International Journal on Digital Libraries (2024) 25:603–622
https://doi.org/10.1007/s00799-023-00383-0

Universities, heritage, and non-museum institutions: a methodological proposal for sustainable documentation

Marina Salse-Rovira[1] · Nuria Jornet-Benito[1] · Javier Guallar[1] · Maria Pilar Mateo-Bretos[2] · Josep Oriol Silvestre-Canut[2]

Received: 21 November 2022 / Revised: 12 September 2023 / Accepted: 15 September 2023 / Published online: 27 October 2023
© The Author(s) 2023

Abstract
To provide a sustainable methodology for documenting the small (and underfunded) but often important university heritage collections. The sequence proposed by the DBLC (Database Life Cycle) (Coronel and Morris, Database Systems: Design, Implementation, & Management. Cengage Learning, Boston, 2018; Oppel Databases a beginner's guide. McGraw-Hill, New York, 2009) is followed, focusing on the database design phase. The resulting proposals aim at harmonising the different documentation tools developed by GLAM institutions (acronym that aims to highlight the common aspects of Galleries, Libraries, Archives and Museums), all of which are present in the university environment. The work phases are based mainly on the work of Valle, Fernández Cacho, and Arenillas (Muñoz Cruz et al. Introducción a la documentación del patrimonio cultural. Consejería de Cultura de la Junta de Andalucía, Seville, 2017), combined with the experience acquired from the creation of the virtual museum at our institution. The creation of a working team that includes university staff members is recommended because we believe that universities have sufficient power to manage their own heritage. For documentation, we recommend the use of application profiles that consider the new trends in semantic web and LOD *(Linked Open Data)* and that are created using structural interchange standards such as Dublin Core, LIDO, or Darwin Core, which should be combined with content and value standards adapted from the GLAM area. The application of the methodology described above will make it possible to obtain quality metadata in a sustainable way given the limited resources of university collections. A proposed metadata schema is provided as an annex.

Keywords University heritage · Cultural heritage documentation · GLAM · University collections

1 Introduction

Throughout their history, universities have collected (and still collect) a lot of items related to their teaching and research activity. For example, faculties offering natural science courses needed plant and biological specimens so that students could learn about the reality of remote environments at a time when travelling was not as easy and fast as it is today, and the Internet was still merely science fiction. Meanwhile, libraries were acquiring books to be used for study and research, and archives were collecting the university's institutional documents. Unfortunately, as science and technology developed, many items ended up in the backrooms or in reserve libraries, giving rise to the so-called Cinderella Collections [47].

At the same time, often for reasons of prestige [17, 25, 28, 30], many universities set up art collections with works

✉ Marina Salse-Rovira
salse@ub.edu

Nuria Jornet-Benito
jornet@ub.edu

Javier Guallar
jguallar@ub.edu

Maria Pilar Mateo-Bretos
pilimateo@ub.edu

Josep Oriol Silvestre-Canut
osilvestre@ub.edu

1 Department of Library and Information Science and Audiovisual Communication, University of Barcelona, Barcelona, Spain

2 Office of the Vice-Principal for Heritage and Cultural Activities, University of Barcelona, Barcelona, Spain

that were sometimes purchased by the university itself, sometimes as a result of donations, and other times self-generated in the fine arts faculties. These collections were also often housed in buildings that were small architectural jewels of their time, as they had to demonstrate the university's status as an institution of knowledge and research.

All this has created a wealth of heritage that is very important for any country. For example, 30% of Britain's heritage recognised as culturally significant by the Department for Culture, Media and Sport is found in universities, yet these institutions house only 4% of the country's museums [24, 42]. However, this heritage is often poorly documented, usually due to a lack of resources and staff. The studies carried out by Salse et al. [35] show that much of this heritage is managed on a part-time basis by enthusiastic individuals who are experts in their field of knowledge but who are not experts in documenting heritage and do not have the time to give this task the level of attention it requires.

On the other hand, there is no shortage of documentation tools available for the task. Archives, libraries, and museums have generated their own often very complete documentation tools and have recognised the points that all these institutions have in common, using unifying terms such as GLAM or "memory institutions" [14, 50]. They have also generated tools that facilitate record sharing, such as metadata exchange schemas (Dublin Core, LIDO) and crosswalks. Finally, they have also found sites such as Europeana or the DPLA (Digital Public Library of America), which make GLAM records accessible from a single location. However, these tools have only occasionally been combined for integrated use in university heritage collections.

2 Objectives

This article proposes a methodology to bridge the gap between the existing documentation tools and the day-to-day reality of collections, especially small university collections with few resources and/or staff, which in this article are referred to as "non-museum collections".

The proposal is based on the experience acquired in the creation of the Virtual Museum of the University of Barcelona, one of Spain's so-called historical universities, founded in 1401 as the *Estudi General de Medicina i Arts*. The University of Barcelona (UB) is a perfect example of an institution that houses non-museum institutions, since it has various collections that fall into the following categories among those defined by UNESCO [39]: **movable cultural heritage, immovable cultural heritage, and intangible cultural heritage**. Furthermore, over the years, the UB has been upholding the importance of harvesting the **human heritage** 3F that has left its mark on the institution [40] for the benefit of the university environment. This last concept is not listed by UNESCO, but it is particularly important in the tourism sector.

With this starting point in mind, this article has two specific objectives:

1. To present a sustainable methodology for the documentation of non-museum collections in a digital environment. 4F[1] The aim is to establish a descriptive procedure that is sustainable and that can ensure that the items described for the future, with standardised schemas, are interoperable at all levels 5F[2] and adapted to the semantic web and the LOD movement. This methodology should make it possible to achieve a unified vision of university heritage that respects the diversity of the different collections but at the same time reveals the relationships that often exist between them (for example, different collections have items associated with medicine).
2. To propose a documentation model that facilitates the description and organisation of items or groups of items in these non-museum structures. This involves identifying the metadata system deemed the most appropriate and recommending the use of different value and content standards for the appropriate composition of the metadata.

The following is not explored in this article:

Specific proposals for documentation in archives, libraries, and museums. These organisations already have their own working systems and strong institutions providing guidelines (ICA, IFLA, ICOM6F[3]). Their work focuses on the processing of records for central system input and on project management at the metadata level.
Specific proposals for human and intangible heritage. We believe that this should be the subject of a specific study to be published in the future.
Specific proposals dealing with processes associated with the life of the heritage item. For example, we would exclude the management of documentation resulting from the processes of selection, acquisition, loan, or management of events such as exhibitions.

[1] This refers to non-museum collections that do not and cannot comply with different points of the *Statement of principles of museum documentation*. Particularly important for our definition is point 1.2.2. of the statement: "The museum must employ or have access to staff with appropriate expertise in documentation procedures, standards and systems" [12].

[2] The levels of interoperability mentioned by Zeng [49], i.e. system (data presentation), syntactic (data language), structural (data structure) and semantic (data context), are applied here.

[3] ICA (International Council of Archives), IFLA (International Federation of Library Associations), ICOM (International Council of Museums).

