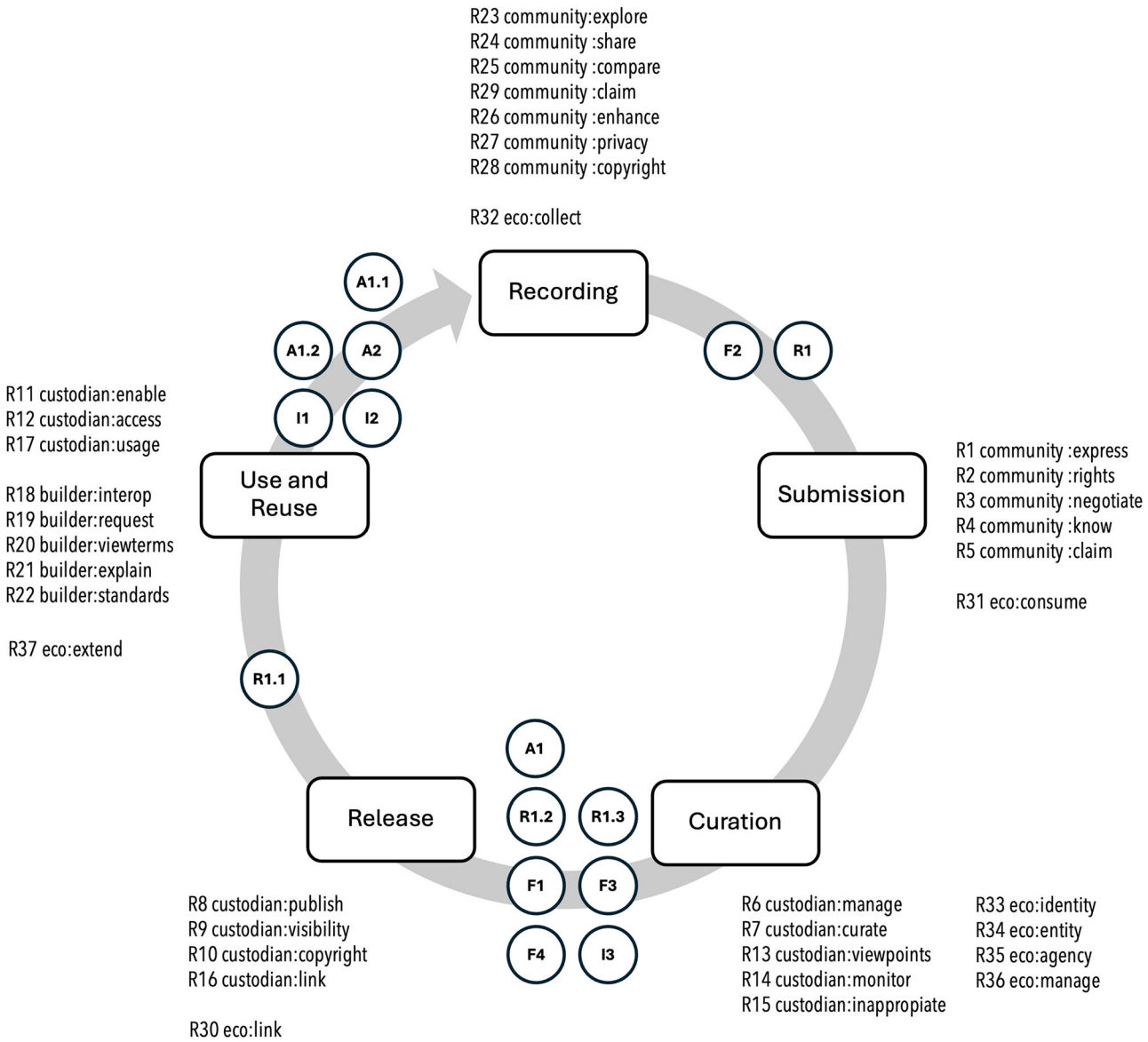

Fig. 3 | Integrated schematisation of the process of community-driven creation of metadata for cultural heritage, representing the stages of the process, the roles and requirements involved per stage, and the FAIR requirements achieved at the end of each stage. The direction of the process is indicated by the grey circular arrow; the stages are represented in rectangular boxes; roles and requirements are placed next to the relevant stage; and FAIR requirements are shown inside circles, as per Table 7.

workshops, and other activities. The museum's context, its emphasis on everyday life, and the practice of focusing on one object per month create favourable conditions for the co-creation of metadata using the community's knowledge and memories.

The community is invited to describe the Object of the Month (an alarm clock in April 2026[53]) in their own words, expressing what they know and remember about it. This recording stage could take place via a dedicated section of the museum's website, where users may submit text and digital files. While contributors are not expected to provide extensive information, the webpage would offer guidance on the types of details that may be included (see Table 3), such as the object's name, type, description, creator, date and place of origin, materials, and so forth. When contributions include digital resources, additional information about the files can be added.

The submitted information is then received by the museum. At the point of submission, contributors select the rights governing the use and management of their contributions before proceeding to the curation stage. Curation is carried out by the custodian, in this case, the individual responsible for the museum's catalogue. The custodian ensures that the contribution complies with institutional guidelines and does not contain inappropriate content. They then organise and manage the community-contributed information within the metadata model, while preserving the original folksonomies. Subsequently, the custodian integrates data already held in the museum's archive, based on taxonomies and controlled vocabularies (see Table 5). In doing so, they establish links between taxonomies and folksonomies.

Rather than relying on a single record, each object is associated with multiple metadata records that share the same taxonomies while reflecting a diversity of folksonomic inputs. Moreover, the inclusion of expert concepts enables the integration of metadata within the museum's broader corpus, facilitating semantic links to related entities (for example, similar clocks in other collections). Once the metadata record has been enriched with these inputs, it can be published with the appropriate licence and intellectual property rights, making it accessible online and marking the release stage.

The use and reuse stage occurs when communities and builder agents interact with the data. For instance, a developer might create an augmented reality interface for the museum, allowing both local and remote visitors to explore and share information related to the Object of the Month more effectively. Such developers should establish interoperable mechanisms for

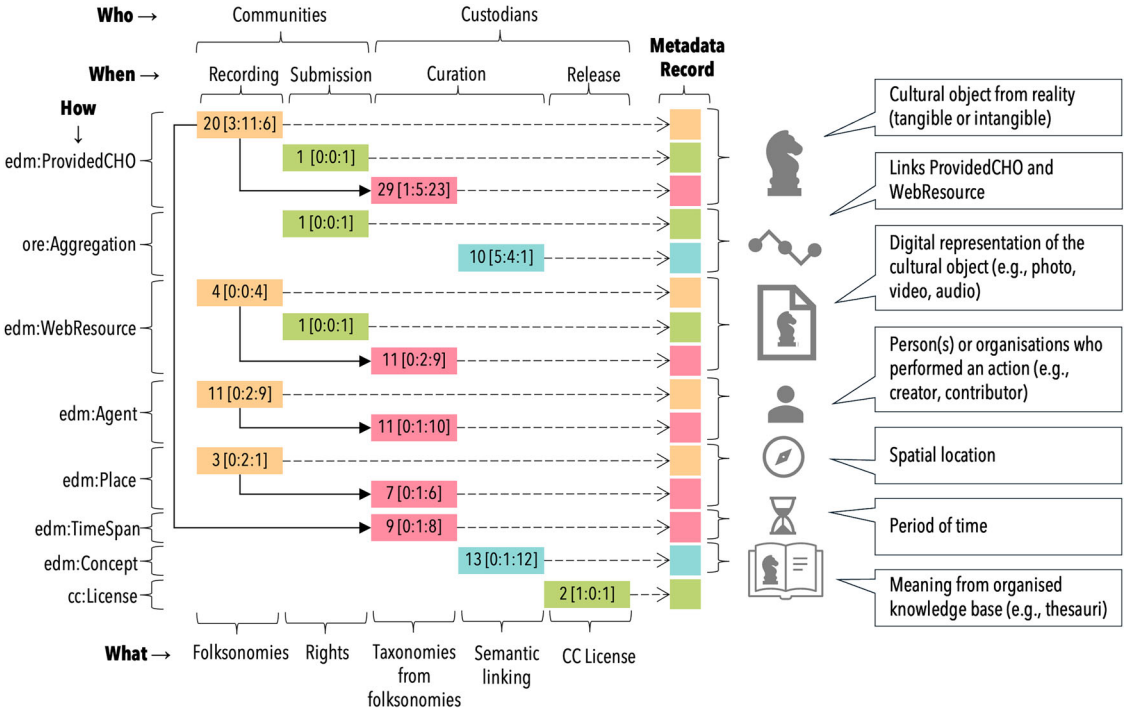

Fig. 4 | Constitution of a metadata record framed in the context of the contributors' roles, process stages, EDM classes, and types of content. Rectangular boxes indicate the number of EDM properties and, inside the brackets, a breakdown indicating the number of those that are mandatory, recommended, and optional, respectively. Fields belonging to folksonomic contents, rights, taxonomies derived from folksonomies, and semantic linking data are respectively presented in orange, green, red, and blue. Metadata fields are also grouped by the agent involved in their creation and stage of contribution (top), the EDM class to which they belong (left), and the type of content (bottom). On the right is a description of what each class of the EDM addresses.

accessing and using the data, preferably through open standards. In addition, any applications built on this data should support the export of metadata in standard formats. Following these stages, the co-created metadata for the Object of the Month would meet the criteria required to be considered FAIR.