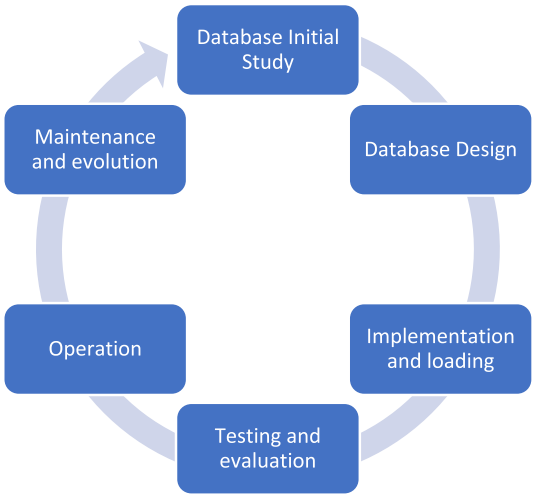

Fig. 1 Database life cycle (DBLC). [13] Source: Authors, based Coronel & Morris

3 Methodology

There are various methodologies that attempt to explain how to create a documentation project, but they are often based on the discipline of origin. The guiding principle for archivists and document managers is the ISO 15489 document management standard [22] and all the ISO standards and recommendations derived from it. In museums there have been various proposals, such as the joint work of UNESCO and ICOM [33]. On the other hand, Zeng [49] proposes a template with a series of steps to be completed in any metadata project. However, in the field of computer science, the SDLC (System Development Life Cycle) is used to generate computer systems [31], while the DBLC (Database Life Cycle) [13] is used to generate databases. The phases comprising this methodology can be seen in Fig. 1.

In our study, we use DBLC as a global framework (since the ultimate objective of our system is the generation of a heritage database), which involves the following phases:

Within this generic structure, we will focus mainly on the Database Design phase since the aim of our study is to propose a sustainable documentation methodology for non-museum university collections. Although it is advisable to bear in mind the indications of institutions such as the aforementioned ISO, ICOM and UNESCO, in this study, we will adapt the recommendations of Valle, Fernández, and Arenillas [29] as they refer more specifically to heritage documentation. These authors structure this design phase as follows:

- Definition of work team
- Definition of objectives of documentation
- Definition of subject of documentation
  - o Classification of heritage
  - o Selection of elements to be documented.
- Definition of documentation tools and methodologies
  - o Documentation methodology
  - o Systematisation and standardisation of information
  - o Information management (not discussed in this article)

The choice of metadata schemas and associated standards is based on the studies by Salse et al. [34, 35] identifying the different metadata schemas used in the field of university heritage, and also on the experience acquired by some of the authors of this study in teaching in the field of metadata and databases, as well as the experience gained throughout the implementation process of the new Virtual Museum of the University of Barcelona, as a case study. This process began at the end of 2018 and its main milestone was the publication in September 2021 of the first version of the new UB Virtual Museum (still under development).

4 Results (methodological proposal)

4.1 Definition of work team

Although this work team is not specifically part of the database design and implementation, it is essential for carrying out the entire DBLC and, as will be discussed below, it is also a key element in the creation of a sustainable university infrastructure for long-term heritage management.

In the initial selection of members of this team, it is very important to consider that the universities already have staff with the scientific and technical knowledge for effective heritage documentation. However, this staff is often dispersed across different departments and only librarians and archivists have documentation as one of their main tasks. Therefore, a working team needs to be created that will bring together the different areas of expertise. The structure proposed in Fig. 2 covers the tasks that should be carried out at the heritage level according to ICOM [12], although it is only a guideline and could include other actors, such as teachers in the marketing area to design dissemination programmes, or specialists in cultural management who could come from the economics area. The key factor to bear in mind is that **it is necessary to involve personnel from different areas**, faculties, or services within the university to reinforce the idea of a shared heritage that needs to be preserved.

Specifically, the following guidelines are proposed:

- The **work team should represent a cross section of the institution,** bringing together professionals from different areas of expertise to achieve a common goal.
- **The management of activities** (acquisitions, transfers, loans, exhibitions, events in general) should remain in the

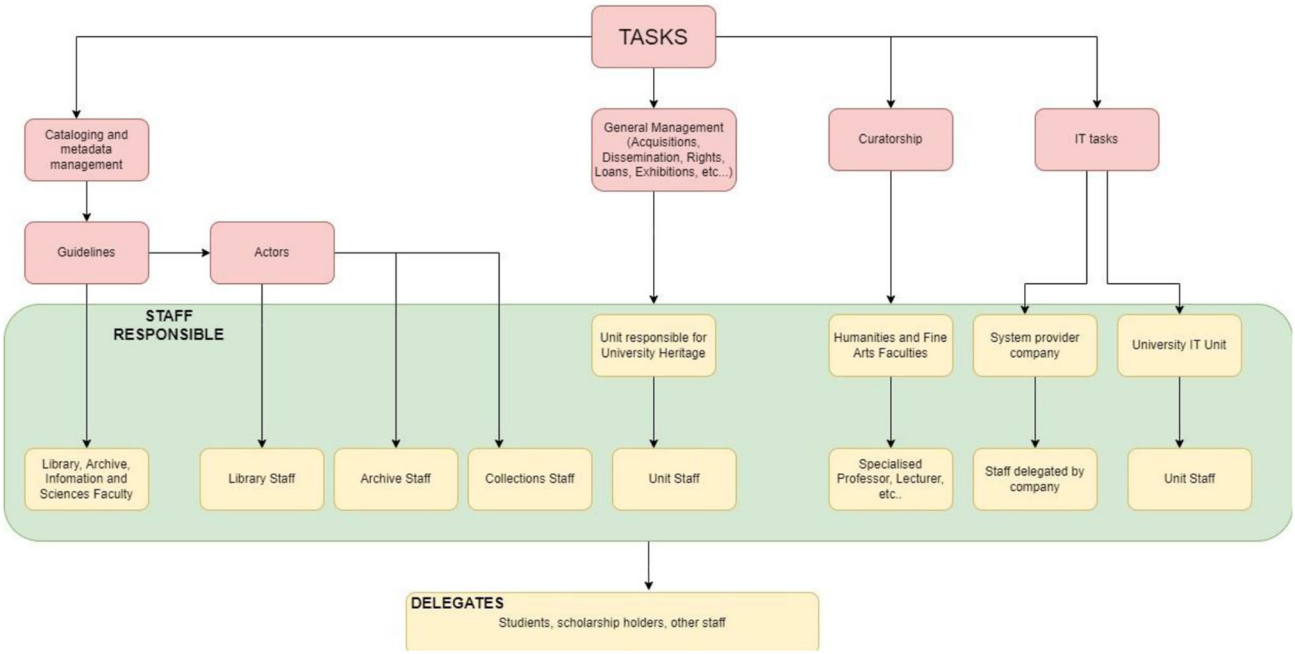

Fig. 2 Work structure proposal. Source: Authors

hands of the central heritage unit, although certain aspects may be delegated to the heads of the collections. It is important, for example, that the first management derives from the heads of the collections, who are the ones who know whether a particular donation or offer can be accepted. These management procedures would have to be specified in an internal regulation of the institution that would establish the specific processes.