Algorithm for a community contribution to FAIR metadata of everyday heritage

This subsection presents in more detail the data collection process for the Recording stage, focusing specifically on the properties to be supplied by communities. The properties fulfilled in subsequent stages by custodians, which depend on internal data management procedures and established institutional practices for curation and publication, are not addressed here.

The metadata properties designated for this stage (see Table 3) are designed for users to populate using their own expressions and resources, thereby generating folksonomies. The custodian's subsequent role is to establish correspondences between these user-generated folksonomies and established controlled vocabularies during the curation stage. Table 8 outlines the properties to be provided by the community, along with descriptive guidance to facilitate their interpretation in accordance with the EDM Mapping Guidelines[29].

To facilitate a more narrative-oriented organisation of the properties, these have been structured into a flowchart (see Fig. 5). Furthermore, the contextual classes `edm:Place` and `edm:Agent` are elaborated upon, illustrating how they can be derived from properties such as `dc:creator` and `dc:contributor`. This flowchart is agnostic, as it does not compromises with specific interfaces or procedures for data registration.

Proof-of-concept survey community knowledge about cultural heritage

The proof-of-concept experiment aimed to validate the feasibility of community participation in the first stage of the process of community-driven creation of metadata for cultural heritage. To do so, we investigated which types of cultural heritage manifestations different individuals provide when prompted with the idea of 'everyday heritage', and sought to understand how contributors articulate their knowledge. The study involved the design and publication of an online survey that remained open for contributions over a period of two months. The responses provided by contributors are then presented and discussed in light of these community contributions within the process of community-driven creation of metadata for cultural heritage.

We asked people to share their knowledge about a manifestation of cultural heritage of personal, familial, or local significance. For this purpose, the fields presented in Fig. 5 were captured and organised as questions with free-text answers. We developed it on the EUSurvey platform, which provides public, free, and fully open-source hosting. The survey and the data collected were deposited in an ad-hoc public repository[54].

The survey structure comprised:

- A preamble disclosing the context of the research project, its motivation, the scope of applicability (i.e., a general definition of Cultural Heritage Object), and instructions for using the survey.
- A field asking participants to submit their contributions under the CC0 1.0 Universal Creative Commons licence[55]. Submission was enabled upon acceptance of these terms.
- Fields addressing the Cultural Heritage Object (18 in total).
- A button to upload one multimedia file (Web Resource).
- Fields addressing the Web Resource (3 in total).
- A text box for contributing as free text (rather than guided by the survey) or for including information beyond the scope of the suggested fields.

While the survey and its accompanying explanations were originally written in English, the platform allows for automatic translation into the 24 official languages of the European Union. We therefore made the survey available in Spanish, Italian, French, and Portuguese using this functionality.

Table 8 | Properties and user guidance for providing information at the recording stage

Class or Property	Guidance for the User
edm:providedCHO	This information describes the heritage object you are documenting
dc:title	Provide a name for the CHO you are sharing.
dc:description	Describe the CHO.
dc:language	In which language is the CHO expressed?
dc:creator	Who created the CHO?*
dc:contributor	Who else made a significant contribution to the CHO?*
dc:type	What type of object is the CHO?
dc:format	What is the specific form of the CHO?
dcterm:extent	What are the dimensions or duration of the CHO?
dcterms:medium	What materials is the CHO made from?
dcterms:provenance	Who has previously owned or been the custodian of the CHO?
dc:source	What is the source or origin of the CHO?
dc:subject	What subject or topic does the CHO relate to?
dcterms:created	When was the CHO created?
dc:date	Provide a significant date in the history of the CHO.
dcterms:coverage	Which location and time period does the CHO represent?
dc:temporal	Which time period corresponds to the coverage of the CHO?
dc:spatial	Which place is represented by, or associated with, the CHO?
dc:publisher	Which organisation published this printed material?
dcterms:issued	When was this printed material published?
dcterms:tableOfContents	What is the structure or contents of this printed material?
edm:WebResource	*This information describes the digital file you are sharing.*
dc:creator	Who created this Web Resource?*
dc:description	Describe the Web Resource.
dcterms:created	When was the Web Resource created?
edm:Agent	**This information further describes a Person or Organisation.*
foaf:name	What is the name of the person/organisation?
rdaGr2:biographicalInformation	What biographical information is known about this person?
rdaGr2:dateOfBirth	When was this person born?
rdaGr2:dateOfDeath	When did this person die?
rdaGr2:dateOfEstablishment	When was this organisation established?
rdaGr2:dateOfTermination	When was this organisation dissolved?
rdaGr2:gender	What is the gender of the person?
rdaGr2:placeOfBirth	Where was this person born?
rdaGr2:placeOfDeath	Where did this person die?
rdaGr2:professionOrOccupation	What is the primary profession or occupation of this agent?
edm:Place	*This information further describes a spatial location.*
wgs84_pos:lat	Latitude coordinate in the WGS84 system.
wgs84_pos:long	Longitude coordinate in the WGS84 system.
wgs84_pos:alt	Altitude coordinate in the WGS84 system.

The symbol * indicates that it is a compulsory field.

The quality of these translated versions was reviewed by native speakers and manually edited where necessary.

Each field of the survey was accompanied by a brief explanation of its meaning (see Table 7). To avoid a prescriptive or formulaic approach, participants were also invited to submit their contributions by using a single and larger text box rather than completing the individual fields.

The survey was disseminated through various channels, including local mailing lists at the authors' affiliated institutions, mailing lists of professional associations, and personal invitations extended to family members and acquaintances. The survey was open to anyone willing to contribute and preserved participants' anonymity by not requiring registration or storing any identifying information.

Results

Contributions submitted via the survey

A total of 49 contributions were recorded. Of these, 17 were submitted in Spanish, five in Italian, two in Portuguese, and the remaining 25 in English. The contribution period was open from 27 November 2025 to 20 February 2026. According to the survey log, four additional contributions were saved as drafts and subsequently withdrawn. The timeline of contributions is shown in Fig. 6.

The contributions encompassed a wide variety of CHOs. There is a balance between tangible and intangible assets, representing 59.2% and 40.8% of the sample, respectively. As expected, contributions made extensive use of folksonomies. The field `dc:Type`, for example, demonstrates the variety found when expressing ideas that can be semantically similar. Table 9 shows the types of CHOs provided by contributors, alongside broader categorisations. According to these categories, the most recurrent were buildings (18.36%), food (16.32%), and practices (12.24%). Categories such as household items, landscape, music, and painting each accounted for four samples (8.16%). While this categorisation was based on the authors' criteria, it often overlaps with the types suggested by the contributors themselves.

The variety of CHOs indicates the wide spectrum of assets that individuals may regard as culturally and historically significant, including abstract concepts such as «*Respect for all living forms*» (sample 15) and «*Old-fashioned human interaction*» (sample 19). Contributions submitted in languages other than English included terms that cannot be directly translated, as they frequently refer to local names and expressions. Examples include the word «*Molcajete*» (sample 14), which denotes a specific type of Mexican traditional mortar; the Venetian oral expression «*Co le ciácoe no se fa fritóe*» (sample 18); and the «*Cachamorra de pastor*» (sample 31), a distinct type of walking stick used in Portugal. These cases exemplify the value and relevance of preserving the integrity of folksonomies in the creation of metadata, regardless of their subsequent pairing with more standardised, taxonomy-based systems and records.

Completeness of the metadata fields

We analysed the amount and type of information that participants contributed. The analysis was based on a categorisation of the responses, distinguishing between those that explicitly provide knowledge, express a lack of knowledge, indicate that the field is not applicable, or were simply left blank. Accordingly, we defined the following categories:

- Blank: The field was empty at the moment of submission.
- Not Applicable: The field contained expressions such as «*NA*», «*N/A*», Not Applicable, etc.
- Unknown: The field contained expressions such as Unknown (in any language) or «?».
- Valid: Any other response providing data.

We found that, on average, respondents completed about 65% of the survey, leaving approximately 30% of fields blank. Responses marked as Not Applicable or Unknown accounted for just over 5% of the total (see Table 10). We did not find significant differences between the levels of completion for tangible and intangible heritage, thereby supporting the hypothesis that community-led metadata recording constitutes a feasible means of revealing and representing diverse manifestations of both tangible and intangible cultural heritage.

However, explicit Not Applicable responses were concentrated in a few specific contributions, as shown in Fig. 7, mostly those related to intangible practices, such as «*Old-fashioned human interaction*» (sample 19), «*Saying in the Venetian dialect*» (sample 18), and «*Respect for all living forms*»

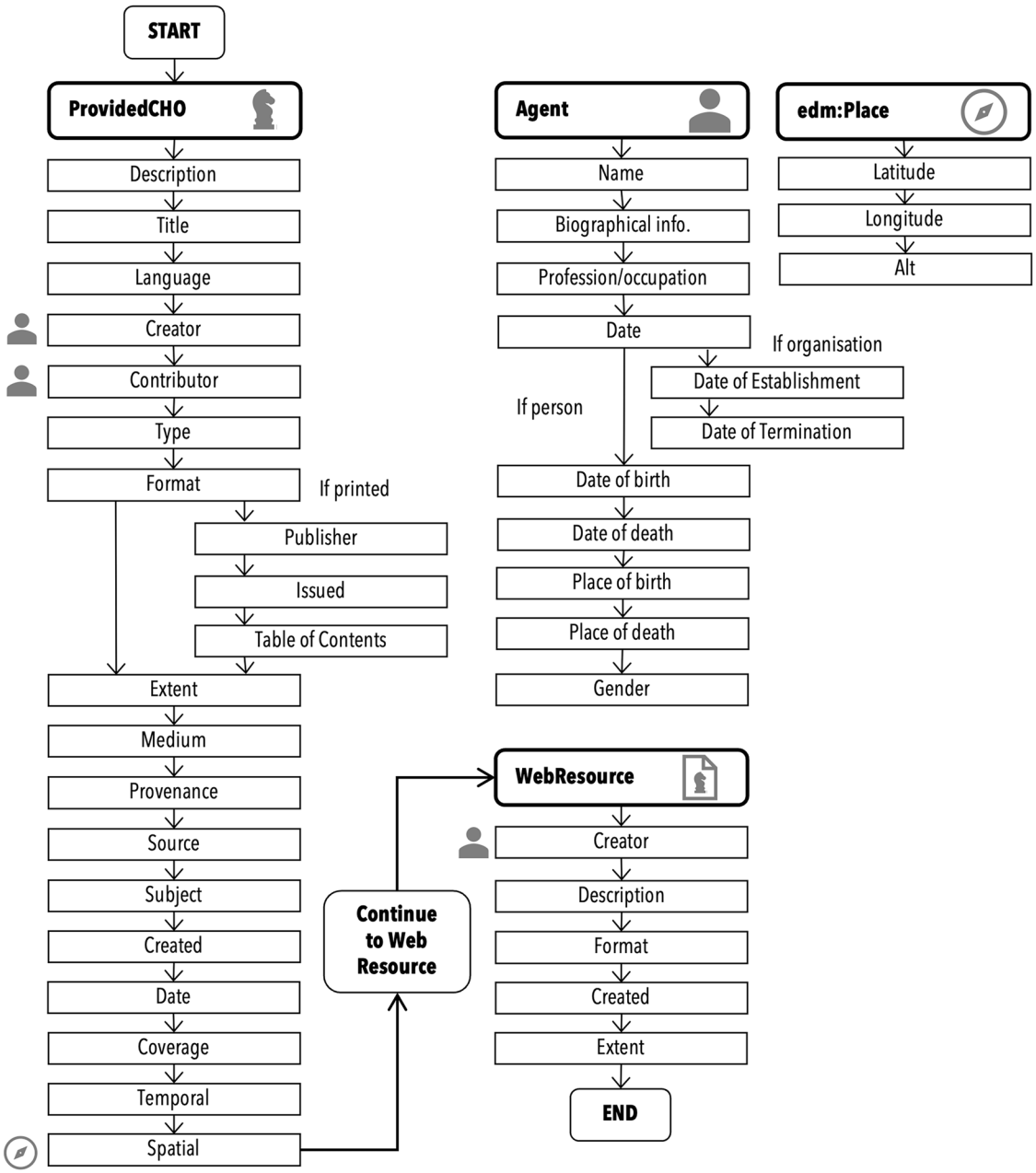

Fig. 5 | Flowchart representing the order in which the fields were structured to create the survey. Classes are represented by bold rectangles, whereas the other rectangles indicate single fields. Arrows represent the process from the start to the end of the survey. The icon of a person indicates that the field enables the user to provide more detailed information belonging to the class `Agent`. Similarly, the compass indicates that the field enables the user to address detailed information related to the class `edm:Place`.

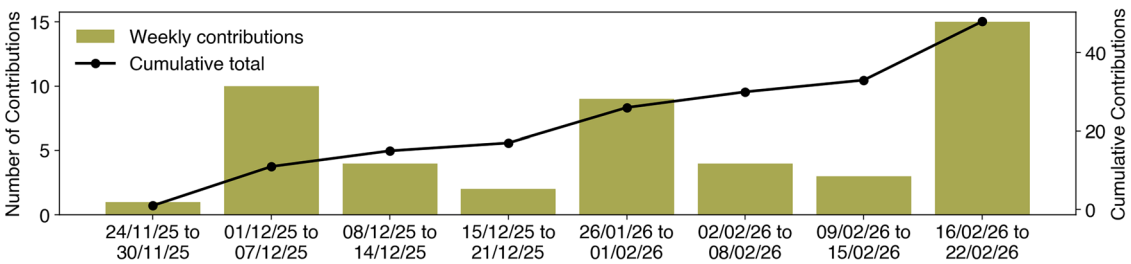

Fig. 6 | Survey contributions per week. The horizontal axis represents periods of time of one week, from 24 November 2025 up to 22 January 2026. The vertical axis represents the number of contributions received in each period. The black line represents the cumulative total of contributions.

Table 9 | Titles (`dc:title`) and types (`dc:type`) of CHOs, as recorded by contributors, together with the broad categories suggested by the authors