- **The cataloguing/inventory guidelines** must remain in the hands of the library and/or archive and/or the corresponding faculty of information science if one exists. The choice of these three actors is because both the library and the archive are specialists in the management of their respective metadata, while the corresponding faculty can provide a broader vision of the world of metadata that responds to the needs of the different collections. If the university has museums with staff dedicated to documentation, these would also be a good option.
- **The cataloguing** itself should be carried out by the people designated within each unit/collection/museum, as they are the ones who know the context and the importance of the catalogued item, with the existence of occasional external partnerships, such as collaboration grants, crowdsourcing [3] or collaborations with final projects in undergraduate programmes or doctoral theses.
- **The conservation** of museum objects should be the responsibility of the university's fine arts centres, while libraries and archives, which often have restoration workshops, should be responsible for written historical documents. In the case of museum work, it should be remembered that most centres that teach studies related to the arts have conservation and preservation courses, and university collections offer excellent opportunities for practical coursework.

At the top of the structure shown in Fig. 2, there should be a steering committee chaired by the position responsible for managing/directing/deciding on the university's cultural heritage (in the case of the University of Barcelona, the Vice-Rector for Arts, Culture, and Heritage) or the person delegated thereby. However, it is essential that this structure should also include at least one cultural heritage expert on the staff of the university itself, as a centralised management figure(s).

One possible composition of this committee is depicted in Fig. 3.

The functions of this committee would essentially be to establish basic periodic lines of action, which would be set out in successive **master plans**, and **specific working procedures**. Elements of analysis for the development of these tools are presented below.

4.2 Definition of objectives of documentation (why we document)

This issue will have to be clearly set out in the master plan, if necessary, at a global level, but it will probably **be necessary**

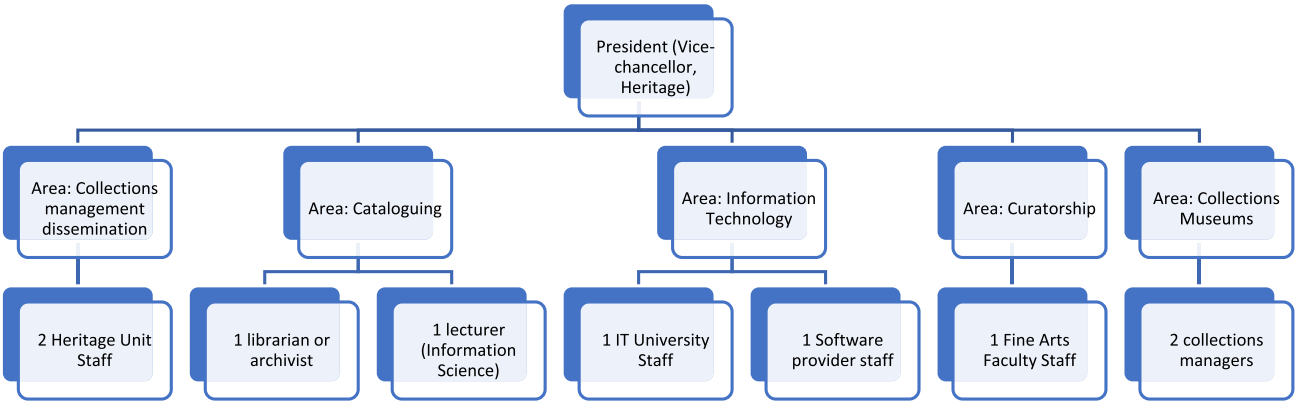

Fig. 3 Steering committee (proposal). Source: Authors

to define objectives at the level of the university, as well as for each collection, as collections are too diverse to be covered by a single, homogeneous objective.

At the global level, Valle, Fernandez, and Arenillas define five main types of documentation objectives in heritage management: as a knowledge strategy, to protect cultural assets, to conduct research, to engage in preservation actions, or to disseminate information. In the case of university heritage, all these objectives are valid, although their importance may vary in specific collections. For example, the objectives of zoological collections have a strong research component, while protection and dissemination are probably key aims for fine arts heritage.

4.3 Definition of subject of documentation (what we document)

Once the objectives of documentation have been established, it is necessary to examine **what** it is that we want to document. While this can be defined specifically for each collection, it is also necessary to establish a common working framework.

Establishing this framework should involve:

1. Defining what it is that we must document. This means defining what we consider to be heritage within the university system, which is part of a broader discussion that aimed specifically at defining what heritage is. Perhaps once we have established this definition (or definitions), some collections may be removed from the list or merged with others.
2. Defining how the different collections are structured to provide a unitary structure to the university system; in other words, generating a classification/taxonomy framework.
3. Prioritising the tasks that should be included in the master plan, both at a general level and for each collection.
4. Defining the levels of description/selection criteria appropriate for the different collections and whether there are relationships between them.

4.3.1 Types of heritage to be documented

The four main types of university heritage elements defined in 2019 by the Vice-Rectorate for Arts, Culture, and Heritage of the University of Barcelona [39] were adopted for this study, on the understanding that these types of heritage are commonly found in most European universities. The first three were previously identified by UNESCO [38]. The four categories are:

- **Movable cultural heritage.** Includes objects of various kinds, books, and archive documents, as well as scientific collections (mineralogy, zoology, herbaria, etc.).
- **Immovable tangible cultural heritage.** Basically includes buildings and botanical complexes such as the Ferran Soldevila Garden at the UB.
- **Immaterial Heritage.** Elements that reflect the spirit of the institution, such as the university hymn, ceremonials, customs, and traditions [38], as well as associated tangible elements.
- **Human heritage**. The contribution of those people who have left their mark on the institution. Not identified by UNESCO, but widely used in the tourism sector,7F[4] human heritage is fundamental to a university environment, where people are important.

In this article, we focus on movable cultural heritage, leaving specific proposals for the management of other types of heritage for future studies. However, some of the parameters

[4] The WTO (Word Tourism Association) defines cultural tourism as "immersion in natural history, human and cultural heritage, arts and philosophy, and institutions of other countries or regions" [21].

set forth here could also be applied to intangible heritage and immovable cultural heritage. Human heritage clearly requires a separate study.

4.3.2 Classification systems

Once we have broadly defined what we must document, it is essential to link it all together by means of a structure that allows the various collections to be interrelated and offers a united vision, normally of a hierarchical nature. Unfortunately, LAM institutions use different classification systems. For example, in museums and libraries, classification structures are usually based on **subjects** (UDC8F[5] or Dewey in libraries, Nomenclature or AAT9F[6] in museums), while in archives there are three different types of classifications: **organic** classifications (based on the principle of provenance of the documentation), **functional** classifications (based on the competences, functions, and activities of the generating institution) and **organic-functional** classifications (combining the first two concepts) [5–7].

We propose a double classification for university collections, which has been adopted for the UB's virtual museum:

1. An **organic approach** by collection. This approach is necessary in order to keep sight of the context of creation, although, if necessary, groupings by type of heritage could be made at higher levels. With this purpose, the property dcterms:isPartOf was reserved.
2. An **approach by subject and by type of object**. Complementary to the organic approach, this approach has the aim of overcoming the fragmentary vision offered by the first. For example, this approach would recover all items related to veterinary medicine regardless of the collection that hosts them. Of the classification systems analysed, the one best suited to the needs of the University of Barcelona was Nomenclature 4.0 for **classification by subject** and the objects facet of the AAT for **types of objects** (in accordance with CDWA guidelines10F[7]). For object type the field dcterms:type was reserved, and for classification by subject a specific field metadadesUB:classification was created.