No.	Title of the CHO (dc:title)	Type (dc:type)	Category
1	Amsterdam and Delft canals and canal walls	A civil structure	Building
2	Bike culture in the Netherlands	–	Practice†
3	Paykan	Car	Vehicle
4	Enrosadira legend (Dolomites)	Oral/written transmitted legend	Oral tradition†
5	Pane cunzato	Traditional recipe	Food†
6	German Bread	Food	Food†
7	El baile de los que sobran	*Intangible heritage asset, music	Music†
8	Museo Nacional de Bellas Artes	*Heritage building and art museum	Building
9	Tostador de marraqueta chileno	*Utensil	Household item
10	Teatro Municipal de Santiago	*Heritage building and scenic arts theatre	Building
11	Rio Mapocho - Tramo Lo Barnechea (Plaza San Enrique)	*A river and natural space	Landscape
12	Parroquia Nuestra Señora de Andacollo	*Religious heritage building	Building
13	Pueblito Las Vizcachas	*Local culture, handcraft, communitary life	Landscape
14	Elaboración de salsas de molcajete y tortillas a mano	*Gastronomy	Food†
15	El respeto a todas las formas de vida	*A value	Practice†
16	Dengbej (the person and song)	A song	Music†
17	Tortillas	Food	Food†
18	Saying in Venetian Dialect: 'Co le ciàcoe no se fa fritoe'	Oral expression in Veneto dialect (traditional saying/proverb)	Oral tradition†
19	Old fashioned human interaction (especially with strangers)	–	Practice†
20	Oostpoort (Delft)	A building	Building
21	Chiesetta di S.Maria degli Angeli	*Building of popular cult	Building
22	Antica fiera di Cavalcaselle	*A popular manifestation	Practice†
23	Ricetta del dolce 'Tiramisù'	-	Food†
24	Parco regionale del fiume Mincio	*Natural park	Landscape
25	Poesie dialettali del poeta veronese Berto Barbarani	*Book	Oral tradition†
26	Mostra di presepi diffusa nel borgo di Campo	*Abandoned neighbourhood	Practice†
27	The song of the bells of Guimaraes	A song	Music†
28	Metate de mi abuela	*Tool	Tool
29	Aurélia de Souza's Self-Portrait	Movable heritage. Artwork.	Painting
30	Dehesa Boyal	*Pasture, a natural park	Landscape
31	Cachamorra de pastor.	*Walking stick	Tool
32	Arepa Ocañera	A dish	Food†
33	Alfiler de traje militar de mi abuelo	*Military uniform	Personal object
34	Iglesia de Tlacotepec	*Temple	Building
35	Ex Voto	Painting	Painting
36	Street Art in Oaxaca	Printing on paper	Painting
37	Tlayuda de Tasajo	*Food	Food†
38	Juego del Vocho Amarillo	*Children's game	Practice†
39	Oxygene	Vinyl disc	Music
40	Taza y plato de café de Sanborns	*Tableware	Household item
41	Tacos al pastor	Food	Food†
42	Orange Mexico City Metro trains	Train	Vehicle
43	Recuerdos de mi primer viaje solo	*Phone card	Household item
44	Pocket watch	Watch	Personal object
45	My first watch	Watch	Personal object
46	My typewriter	Typewriter	Household item
47	Relógio da estação de Porto São Bento	*Building Clock	Building
48	Buzludzha monument	A monument	Building
49	Painting of wooden fishing boat	Painting	Painting

The sample number indicates the order in which the contributions were received. The title of each CHO is provided in its original language, while the type has been translated—where it was not originally expressed in English—and marked with an asterisk. The † symbol indicates that the CHO is intangible.The symbol * indicates that it is a compulsory field.

(sample 15). This apparent incompleteness could suggest that this metadata model may be excessively defined when describing heritage manifested as rather abstract concepts, thus exceeding contributors' needs.

When analysing the proportion of each category per field (Table 11), we found 100% valid responses in the `dc:title` and `dc:description` fields, which were mandatory at the point of submission. Valid responses for Subject and Type were present in 94% of the samples, suggesting that contributors felt confident in disclosing these properties regardless of the nature of the cultural heritage object.

As shown in Fig. 8, missing valid responses were rarely compensated by explicit Not Applicable or Unknown answers and instead more often resulted in blank fields. This represents a limitation of this approach, as it does not allow us to reliably determine what prevented participants from providing this information. Only the field `Publisher, issued, table of contents` received a comparatively higher number of Not Applicable responses, which can be attributed to the explicit description provided in the survey: «*Applicable if the CHO is a printed or published material*». Similarly, the language of the CHO was frequently left blank. While this may reflect common-sense reasoning on the part of participants, this experiment does not provide sufficient evidence to substantiate this interpretation.

The matrix relating the type of response per field and sample (Fig. 9) shows that the absence of a web resource (e.g., a photograph) attached to the contribution is, in most cases, followed by subsequent empty fields addressing the description, author, and creation date of such a digital file. This matrix also reveals that information relating to the scope and spatio-temporal context of contributions was frequently missing, potentially reflecting a conceptual barrier that could deserve more attention in future interface and participatory mechanism design.

Surprisingly, only four contributions used the free-text field at the end of the survey. Nonetheless, when analysing the length of responses per field, it is clear that contributors tended to provide relatively long texts in the `dc:description` field. In this light, the «Free text» field may have been perceived as redundant by contributors. This finding invited us to pay special attention to the `dc:description` component of the metadata, as the knowledge that participants wish to share about the cultural heritage object appears to be denser in that field.

Table 10 | Descriptive statistics of the categorised responses

Category	Total (49 samples)	Average per sample	(%)	σ
Valid	703	14.37	62.38%	4.1460
Blank	365	7.45	32.39%	4.4678
Not applicable	42	0.8571	3.73%	2.5658
Unknown	17	0.3469	1.51%	1.1096
Total	1078	22		

The value of descriptions in improving record completeness

When following a detailed review of the contents of the `dc:description` field, we identified multiple instances where the descriptive text contained information corresponding to metadata fields marked as Blank, Unknown, or Not Applicable. As summarised in Table 12, these cases frequently involve references to the spatial or temporal context of the cultural heritage manifestations. The relative concentration of semantically rich information within the `dc:description` field suggests that it may function as an important reservoir for metadata elements that are otherwise missing from the strctured fields.

The density of data in the `dc:description` field throughout the samples, and the potential transferability of information to other fields (otherwise marked as Blank, Unknown, or Not Applicable), are illustrated in Fig. 10. This indicates that the overall completeness of metadata exceeded our initial assessment. In light of the substantial number of empty or non-informative fields, this pattern may further suggest that contributors find it more intuitive to provide information in narrative form than through strictly predefined categories. Accordingly, approaches for tagging and extracting metadata elements from free-text descriptions should be considered in future contribution mechanisms.

Discussion

This work set out to explore how a community driven framework for metadata co-creation can be conceptually defined and operationalised at the point of community recording, particularly for forms of everyday heritage that fall outside the custody of formal memory institutions, showing

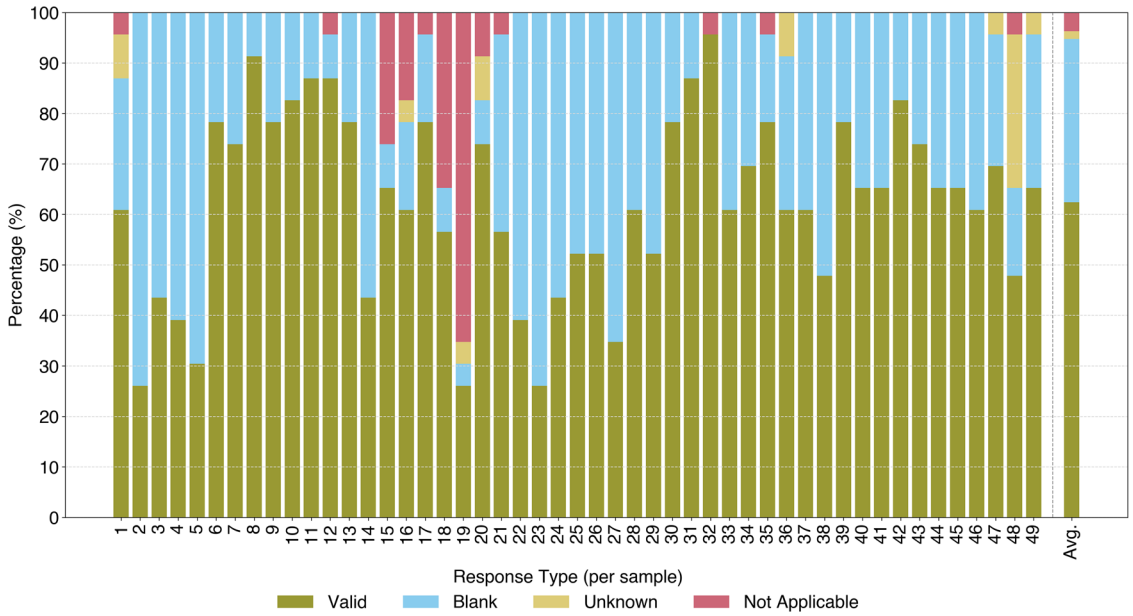

Fig. 7 | Distribution of response types per sample. The horizontal axis shows each one of the 49 entries registered. For each sample, the stacked bar represents the proportion of fields containing valid, blank, unknown, and non-applicable responses, respectively shown in green, blue, yellow, and red, totalling 100% of the survey fields. The bar on the right indicates the average values.

Table 11 | Proportion of categorised responses per field

	Responded		Unknown		Not applicable		Blank	
Field	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**
Title	49	100.00%	0	0.00%	0	0.00%	0	0.00%
Description	49	100.00%	0	0.00%	0	0.00%	0	0.00%
Language	17	34.69%	0	0.00%	0	0.00%	32	65.31%
Creator	36	73.47%	4	8.16%	1	2.04%	8	16.33%
Contributor	30	61.22%	1	2.04%	0	0.00%	18	36.73%
Provence	36	73.47%	0	0.00%	2	4.08%	11	22.45%
Source	36	73.47%	1	2.04%	0	0.00%	12	24.49%
Type	46	93.88%	1	2.04%	0	0.00%	2	4.08%
Format	42	85.71%	0	0.00%	1	2.04%	6	12.24%
Publisher, issued, table of contents	6	12.24%	0	0.00%	9	18.37%	34	69.39%
Extent	35	71.43%	0	0.00%	4	8.16%	10	20.41%
Medium	34	69.39%	0	0.00%	3	6.12%	12	24.49%
Subject	46	93.88%	0	0.00%	1	2.04%	2	4.08%
Created	43	87.76%	1	2.04%	1	2.04%	4	8.16%
Date	20	40.82%	1	2.04%	4	8.16%	24	48.98%
Scope	23	46.94%	0	0.00%	2	4.08%	24	48.98%
Temporal	20	40.82%	2	4.08%	1	2.04%	26	53.06%
Spatial	19	38.78%	1	2.04%	1	2.04%	28	57.14%
Web resource	29	59.18%	1	2.04%	1	2.04%	18	36.73%
Creator (Web Resource)	31	63.27%	1	2.04%	3	6.12%	14	28.57%
Description (Web Resource)	28	57.14%	1	2.04%	3	6.12%	17	34.69%
Created (Web Resource)	24	48.98%	2	4.08%	4	8.16%	19	38.78%
Free text	4	8.16%	0	0.00%	1	2.04%	44	89.80%

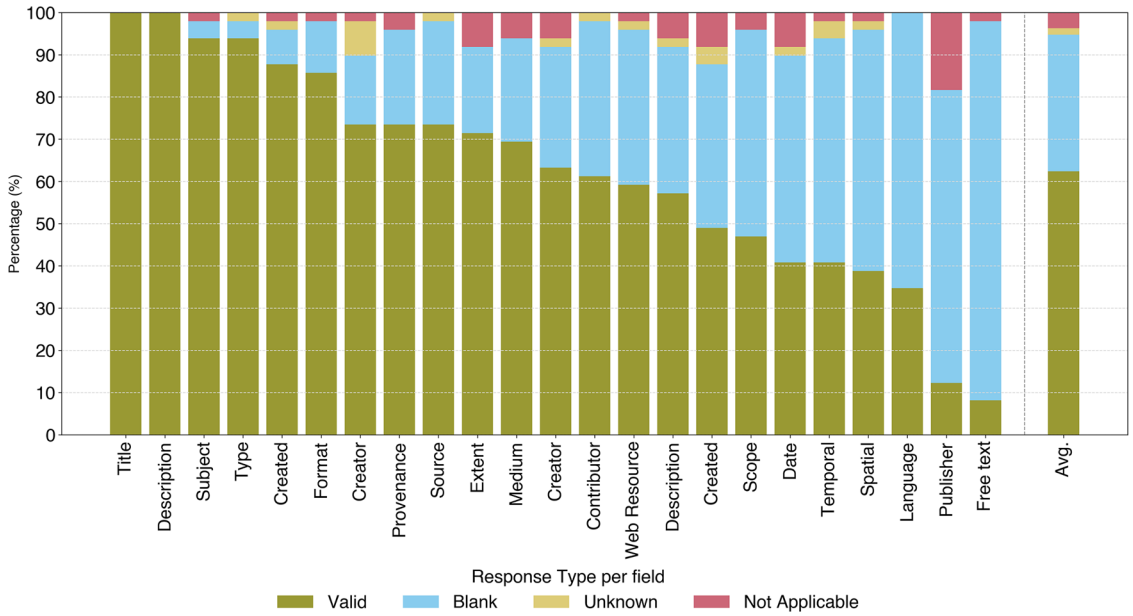

Fig. 8 | Distribution of response types per survey field. The horizontal axis shows each of the 23 fields. For each field, the stacked bar represents the proportion of entries containing valid, blank, unknown, and non-applicable responses, respectively shown in green, blue, yellow, and red, totalling 100% of the 49 registered entries. The bar on the right indicates the average values.

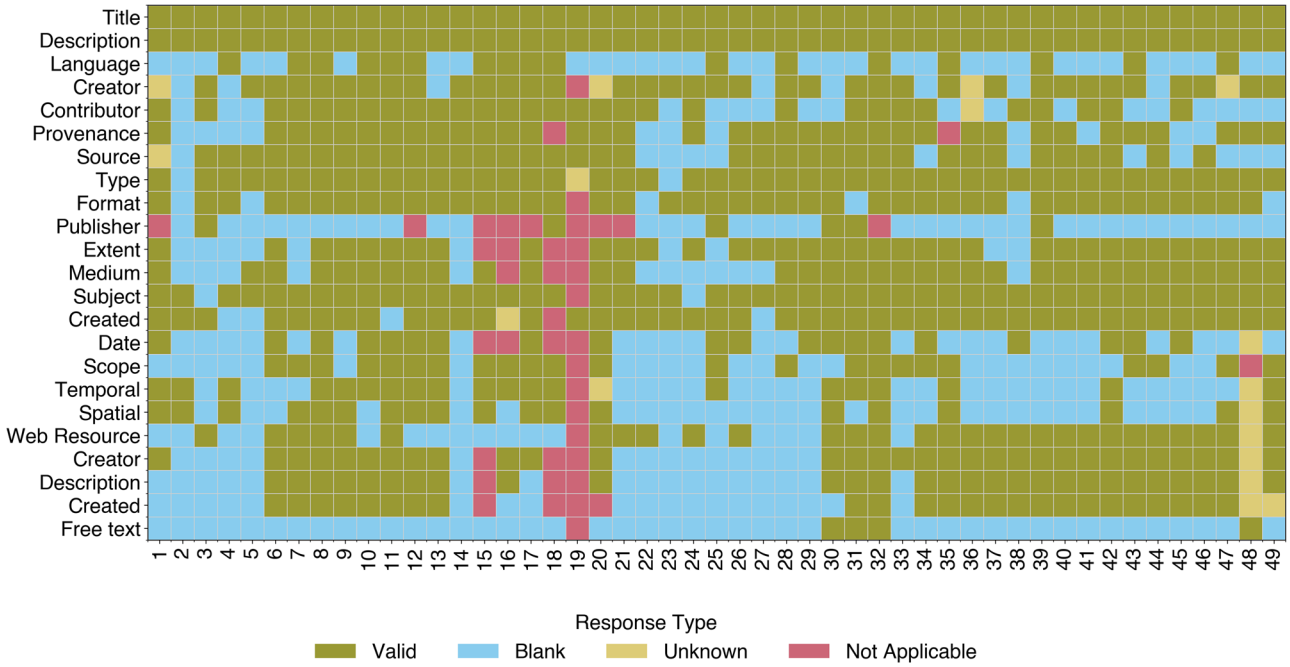

Fig. 9 | Matrix correlating each sample and the response type per field. The horizontal axis shows each of the 49 registered entries. The vertical axis shows each of the 23 fields. The colour of the squares at each intersection represents valid, blank, unknown, and non-applicable responses, respectively shown in green, blue, yellow, and red.

conceptually that FAIR principles can be taken beyond expert-driven, institutionally held collections to encompass the places, objects, practices, and memories that communities themselves value, thus challenging the prevailing current paradigm in which only heritage under formal custodianship can aspire to be FAIR.

We proposed a process-oriented framework grounded in citizen curation and aligned with the Europeana Data Model, and empirically explored its initial stage through a proof-of-concept survey. This first stage enables communities to act as primary contributors to the creation of FAIR metadata for everyday heritage and repositions memory institutions as co-authors and stewards of collectively produced knowledge. This alternative model conceives the metadata as a shared space for multiple interpretations rather than a single authoritative narrative.