4.3.3 Levels of description / cataloguing

It should be borne in mind that the items to be described in a university environment are sometimes isolated items, while others are part of a whole or have parts that are worth cataloguing and always belong to a specific collection. This makes it necessary to consider different levels of description/cataloguing. This is an issue that has been addressed by various GLAM institutions that have designed specific conceptual models and/or descriptive standards to provide a solution. For example, the International Federation of Library Associations (IFLA) has established the LRM (Library Reference Model), which outlines four broad classes of objects to be described: work, expression, manifestation, and item. In this classification, "item" refers to the concrete object held in the library, while "work" is the more abstract idea in the mind of the creator.

In historical archives, the levels of description have traditionally been delimited by the ISAD G, which recommends different levels of description created from documentary groups. The most generic is the fonds and the most specific is the simple documentary unit [20], although there may be numerous intermediate levels. This standard is the basis for the creation of the new conceptual model, Records in Context (ICA RIC-O), which restructures conventional groupings by integrating them into three large groups: RecordSet, Record, and RecordPart. This conceptual model will serve as the basis for the new standard archival description.

In museums, the work is normally considered a material element rather than a work in the mind of the author as it is in the bibliographic world [19]. Furthermore, we often cannot define as many document groups as can be defined in the case of an archival environment. Moreover, the conceptual models associated with museums, such as the CIDOC CRM model11F[8] and the CCO model,12F[9] do not define specific levels; instead, they define objects that contain other objects. In fact, museum metadata schemas do not use the same criteria to establish levels of description. For example, in its *Catalog Level* element, CDWA recognises 19 levels of description, while VRA13F[10]core, only recognises three: work, collection, and image, and LIDO,14F[11] the interchange schema for museums promoted by ICOM (International Council of Museums), distinguishes between individual item, part of an item, and group of items.

Having evaluated the different conceptual models, we propose the adoption of the solution provided by LIDO, as its fewer levels make it more feasible for infrastructures with the limited resources typical of non-museum collections. To its three levels we would add two more: the collection (to collect the usual structure of the university heritage) and

[5] UDC: Universal Decimal Classification.

[6] AAT: Art and Architecture Thesaurus.

[7] CDWA: Categories for Description of Works of Art. Data Structure Standard, created by the Getty Foundation and used mainly in museums.

[8] CIDOC CRM—Conceptual Reference Model from ICOM for managing museums.

[9] CCO: Cataloging Cultural Objects. Data Content Standard for heritage.

[10] VRA: Visual Resources Association. Promoting institution of the VRA Core Standard.

[11] LIDO: Lightweight Information Describing Objects.

the institution, which would be used exclusively to control the physical manifestations of the intangible heritage (photographs, videos, recordings) [8, 44].

Therefore, **in the case of university collections, we would define four main levels of cataloguing or description plus a level of instantiation for intangible heritage**. The cataloguing approach we propose goes from the general to the specific, in accordance with the principles of archival science, to ensure that the higher levels are catalogued or at least inventoried. The five levels are:

Collection: the first thing to be catalogued if nothing else could be, which could also have sub-collections.
Group of items sets of items that for various reasons may be linked and that it is not sustainable to work on separately (e.g. an eighteenth-century doctor's bag that includes all the doctor's working tools inside it). This would constitute a second level in cataloguing, and it would be interesting to consider it as a descriptive unit in very large collections for which it would not be sustainable to catalogue individual items.
Item: a separate item that is considered necessary to catalogue because of its value, although in large collections it is only possible to do this for single items.
Part of the item: only in the case of architecture or very specific items of exceptional value.
Instantiation (intangible heritage only): the registration and cataloguing of intangible heritage have been promoted especially since the 2003 Convention for the Safeguarding of the Intangible Cultural Heritage,15F[12] although the ethnographic aspects of disappearing civilisations were already being collected long before that. However, these intangible elements need to have a physical medium that would make it possible to record their existence, and this medium would have to be catalogued.

This structure of 4 + 1 levels should not undermine the work done by museums, libraries, and archives, as it is possible to map out correspondences between the different levels.

The choice of the level of work for each collection would have to be determined in accordance with:

1. The proposal of the head of the collections, in consultation with experts.
2. Pre-existing tools (inventories or catalogues);
3. The internal structure of the collection;
4. The sustainability of the proposal (are there human, material, and economic resources available to carry out the project?).

The final proposal for the level of description must be approved by the work team (see Sect. 4.1.).

4.3.4 Selection criteria

It is often not sustainable to consider cataloguing at item level due to the size of some collections. For example, at the University of Barcelona, as shown in Table 1, there are different collections with more than 5,000 items [41].

In these cases, it is essential to prioritise. It will be the experts on each collection who will establish the criteria for selecting the items or levels to be documented, based on the existing resources, the age and uniqueness of the items, the demands of users, the state of conservation, and the previous state of the catalogue/inventory.

4.4 Definition of documentation tools and methodologies (how we document)

This section defines **HOW** we document in accordance with the following basic sections:

- Documentation methodology
- Systematisation and standardisation of information

4.4.1 Documentation methodology

This section is a continuation of the section on selection criteria (see 4.3.4.). It involves an in-depth analysis of the needs of each collection and how we can balance scientific rigour and sustainability.

In general, there are three main types of university heritage collections:

- New/non-inventoried collections.
- Collections with total or partial inventory.
- Collections fully catalogued or inventoried.

Depending on the descriptive stage of the collections, the methodology will vary, incorporating more or fewer of the following steps:

1. Establishment of the starting point for each collection.
2. Decision on how the work will continue from now on for each collection, which will involve the following:
3. Drawing up **a work schedule** that prioritises the tasks to be carried out and integrate them into the general schedule.

[12] The "intangible cultural heritage" means the practices, representations, expressions, knowledge, skills—as well as the instruments, objects, artefacts, and cultural spaces associated therewith—that communities, groups, and in some cases individuals recognise as part of their cultural heritage [38].

Table 1 Collections of the University of Barcelona that are featured in the virtual museum (only notable pieces). There are others that have not been incorporated (Criminology, Mathematics, Intangible and Human Heritage, etc.) which will be added later

A list of some UB collections is shown in Table 1: Collection name	Subject	Volume
Centre for Animal Biodiversity Resources (CRBA)	Zoology	152.235 items[a] (56.155 recorded)
Living Plants Collection (Ferran Soldevila Garden)	Botanic	250 taxa
Fine Arts Collection (Faculty of Fine Arts)	Art	1690 items and documents
Art Collection and Research Point «La Relació», de Duoda	Art	9 items
Scientific Tools and Instruments	Science and Technology	962 items inventoried
Collection of Mineralogy	Mineralogy	20,000 hand samples, approximately 7,000 thin sections, and about 3,000 core samples
Center for Plant Biodiversity Documentation (CeDocBiV)	Botanic	400.000 specimens
Catalan Pharmacy Museum	Pharmacy and medicine	4,105 items of materials linked to the world of medicine and pharmacy
Antique or Rare Book Collection (CRAI Fons antic)	Diverse	130,000 items predating the nineteenth century, primarily books
CRAI Pavelló de la República Library	History (Spanish Civil War, Second Republic, and Francoism)	152,000 items that bear witness to the Second Republic, Francoism, the Spanish Civil War, and the Transition (post-Franco period) (posters, banners, coins, badges, etc.)
Josep Artigas Collection	Graphic Arts	100 items related to the work of this poster artist, publicist, and sign writer
Sabater Pi Collection	Ethnology	8992 items
Miquel Porter i Moix Collection	Art (cinema)	Unaccounted photographic and film documentary collection. 4,000 bibliographic documents
Oriol Martorell Collection	Art (Music)	Approximately 9,389 items
Art (UB Collection)	Art	300 items
Architecture	Architecture	10 buildings and architectural complexes that have some level of official protection
UB Lithotheque	Geology	Approximately 5.350 items

[a]The concept of "items" is extracted from internal documents of UB. It refers to individual objects or sets of objects that cannot function separately

4. Deciding on the **tools to be documented** (metadata schemas, controlled vocabularies, programming) and trying to find a common meeting point. In the case of the UB, this meeting point was the institution's new virtual museum. See 4.4.2.
5. Establishing the **depth of the description** (for systems that are not inventoried/catalogued or are only partially inventoried/catalogued). Traditionally, both archival science and museology distinguish between the concepts of **inventory** (a more superficial description and/or with a recording function) and **catalogue** (more in-depth, limited to collections or single items, and with a strong influence on the historical or contextual aspect), although the definitions tend to vary slightly depending on the field of knowledge [4]. However, in a heritage environment as diverse as the university, these distinctions can become very ambiguous, which is why we advocate Santana's proposal [35] to collapse these terms into the concept of *General Cultural Heritage Registers.* Based on this single model of register, we would need to decide whether to do an **in-depth**, research-oriented **cataloguing** (a "catalogue" in the conventional sense), or **a quick cataloguing** that captures the fundamental characteristics of the item (or group of items) to be documented, registers it, and prepares it for a more in-depth cataloguing. It would even be possible to simply consider entering a minimal **record**, which is little more than a code and a title, simply to

provide evidence of the material existence of a heritage element.

However, an important question to consider is that in many centres the university heritage (HERITAGE) is linked to a more general heritage unit (ESTATE), which inventories all university properties. In these cases, it would be advisable to analyse the inventory needs of this unit and ensure that our records have the necessary fields, to ensure compatibility between all databases.

5. Taking the following actions in the case of systems that are already fully inventoried or previously catalogued in computer format:
   i. apply clean-up techniques and reconcile metadata with Open Refine [48] systems to avoid inconsistencies between the controlled languages that would have to serve all collections and the languages that have been used so far, if any; and
   ii. design/adapt crosswalks to facilitate the entry of records and their collection within the virtual museum.
2. Considering a cataloguing approach that integrates the semantic web and Linked Data as basic operating principles. This entails, among other things, treating each document grouping as a resource and therefore generating a permanent IRI17F[13] for each one, searching for documentary languages that are adapted to LOD and applying them, and taking advantage of the data they contain for the semantic enrichment of our records [2].

4.4.2 Systematisation and standardisation of documentation

As noted above, one of the basic objectives of our project is to define the tools for documentation. The selection of tools should be subject to the following guidelines:

Adopt the most widely accepted standards for documentation. These tools will be evaluated by the project management team to assess their suitability. For this task, we will use Boughida's terminology, as quoted by Gilliland [18], to characterise the different types of standards to be defined: **data structure standards are** metadata schemas used in a community comprising the structures of relationships, records, and fields that a database will have; **data value standards** are those controlled languages used within a community to populate certain fields of metadata schema; **data content standards** facilitate the definition of the format of the values; and **data format/technical interchange standards** are those formats that allow us to save the records and share them with other institutions (XML, JSON, CSV, etc.).

Table (Table 2) contains some examples of standards adopted by GLAM institutions in each category, bearing in mind that the list is not exhaustive, and the examples shown do not all fall entirely into a single category.27

If each institution has its own standards, which ones should we use in a university environment? We will try to answer this question in the following sections.

- **Produce quality metadata.** Good metadata are essential to be able to pursue quality research. If the metadata are poor, the research results will be poor as well [10, 11, 16, 18, 26, 27, 37].
- **Produce metadata that adapt to different types of users.**
- **Produce metadata** that can be easily located, accessed, reused, and shared. In other words, they must comply with the **FAIR** principles,18F[14] which, although they were created in a scientific environment, are fully applicable here [46]. A final objective should be the aggregation of data for submission to large institutional repositories such as DPLA, Europeana or GBIF.19F[15]
- Create/adapt a metadata schema or application profile that is **modular and extensible**, adapted to the diversity of the collections and developed from a **simple conceptual model** that provides the information with a basic structure.
- Ensure that **the metadata can be developed by people outside libraries and archives who are not information specialists**, but who already do an excellent job in their fields on the advice of information professionals. This means that the system chosen must allow the records corresponding to the different collections to be edited and added in a simple way.
- Facilitate **metadata ingestion** from other systems, such as libraries, archives, legally established museums, and collections that have developed their own systems. To this end, it is essential that crosswalks be developed to act as bridges between old systems and the new one.

This section can be summarised as follows (Fig. 4):

The next section outlines our proposals for non-museum university collections, in terms of both conceptual models and standards.

Conceptual data models for non-museum university collections

The conceptual models existing in the GLAM environment are characterised by a high level of complexity that

[13] IRI: Internationalized Resource Identifier.

[14] FAIR: Findable, Accessible, Interoperable, Re-usable.

[15] GBIF: Global Biodiversity Information Facility.

Table 2 Metadata standards

Standards	Libraries	Archives	Museums	Humanities, arts and social and applied sciences collections	Natural sciences collections
Structure standard	MARC	EAD/EAC	CDWA VRA Spectrum Other	CDWA VRA Own metadata schemas	DARWIN CORE, ABCD, ABCD_EFG Own metadata schemas
Content standard	RDA/AACR	ISAD G and national standards DACS	CCO	No uniformity in use	No uniformity in use
Value standards	LCSH, VIAF, UDC, DDC	No uniformity	ULAN, TGN, ICONCLASS, AAT	No uniformity in use	No uniformity in use
Exchange standards	Dublin Core, LIDO, CDWA Lite, VRA, Darwin Core (scientific collections)				
Coding standards	XML, txt, JSON, JSON-LD, CSV, RDFa, Turtle				
Conceptual model	FRBR	RiC (Records in Context)	CIDOC-CRM CCO	No uniformity in use	No uniformity in use

Fig. 4 Requirements for metadata in university collections. Source: Authors

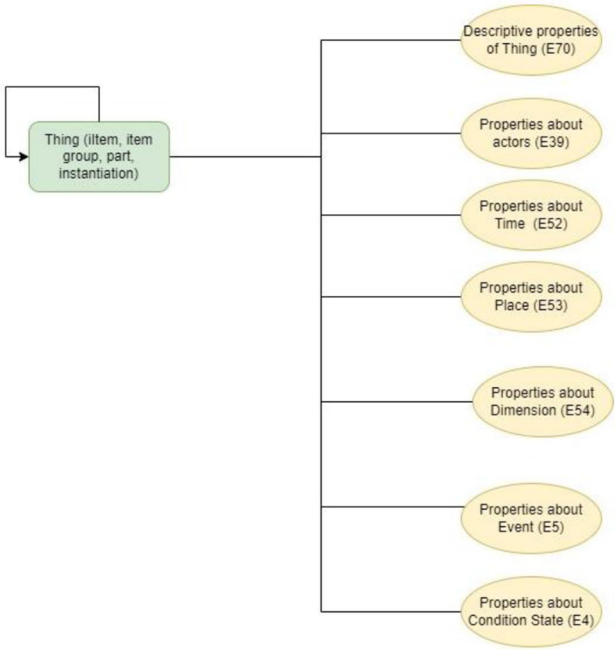

Fig. 5 Our proposal: E-R Non-Museum conceptual model

can only be sustained by institutions with sufficient staff and resources. Our proposal would be based on a simplified version of the CIDOC CRM model,20F[16] which we consider to be the most appropriate for cultural heritage as a whole because of its treatment of the day-to-day management of items. The difference is that while CIDOC CRM defines a series of basic classes, subclasses, and properties, our proposal includes most of the basic entities of the CIDOC CRM model but transforms them into properties to be reported (Fig. 5).