To operationalise this approach, we introduced a novel and structured yet flexible framework grounded in citizen curation and implemented through the Europeana Data Model. The framework defines roles, responsibilities, and a five-stage lifecycle (recording, submission, curation, release, and use/reuse) explicitly linking actors' actions to metadata properties at each stage. A key design principle is the coexistence of community-generated folksonomies and institutional taxonomies. Rather than treating these as competing approaches, the model enables them to function symbiotically: folksonomies preserve local language, meaning, and lived experience, while taxonomies support interoperability, long-term stewardship, and integration within broader cultural heritage infrastructures. Crucially, institutional enrichment is conceived as additive and relational, not corrective or substitutive, thereby safeguarding the integrity of community expression.

The proof-of-concept study provides empirical support for the feasibility of the recording stage of the proposed framework. Participants successfully documented a wide spectrum of tangible and intangible cultural heritage manifestations, including highly localised, abstract, and linguistically specific forms of cultural value. The diversity of contributions confirms that communities are capable of articulating what constitutes heritage in their own terms when provided with an open and non-prescriptive framework. To illustrate the qualitative richness of the contributions generated during the Recording stage, we briefly examine the narrative content of entries submitted to the `dc:description` field. While a full qualitative analysis falls outside the scope of this study, even a superficial analysis reveals that participants frequently framed «everyday heritage» in experiential and relational terms, rather than through institutional or taxonomic language.

For example, the description of the «*Antica fiera di Cavalcaselle*» (sample 22) suggests that the place is significant because of its continuous use as a meeting point during the seasonal migration of livestock, describing it as a millenary tradition. Beyond describing the place and its temporal continuity, the submission adds that it «renews the sense of community of those who live in the territory», thereby highlighting its underlying communal value. This type of description foregrounds social practice, affect, and collective memory, rather than architectural features, chronology, or formal classification, which are elements that typically dominate institutional heritage metadata. Similar narrative patterns recur across multiple submissions, emphasising use, habit, and personal or communal attachment.

In «*The Song of the Bells of Guimarães*» (sample 27), the contributor describes this manifestation through the personal relationship sustained by memory and nostalgia: «*For me, the song is now nostalgic and tied to my memories of the city*». This type of contribution contains narrative content that reveals how participants conceptualise heritage and offers insights into how community narratives may differ from institutional metadata practices, as personal memories are central to the valorisation of these manifestations.

As a final example, the «*Orange Mexico City Metro trains*» illustrates how a specific feature of a system (the train's colour) is perceived as something of cultural value, thus diluting the boundaries between urban heritage and landscape. While such an analysis falls beyond the scope of this work, this and other contributions suggest that a less rigid conceptualisation of cultural heritage by non-expert communities could shed new light on the conceptions and categorisations that experts often resort to when documenting cultural heritage.

From a metadata perspective, these accounts populate fields such as `dc:description` with semantically rich yet folksonomic content which, although not immediately aligned with controlled vocabularies, constitutes valuable input for later enrichment during the curation stage. This observation reinforces the rationale of the proposed framework: to capture meaning-laden, community-authored narratives early in the lifecycle,

Table 12 | Information found in the description (`dc:description`) field that corresponds to metadata fields marked as blank, unknown, or not applicable at submission

Sample	Field	Original value	Extract from description
2	Type	Blank	«*…city image.*»
5	Language	Blank	«*…Sicilian dialect.*»
5	Format	Blank	«*…bread seasoned with plenty of oil, salt, and oregano, and then completed with tomatoes, anchovies, and cheese.*»
6	Spatial	Blank	«*…in Germany*.»
14	Scope	Blank	«*…cocina mexicana.*»
18	Provenience	Not Applicable	«*It evokes my grandparents and older generations.*»
21	Temporal	Blank	«*XV century*.»
21	Spatial	Blank	«*Comune di Castelnuovo del Garda, provincia di Verona, sul lago di Garda*.»
22	Temporal	Blank	«*…testimoniato fino dall'anno mille …la terza domenica e lunedí di novembre*.»
22	Spatial	Blank	«*Monte Baldo.*»
23	Type	Blank	«*Dolce tipico*.»
23	Medium	Blank	«*500 g di mascarpone, 5 uova, 5 cucchiaini di zucchero, biscotti tipo pavesini, caffé e crema marsala.*»
23	Spatial	Blank	«*Veneto, Treviso.*»
24	Spatial	Blank	«*Il corso del fiume Mincio dal lago di Garda alla confluenza nel fiume Po*.»
25	Source	Blank	«*Tradizioni popolari veronesi.*»
25	Spatial	Blank	«*Verona.*»
26	Medium	Blank	«*Presepi.*»
26	Spatial	Blank	«*All'interno di vecchie case e stalle, tra i vicoli e nei sottoscala.*»
27	Spatial	Blank	«*Guimarães.*»
27	Temporal	Blank	«*The closer the clock gets to a full hour, the longer the part played.*»
29	Scope	Blank	«*Portuguese modern art.*»
29	Spatial	Blank	«*Soares dos Reis National Museum, Porto*.»
31	Format	Blank	«*…uma vara de madeira*.»
48	Temporal	Unknown	«*Socialist regime in Bulgaria.*»
48	Spatial	Unknown	«*On the top of the mountain.*»
49	Format	Blank	«*Oil painting.*»

Extracts are kept as originally expressed.

before their translation into standardised or interoperable forms. At the same time, this material highlights the importance of subsequent stages in the framework (particularly curation) in mediating between community expression and institutional interoperability requirements. A more systematic qualitative analysis of these narratives, ideally conducted through participatory and community-led interpretive methods, is identified as an important direction for future research.

Participants were recruited through mailing lists and personal invitations, which introduces limitations in terms of representativeness and potential sampling bias. Although the survey was also disseminated beyond academic and professional research networks, its anonymous design prevents us from determining the extent to which individuals outside these environments participated. Consequently, the sample may over-represent individuals already familiar with cultural heritage or research practices and may not fully reflect the diversity that this framework seeks to capture. Furthermore, the mode of participation may itself have limited engagement from individuals with less experience in research contexts, lower digital literacy, or restricted access to technical resources. These constraints should be considered in future applications and implementations of the framework.

Importantly, the empirical component does not implement or validate the full lifecycle framework. Rather, it examines the feasibility and expressive characteristics of community led metadata recording under FAIR aligned design assumptions. The findings therefore inform the design rationale of the framework (specifically its emphasis on narrative expression, folksonomic input, and incremental enrichment) without claiming validation of downstream stages such as curation, release, or reuse. At the same time, the uneven completion of certain metadata fields highlights important challenges that deserve further research. These findings suggest that future implementations should prioritise narrative-oriented interfaces, flexible