Data structure standards for non-museum university collections

Our proposal would be to use **interchange metadata schemes** for the daily management of university collections, while maintaining the original metadata schemas in those

[16] The complete model was applied in the first version of the University of Barcelona's virtual museum, but it proved to be unsustainable in the environment. The people in charge of the collections ended up abandoning it (they continued to use the old systems) or substantially simplifying the cataloguing method.

environments with consolidated standards. These schemas are used for the following reasons:

1. Their learning curve is low. With a minimum of training and good documentation, it is possible to achieve the documentation objectives set out for each collection.
2. They are supported by most programming that is based on more complex conceptual systems. This would allow the library, archives, and museums (if any) to work with the standards specific to their environment and periodically feed the common system with the processing of new datasets.
3. They allow scaling up to national and international cooperation environments, if the appropriate systems and protocols are in place to do so (collective catalogues, international repositories).
4. They are usually conceptually prepared for an LOD environment. The schemas we propose in this article are mostly application ontologies [49].
5. They usually include properties or attributes that are well adapted to the needs of the proposed conceptual model.
6. They facilitate the creation of specific application profiles for certain environments. This would allow us to include those properties required by our conceptual model, but which do not appear by default.
7. They usually already have crosswalks/maps that facilitate migration between schemas.

Data structure standards for cultural heritage

There are two main Data Structure Standards that we have considered using for collections of tangible and intangible cultural heritage.21F[17] The first is Dublin Core, which has emerged from the library environment for the cataloguing of electronic resources on the Internet and is now consolidated at the level of data interchange [26], since it is the basis for migrating data to a multitude of digital repositories, such as Europeana or DPLA. The second is LIDO, which developed out of the museum environment (ICOM) and is closely linked to CDWA and the CIDOC-CRM ontology, of which it is an expression.

We ruled out other specific systems for museums/heritage materials, such as VRA Core and CDWA Lite, because they are superseded by LIDO, which is more of an institutional initiative through ICOM.

However, LIDO has a narrower range of application than Dublin Core, as reflected in Waldron & Webster's report on VRA cataloguing and metadata [43], because it is used exclusively in a museum context and is a more complex system, although numerous initiatives have been established to promote its use [1].

[17] The article does not study human heritage, which, due to its complexity, deserves an in-depth analysis of its own.

With the above in mind, **our proposal focuses on creating an application profile based on Dublin Core for the** following reasons:

1. It is easy to use and easily adapted by collections managers, which should be our basic objective.
2. It is very well established and is the basis for large repositories (Europeana, DPLA).
3. It allows data to be made ready for scaling and migration to other systems.
4. It allows data to be prepared for LOD environments, as reflected in the various guides prepared by the *Dublin Core Metadata Initiative* [15] and in the fact that it is currently constituted as a formal ontology.
5. It allows the establishment of different relationships between the items using the properties *relation*:isPartOf and relation:hasPart. This will allow the implementation of our levels of description (see 4.3.3.).
6. It encourages the creation of application profiles that allow us to add management aspects of those recognised by ICOM/CIDOC and that we have added to our conceptual model.

The proposed application profile (see "Appendix 1") offers a flat structure for the user that connects with the different parts of our adaptation to the CIDOC CRM conceptual model.

Data structure standards for scientific heritage

Despite its advantages, Dublin Core does not adapt well to natural science collections, which are very common in the university environment, as they have very specific characteristics, such as the existence of taxonomies (botany, zoology) and/or of multiple descriptive fields (mineralogy).

Our analysis of existing metadata schemas has led to the identification of three main types of approaches:

- **Museum approach**. Particularly noteworthy is the case of Italy, where the ICCD (Istituto Centrale per il Catalogo e la Documentazione) has developed specific cataloguing standards for botany, mineralogy, petrology, planetology, palaeontology, and zoology [32]. These are the basis for the *Catalogo Generali dei Beni Culturali*, which compiles Italy's impressive cultural and scientific heritage. Other initiatives can be found in Canada (Chin Natural Sciences Data Dictionary) and ICOM itself, which has developed CRMsci, an extension of the CIDOC-CRM model for science that is still in the draft phase.
- **Natural sciences approach.** There are numerous metadata schemas that have been designed to deal with very specific areas (e.g. agriculture, mineral resources, biodiversity, and ecology). Many of these are collected at the DCC (Digital Curation Centre). Because of their presence in a university environment and in collections

management, the structural standards developed by the Biodiversity Information Standards (TDWG)22F[18] are worth highlighting here, particularly the following:

- Darwin Core, a Dublin Core application profile, ratified as a standard in 2009, which has been used for cataloguing zoological collections to record taxonomies. Its purpose, however, is the cataloguing of biodiversity, not the cataloguing of specific items in a museum environment. In other words, what is important in this schema is the event (the sighting of a bear in the Pyrenees, for example), rather than the item itself. Darwin Core is currently used to feed international biodiversity repositories, such as GBIF, in which many universities participate.
- ABCD (Access to Biological Collection Data), initiated as a project in 2000 and consolidated as a standard in 2005, which allows the collection of information on natural science collections and specific specimens. It feeds the BioCase and GBIF portals and has developed crosswalks to Darwin Core that allow interoperability between the two schemas. It also allows data to be exported to LIDO.
- ABDC_EDF (extension for geosciences), was developed in 2005 out of ABDC, since there was a lack of standards in the field of earth sciences. This metadata schema is the one used in the GeoCASE portal, the equivalent of GBIF in the field of geoscience.

None of the above schemas allow museum management; at most there are elements linked to conservation and preservation within schemas linked to palaeontology.

- **Other approaches.** There are approaches that do not fall into either of the two categories above, but which can be very interesting for the construction of databases. For example, Wikipedia's infoboxes [45] have very elaborate templates that could be useful as metadata schemas in the natural sciences, as shown in Fig. 6.

It should be noted that although at the time of writing Schema.org had not yet developed metadata for the natural sciences, its unstoppable growth suggests that at some point this will become an option.

In addition, there is a whole series of approaches that could be described as "hand-made", something very common in university collections. Many of these began being automated when there were no metadata standards, using simple resources such as Microsoft Excel or Microsoft Access. Some attempts have also been made to generate application profiles of widely used metadata schemas, such as the case of the Colorado School of Mines [9], which adapted Dublin Core to the cataloguing of minerals.

From our point of view, based on the principles outlined in Sect. 4.4.2., there is a need for collections in the natural sciences, applying a metadata schema that is simple but that can also be migrated to established standards. We consider two options to be worth considering:

1. Adapt the environment's own systems (ABCD, ABCD_EFG and Darwin Core) in accordance with the managers' needs and with the pre-existing metadata schemas of the various collections.
2. Adapt Dublin Core to the needs of scientific collections.

In both cases, it would be necessary to add the fields that have been determined for the museum management of the items (it is important to bear in mind that these types of items have a great deal of movement due to exchanges, donations, loans, and transfers) and the necessary mapping to be able to enter the records into major international repositories.

From our point of view, the first is the better option, as metadata schemas are more appropriate. However, the use of all of the fields may exceed the capacities of cataloguers, and may also often be unnecessary. This question requires in-depth analysis in the future research.

Data content standards

If Dublin Core is chosen as the standard, we have a problem of content, as it avoids "forcing" and only gives very generic recommendations that do not have to be followed [15]. The original purpose of such flexibility was to facilitate the implementation of the standard, but it can lead to major inconsistencies in the data. It is therefore essential to limit the values that users can write, either through Data Value Standards or through Data Content Standards.

For content standards, there are three main options:

1. Adapt the content standards specific to one of the GLAM institutions, which are listed in Table 2.
2. Use the content recommendations of specialised structural standards, such as CDWA, which are so complete that in many cases these same standards can serve as a content standard.
3. Use specific international standards for specific aspects in certain fields, such as ISO standards for formatting dates or geographical coordinates.

Our proposal is to adopt the content standards recommended by Dublin Core. However, if they are not suitable, or if no standard is recommended, we propose the use of the content recommendations of the museum standards (i.e. CDWA and CCO), filtering them to adapt them to the characteristics of our collection and our software. In other words, the proposal is to:

[18] TDWG: Taxonomic Databases Working Group.

```
{{Infobox mineral
| name        =
| boxwidth    =
| boxbgcolor  =
| image       =
| imagesize   =
| alt         =
| caption     =
| struct image     =
| struct caption   =
| struct imagesize =
| struct2 image    =
| struct2 caption  =
| struct2 imagesize=
| SMILES           =
| Jmol             =
| category    =
| formula     =
| molweight   =
| strunz      =
| dana        =
| system      =
| class       =
| symmetry    =
| unit cell   =
| color       =
| colour      =
| habit       =
| twinning    =
| cleavage    =
| fracture    =
| tenacity    =
| toughness   =
| mohs        =
| luster      =
| streak      =
| diaphaneity =
| gravity     =
| density     =
| polish      =
| opticalprop =
| refractive  =
| birefringence =
| pleochroism =
| 2V          =
| dispersion  =
| extinction  =
| length fast/slow =
| fluorescence =
| absorption  =
| melt        =
| Curie temp  =
| fusibility  =
| diagnostic  =
| solubility  =
| impurities  =
| alteration  =
| other       =
| prop1       =
| prop1text   =
| references  =
| var1        =
| var1text    =
| var2        =
| var2text    =
| var3        =
| var3text    =
| var4        =
| var4text    =
| var5        =
| var5text    =
| var6        =
| var6text    =
}}
```

Ruby

A naturally occurring ruby crystal

General	
Category	Mineral variety
Formula (repeating unit)	aluminium oxide with chromium, Al_2O_3:Cr
Crystal system	Trigonal
Crystal class	Hexagonal scalenohedral ($\bar{3}$m) H-M symbol: ($\bar{3}$ 2/m)
Space group	$R\bar{3}c$ (No. 167)
Unit cell	unit cell
Identification	
Color	Red, may be brownish, purplish or pinkish
Crystal habit	Varies with locality. Terminated tabular hexagonal prisms.
Cleavage	No true cleavage
Fracture	Uneven or conchoidal
Mohs scale hardness	9.0
Luster	Vitreous
Streak	white
Diaphaneity	transparent
Specific gravity	4.0
Refractive index	n_ω=1.768 - 1.772 n_ε=1.760 - 1.763, Birefringence 0.008
Pleochroism	Orangey red, purplish red
Ultraviolet fluorescence	red under longwave
Melting point	2044 °C
Solubility	none
Major varieties	
Sapphire	Any color except red and violet
Oriental amethyst	violet color
Corundum	various colors (mineral species)
Emery	Granular (mineral mixture)

Fig. 6 Mineral infobox template and example. Source: https://en.wikipedia.org/wiki/Template:Infobox_mineral

1. Develop a data dictionary/metadata record in which the characteristics stated for each of the properties are briefly adapted to CDWA/CCO (see 4.6).
2. Link the data dictionary to the specific point of CDWA or CCO where this property is treated in case the cataloguer wants to obtain more information.

Data value standards

Value standards, also known more recently as KOS (Knowledge Organisation Systems), indicate the accepted values for a given property. In their simplest version, they constitute what we call lists of terms, but they can include complex language systems such as thesauri. They work in conjunction with content standards because when a term does not appear in the list of values, it is then necessary to construct the term anew using the recommendations given by the content standard.

It is important to emphasise that we must avoid creating our own lists as much as possible. There are various vocabularies available today that are adapted to the LOD environment, which can be perfectly useful for our purposes. In the field of cultural heritage, international tools such as VIAF23F[19] (people and institutions), Nomenclature 4.0 (classification), Geonames (geographical locations) or AAT (materials or types of objects) are highly recommended. The use of these tools involves something of a learning curve, but it is better for collections managers to learn how to use these tools rather than spending time and effort creating new ones. However, there may be exceptions to this rule, such as:

In some cases, where there are well-developed local tools, their use may be preferable to global standards. For instance, at our centre, rather than using AAT directly, we chose to use the UB Thesaurus, developed by the university library for subjects, as it was already adapted to the linked data, had links to international standards and was fully adapted to our needs in terms of concepts. Moreover, if it was necessary to add new terms, the library could add them without any problem. On the other hand, for creators we opted for VIAF, which was also used by the library itself.

2. In some cases, where the list may be shorter than it is for subjects, it may be useful to make a pre-selection of terms to give more guidance to cataloguers. For example, in our case, in terms of materials and techniques, the AAT was adequate, but a selection was made of the terms that might appear more frequently as a way of facilitating the work and ensuring consistency. The same was done with

[19] VIAF: Virtual International Authority File.

Fig. 7 Summary of procedures for non-museum university collections

1. Define the working team.	Libraries, archives, or metadata experts in Library and Information Science Faculties must lead the metadata project.
2. Define the typologies of heritage and the collections they contain.	
3. Define the classification that we will apply to our instituttion.	Nomenclature 4.0 is suggested to functionally classify the items. A metadata schema property must be reserved for it.
4. Evaluate each collection of the collections and define:	Levels of description. Selection criteria. Work schedule.
5. If suitable, adapt Dublin Core for cultural and intangible heritage.	Fem application profile, if it is affected by certain collections, the specific elements demanded. Choose content standards for certain properties. Choose value standards for certain properties.
6. Adapt other schemas for natural heritage and human heritage.	
7 Look for a system that supports:	Dublin Core with enough flexibility to create application profiles. Interoperability. Linked Open Data. IIIF Virtual Reality 3D simulations
8. Create a manual and general data dictionary for the entire collection.	Create an annex with the specifics of each collection.
9. Train all actors and those involved	

the levels of Nomenclature 4.0, which are too detailed for our needs.