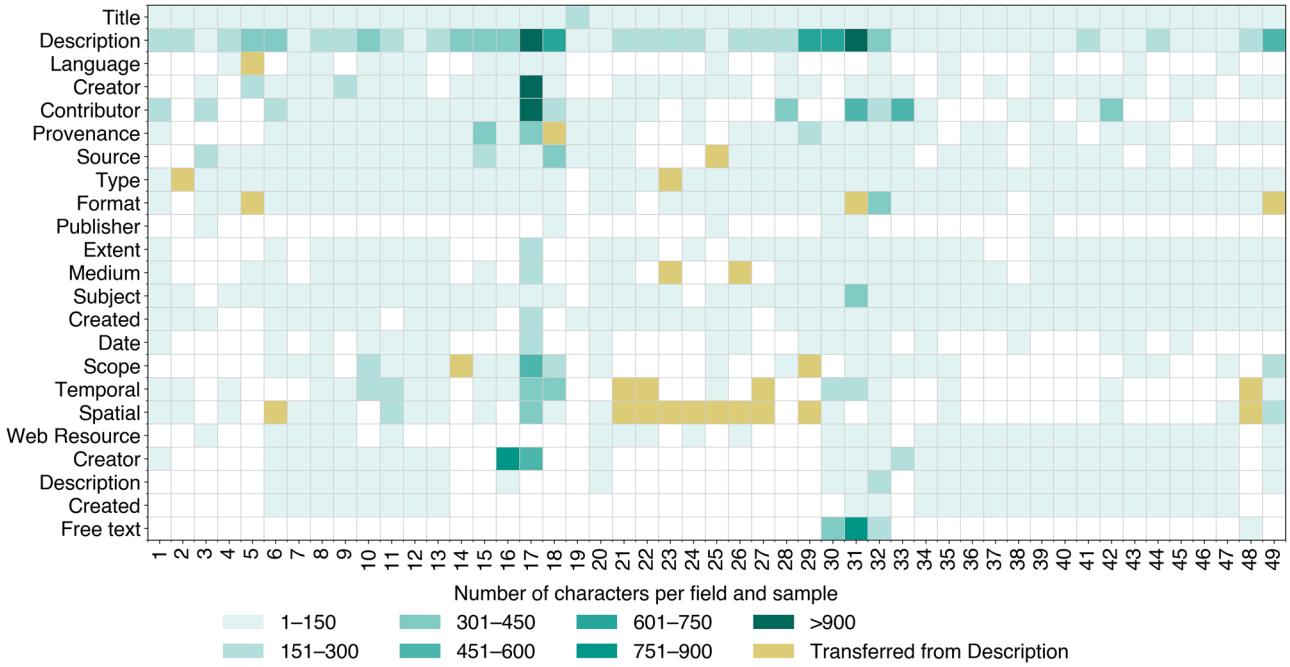

Fig. 10 | Matrix of response types illustrating response length and information transferability from the dc:description fields. The horizontal axis shows each of the 49 registered entries. The vertical axis shows each of the 23 fields. The colour of the squares at each intersection represents the length (in number of characters) of each response, classified in intervals of 150 characters from 1 up to more than 900. Squares highlighted in yellow indicate where initially empty fields received information transferred from the `dc:description` field.

data entry mechanisms, and participatory designs that do not presume uniform digital literacy or fixed conceptual models of heritage.

Despite the relatively small sample, it was possible to test the feasibility and conceptual soundness of this stage. However, further research needs to evaluate the long-term dynamics of the curation, release, and reuse stages, particularly in institutional settings. Future work should also examine how the framework performs in sustained collaborations with specific communities and memory institutions, including contexts involving sensitive, contested, or Indigenous heritage, where ethical considerations about access, authority, and representation are especially critical.

This study demonstrates that designing a FAIR-aligned, community-driven metadata framework is timely, technically viable, and socially meaningful. By aligning participatory heritage practices with established metadata standards, the proposed approach supports more inclusive, equitable, and plural cultural memory ecosystems. Making everyday heritage FAIR is not merely a matter of extending technical infrastructures, but a cultural and ethical commitment to recognising whose knowledge is documented, preserved, and made visible for the future.

Data availability

The datasets generated and analysed during the current study are available in the following repository: Ramirez Eudave, R., Ferreira, T. M., & Giardina, G. (2026). CORDHISK Survey on Community-driven knowledge about cultural heritage [Data set]. Zenodo. https://doi.org/10.5281/zenodo.18678879.

Received: 3 March 2026; Accepted: 31 May 2026;
Published online: 11 June 2026

References

1. Diara, F. Cultural Heritage Digital Data: Future and Ethics. In *Digital Cultural Heritage*, (ed Kremers, H.) 303–314 https://doi.org/10.1007/978-3-030-15200-0_20 (Springer International Publishing, 2020).
2. Larsson, sM., Bornsäter, B. & Hacke, M. Developing practices for FAIR and linked data in Heritage Science. *npj Herit. Sci.* **13**, 53 (2025).
3. Ranjgar, B., Sadeghi-Niaraki, A., Shakeri, M., Rahimi, F. & Choi, S.-M. Cultural heritage information retrieval: past, present, and future trends. *IEEE Access* **12**, 42992–43026 (2024).
4. Gibson, R. C., Chowdhury, S. & Chowdhury, G. User versus institutional perspectives of metadata and searching: an investigation of online access to cultural heritage content during the COVID-19 pandemic. *Int. J. Digital Libraries* **25**, 105–121 (2024).
5. Alivizatou-Barakou, M. et al. Intangible cultural heritage and new technologies: challenges and opportunities for cultural preservation and development. In *Mixed Reality and Gamification for Cultural Heritage*, (eds Ioannides, M., Magnenat-Thalmann, N. & Papagiannakis, G.) 129–158 https://doi.org/10.1007/978-3-319-49607-8_5 (Springer International Publishing 2017).
6. Barrére, C. Cultural heritages: from official to informal. *City, Cult. Soc.* **7**, 87–94 (2016).
7. Dahlgren, A. & Hansson, K. The diversity paradox: conflicting demands on metadata production in cultural heritage collections. *Digital Cult. Soc.* **6**, 239–256 (2020).
8. Dahlgren, A. & Hansson, K. What an image is: the ontological gap between researchers and information specialists. *Art. Documentation J. Art. Libraries Soc. North Am.* **40**, 21–32 (2021).
9. Barabucci, G., Tomasi, F. & Vitali, F. Modeling data complexity in public history and cultural heritage. In *Handbook of Digital Public History* (eds Noiret, S.,Tebeau, M. & Zaagsma, G.) 459–474 https://doi.org/10.1515/9783110430295-041/html (De Gruyter, 2022).
10. Shi, J., Nason, M., Tullney, M. & Alperin, J. Identifying metadata quality issues across cultures. *Coll. Res. Libraries* **86**, 101–134 (2025).
11. Kim, J. Metadata as radical care: how community-led archives reimagine descriptive practices. *IJournal Student J. Faculty Inf.* **10**, https://theijournal.ca/index.php/ijournal/article/view/45416 (2025).
12. Allard, D. & Ferris, S. Antiviolence and marginalized communities: knowledge creation, community mobilization, and social justice through a participatory archiving approach. *Libr. Trends* **64**, 360–383 (2015).

13. Stevens, M., Flinn, A. & Shepherd, E. New frameworks for community engagement in the archive sector: from handing over to handing on. *Int. J. Herit. Stud.* **16**, 59–76 (2010).
14. Roeschley, A., Kim, J. & Zavalina, O. L. An exploration of contributor-created description field in participatory archives. In *Sustainable Digital Communities*, vol. 12051 (eds Sundqvist, A.,Berget, G., Nolin, J. & Skjerdingstad, K. I.) 638–648 (Springer International Publishing, Cham, 2020). https://doi.org/10.1007/978-3-030-43687-2_54.
15. Barber, S. T. The ZOONIVERSE is expanding: crowdsourced solutions to the hidden collections problem and the rise of the revolutionary cataloging interface. *J. Libr. Metadata* **18**, 85–111 (2018).
16. Roeschley, A. & Kim, J. "Something that feels like a community": the role of personal stories in building community-based participatory archives. *Archival Sci.* **19**, 27–49 (2019).
17. Howell, R. & Chilcott, M. A sense of place: re-purposing and impacting heritage resource evidence through digital heritage and interpretation practice. *Int. J. Intang. Herit.* **8**, 165–178 (2013).
18. Council of Europe. *Council of Europe Framework Convention on the Value of Cultural Heritage for Society* (Council of Europe, 2025).
19. Colomer, L. Exploring participatory heritage governance after the EU Faro Convention. *J. Cultural Herit. Manag. Sustain. Dev.* **13**, 856–871 (2023).
20. Lauso Zagato. The Notion of “Heritage Community" in the Council of Europe’s Faro Convention. Its Impact on the European Legal Framework. In *Between Imagined Communities of Practice: Participation, Territory and the Making of Heritage* (eds Adell, N.,Bendix, R. F., Bortolotto, C. &Tauschek, M.) (Göttingen University Press, 2015). http://books.openedition.org/gup/191.
21. Kartawihardja, D. S., Pee, A. N. C. & Zakaria, M. H. A modification on existing cultural metadata management system to accommodate community-driven and holistic cultural metadata. *J. Phys. Conf. Ser.* **1201**, 012037 (2019).
22. Kaldeli, E. et al. CrowdHeritage: crowdsourcing for improving the quality of cultural heritage metadata. *Information* **12**, 64 (2021).
23. Daga, E. et al. Integrating citizen experiences in cultural heritage archives: requirements, state of the art, and challenges. *J. Comput. Cultural Herit.* **15**, 1–35 (2022).
24. Musen, M. A. et al. Modeling community standards for metadata as templates makes data FAIR. *Sci. Data* **9**, 696 (2022).
25. Pasciuto, T., Albertoni, R. & Maggi, R. Metadata schemas in cataloguing the GLAM domain: an overview. *J. Inf. Sci.* https://journals.sagepub.com/doi/10.1177/01655515251365580 (2025).
26. Ferreira-Lopes, P. & Gracia, E. G. Performing a systematic literature review on the implementation of the CIDOC CRM in cultural heritage. *J. Comput. Cultural Herit.* **18**, 1–16 (2025).
27. Dorobăț I. C. & Posea, V. Enriching the cultural heritage metadata using historical events: a graph-based representation. In *Digital Libraries for Open Knowledge*, vol.11799 of *Lecture Notes in Computer Science* (eds Doucet, A.,Isaac, A., Golub, K.,Aalberg, T. & Jatowt, A.) 344–347 (Springer International Publishing, 2019).
28. Peroni, S., Tomasi, F. & Vitali, F. Reflecting on the Europeana Data Model. In *Digital Libraries and Archives*, vol. 354, (eds Agosti, M.,Esposito, F., Ferilli, S. &Ferro, N.) 228–240 https://doi.org/10.1007/978-3-642-35834-0_23 (Springer, 2013).
29. Adina Ciocoiu. EDM Mapping Guidelines. Tech. Rep., https://europeana.atlassian.net/wiki/spaces/EF/pages/987791389/EDM+-+Mapping+guidelines (2017).
30. Regazzoni, L. Unintentional monuments, or the materializing of an open past. *Hist. Theory* **61**, 242–268 (2022).
31. García-Esparza, J. A. & Nikšič, M. Revealing the community’s interpretation of place: integrated digital support to embed photovoice into placemaking processes. *Urban Planning* **9**, https://www.cogitatiopress.com/urbanplanning/article/view/7074 (2024).
32. Chiripanhura, P., Sibanda, M. & Chabata, F. Digital repatriations: a temporary solution to improve access to Zimbabwe’s heritage. In *Revaluation and Preservation of Indigenous Knowledge Systems in Modern Society* (eds Masenya, T. M. & Bhuda, M. T.) 17–38 https://doi.org/10.4018/979-8-3693-7964-6.ch002 (IGI Global, 2025).
33. UNESCO. UNESCO Virtual Museum of Stolen Cultural Objects, available at https://museum.unesco.org/ (2025).
34. Pilav, A. Before the war, war, after the war: urban imageries for urban resilience. *Int. J. Disaster Risk Sci.* **3**, 23–37 (2012).
35. Veronese, G., Hamamra, B., Mahamid, F. & Bdier, D. Heritage fractured: violence, displacement and the disintegration of cultural identity. *J. Cultural Heritage Manag. Sustain. Dev.* 1–14, https://www.emerald.com/jchmsd/article/doi/10.1108/JCHMSD-02-2025-0026/1272383/Heritage-fractured-violence-displacement-and-the (2025).
36. Wagner, A. & De Clippele, M.-S. Safeguarding cultural heritage in the digital era - a critical challenge. *Int. J. Semiotics Law Rev. Int. de. Sémiotique Jurid.* **36**, 1915–1923 (2023).
37. Relph, E. Digital disorientation and place. *Mem. Stud.* **14**, 572–577 (2021).
38. Szabo, V., Lacedelli, S. Z. & Pompanin, G. From landscape to cities: a participatory approach to the creation of digital cultural heritage. *Int. Inf. Libr. Rev.* **49**, 115–123 (2017).
39. Echavarria, K. R., Samaroudi, M., Dibble, L., Silverton, E. & Dixon, S. Creative experiences for engaging communities with cultural heritage through place-based narratives. *J. Comput. Cultural Herit.* **15**, 1–19 (2022).
40. Molho, J. The promises and limitations of digital participation in heritage: Planning transmedia heritage districts in superdiverse cities. *Int. J. Herit. Stud.* **29**, 860–876 (2023).
41. Mason, M. & Vavoula, G. Digital cultural heritage design practice: a conceptual framework. *Des. J.* **24**, 405–424 (2021).
42. Shi, X., Beza, B. B., Liu, C., Wu, B. & Qiu, N. Differences in perspectives between experts and residents on living heritage: a study of traditional Chinese villages in the Luzhong region. *Buildings* **14**, 4022 (2024).
43. Cerquetti, M. & Ferrara, C. Marketing research for cultural heritage conservation and sustainability: lessons from the field. *Sustainability* **10**, 774 (2018).
44. Silva, P. I. & Terra, A. L. The role of users in the organization of digital information: a Portuguese experience in an academic museum and archive setting. *IFLA J.* **50**, 64–74 (2024).
45. Seth van Holland. Spectator becomes annotator : possibilities offered by user-generated metadata for image databases. In *Proceedings of the CILIP Cataloguing & Indexing Group Annual Conference* http://eprints.rclis.org/8080/ (2006).
46. Stock, W. G. Folksonomies and science communication: a mash-up of professional science databases and Web 2.0 services. *Inf. Serv. Use* **27**, 97–103 (2007).
47. Bullard, J., Crowston, K., Jackson, C., Smith, A. O. & Østerlund, C. Folksonomies in crowdsourcing platforms: three tensions associated with the development of shared language in distributed groups. In *Proceedings of 22nd European Conference on Computer-Supported Cooperative Work* (European Society for Socially Embedded Technologies (EUSSET), 2024). https://dl.eusset.eu/handle/20.500.12015/5112.
48. De Oliveira Souza, G. & Vicentini Jorente, M. J. Information design and digital curation at the museum of the person: folksonomy converged to the social technology of memory. *Can. J. Inf. Libr. Sci.* **47**, 128–136 (2024).
49. Open Heritage Coalition. Open Heritage Statement, available at https://openheritagestatement.org/statement (2025).
50. Vézina, B. & Muscat, A. Sharing Indigenous Cultural Heritage Online: An Overview of GLAM Policies https://creativecommons.org/2020/08/08/sharing-indigenous-cultural-heritage-online-an-overview-of-glam-policies/ (2020).

51. Nuohtti Project Consortium. Ethical Guidelines. Tech. Rep., Nuohtti Project Consortium. https://nuohtti.com/Content/ethical-guidelines?lng=en-gb (2021).
52. Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).
53. Museu da Luz. Relógio - Objeto do mês de abril https://museudaluz.org.pt/2026/04/08/relogio-objeto-do-mes-de-abril-2026/ (2026).
54. Ramirez Eudave, R., Ferreira, T. M. & Giardina, G. CORDHISK Survey on Community-driven knowledge about cultural heritage. https://doi.org/10.5281/zenodo.18678878 (2026).
55. Creative Commons. CC0 1.0 Universal Deed, available at https://creativecommons.org/publicdomain/zero/1.0/ (2015).

Acknowledgements

This project has received funding from the European Union's Horizon Europe 2023 research and innovation programme under the Marie Skłodowska Curie grant agreement No 101149833 for the project «Community-driven Digitisation for Heritage at Risk» (CORDHISK). This research was funded in part by the Fundação para a Ciência e a Tecnologia, I.P. under Grant UID/06438/2025 of the research unit CERIS.

Author contributions

R.R.E., T.M.F., and G.G. contributed to the conceptualisation of this work (ideas; formulation or evolution of overarching research goals and aims) and to Writing – Review & Editing (preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre- or post-publication stages). R.R.E. and G.G. contributed to Investigation (conducting a research and investigation process, specifically performing the experiments, or data/evidence collection) and Project Administration (management and coordination responsibility for the research activity planning and execution). G.G. and T.M.F. contributed to Supervision (oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team). R.R.E. contributed with Data Curation (management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later reuse); Formal analysis (application of statistical, mathematical, computational, or other formal techniques to analyse study data); Funding Acquisition (acquisition of the financial support for the project leading to this publication); Methodology (development or design of the methodology); Software (programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components); Validation (verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs); Visualization (preparation, creation and/or presentation of the published work, specifically visualization/data presentation); and Writing – Original Draft Preparation (creation and/or presentation of the published work, specifically writing the initial draft, including substantive translation). All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Rafael Ramírez Eudave.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2026