4.5 Before starting to catalogue: the choice of computer system

There are numerous systems, including open-source systems, that may be suitable for our purposes and allow us to implement the designed application profile and the selected standards. Alcaraz offers a good comparison of some of these systems [6]. Although it is not the aim of this article to indicate which system should be used, we can point out some of the features it should have:

- Support for different metadata schemas, especially Dublin Core.
- Integrated mappings to other metadata standards.
- Capacity to create application profiles from certain metadata schemas.
- Capacity to define your own controlled values and to ingest externally controlled values.
- Support for OAI-PMH,24F[20] the protocol that will allow aggregators to collect your records and send them to Europeana.
- Support for image management using IIIF.25F[21]
- Use of semantic web standards, which should allow for features such as: introduction and application of ontologies, IRI generation for resources, connection to LOD datasets, SPARQL26F[22] EndPoint, semantic enrichment options, etc.

[20] OAI-PMH: Open Archive Initiative-Protocol for Metadata Harvesting.

[21] IIIF: International Image Interoperability Framework.

[22] SPARQL: SPARQL Protocol and RDF Query Language.

Table 3 Data dictionary

Identifying	
Identifier	http://purl.org/dc/terms/identifier
Version	V1
Naming	
Name	Dcterms:type
Label	Work type
Context description	In the LAM environment, to organise by object type is a basic element in all institutions. In the context of university heritage, it acquires even more importance because it allows elements of very diverse origins to be grouped together, even if the institutions are different. For example, we can group all the books of a pharmacy museum and retrieve them together with the documents hosted by the library, the archive, or another university collection. The same would apply to scientific instruments such as microscopes or with historical academic documents, to name a few examples
DEFINITION	
Definition	Worktype of the item being catalogued
ADMINISTRATIVE	
Comments	For any question contact museuvirtual@ub.edu
Status	Active
Responsible organisation	Vice-rector of Arts, Culture and Heritage. University of Barcelona
RELATIONAL	
Classification Scheme	Dublin Core
Related metadata schemas	–
Other	
Obligation	Mandatory
Datatype	String, URI
Multiplicity	1..*
Value Domain	AAT Thesaurus. Catalan version available at Tesaurus d'Art i Arquitectura (gencat.cat)
Explanatory comment	1. 1. Select the most appropriate term from the list 2. If you have several options, try to use the most specific term possible. For example, if you have a binocular microscope, use the term Binocular Microscope, not simply Microscope 3. If the term does not appear in the list, you can consult the original thesaurus, Tesaurus d'Art i Arquitectura (gencat.cat), to see if the term exists. If you cannot find it or it is not translated, consult its English version (https://www.getty.edu/research/tools/vocabularies/aat/) or its Spanish version (Tesauro de Arte Arquitectura | (aatespanol.cl)). You can submit the translated term/new term to museuvirtual@ub.edu to have it incorporated, if applicable, in the list of accepted terms
Examples	Microscope Manuscript

- User interfaces that are easy to understand, visual, and customisable.
- Importing systems that allow you to work with the main serialisations of the work environment. Normally, CSV, XML (and its variants), JSON, or Turtle.

4.6 Before starting to catalogue documenting the process

The end point of the whole system should be a document containing all the decisions relating to the final database. Depending on the environment, this document will have different names. Computer scientists call them **data dictionaries** [13], while information specialists, especially in a semantic web environment, refer to them as **metadata registries** [49]. Other names include **systems catalogues**, **glossaries**, and **database manuals**. But whatever the name, this document must be a tool that helps collections managers to enter the data, property by property, keeping doubts to a minimum and in the simplest possible way, without the need for a potentially overwhelming level of technical knowledge. For example, in a field in which users must enter URIs that facilitate subsequent semantic enrichment, we do not need to explain in detail what this enrichment is, but we do need to explain how we want them to enter the data.

We propose a data registry/dictionary model based on ISO 11179 -2013 [23], a standard that supports the development of metadata registries and is accompanied by some extreme elements of Coronel and Morris' data dictionary [13]. Table 3 presents an example of this.

4.7 Summary of the work methodology

See Fig. 7.

5 Conclusion

The proposal for the documentation of university heritage outlined in this article is designed to respond to the following unquestionable facts:

- The metadata must be of high quality.
- The metadata must comply with the FAIR principles.
- There are numerous tools and standards for quality metadata, but they are often sectorial in nature.
- University heritage is highly interdisciplinary.
- The people in charge of most university collections work part-time and need simple and very targeted systems.

Consequently, we propose the creation of an application profile based on Dublin Core, the most widely used of the structural interchange standards, while preserving elements of the CIDOC CRM conceptual model, which will become properties of this profile. We also propose the use of value and content standards specific to the sector to populate the properties that make up the profile.

We believe that our proposal allows these professionals to work part-time but does not prevent full-time library and archive professionals from submitting their records to the central heritage system.

To manage this system, we propose the creation of a steering committee made up of various university professionals under the direction of the relevant vice-rectorate. Decisions on metadata should be left in the hands of information professionals, be they librarians, archivists, or professors in the university's information science faculties. The composition of the steering committee and its functions should probably be set out in the regulations of the university's heritage unit as a way of consolidating its existence.

We believe that our proposal, summarised in point 4.7, offers a sustainable approach to the documentation of university collections, although it only focuses on one specific type of heritage, a limitation that will need to be addressed in subsequent studies.

6 Proposals for the future

Although this paper focuses mainly on tangible cultural heritage, it acknowledges that there are other types of heritage that need to be looked at in greater depth. For example, as far as natural science collections are concerned, Darwin Core could be a suitable system for zoological and botanical collections (many of which already collaborate with GBIF), but further study is necessary to identify the right system for geosciences and other natural sciences. A more in-depth study would therefore be needed to determine whether a common application profile could be used for all natural science fields.

Human heritage is a more complex question. This type of heritage is not defined by UNESCO and has been discussed mainly in the field of tourism. Research on how to deal with this type of heritage is still in its infancy, as it is difficult to adapt it to standard scientific criteria. It is worth considering, however, that the analysis would probably have to start with metadata standards that collect biographical data on people, such as Schema.org's Person (or its predecessor, FOAF), ISAAR CPF, or CDWA's Person/Corporate Body Authority.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. There are no financial relationships with third parties.

Declarations

Conflict of interest There are no conflicts of interest since our work is carried out in a public institutional environment without the intervention of third parties in design and development.

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/.

Appendix 1. Proposed application profile for non-museum university institutions

Label	Curie	CDWA equivalence	CIDOC CRM classes
Creator	dcterms:creator	4.1.3. Creator Identity	Actor
Publisher or manufacturer	dcterms:publisher	4.1.3. Creator Identity	Actor
Rights	dcterms:rights	22.1. Copyright Statement	Actor
Cataloger name	notMuseumUC:autoriaRegistre	25.2. Cataloger Name	Actor
Item condition	notMuseumUC:condition	14.1. Condition/Examination Description	Condition State
Remarks	notMuseumUC:remarks	1.5. Remarks	Thing
Dimensions	dcterms:extent	6.1. Dimensions Description	Dimensions
Monetary valuation	notMuseumUC:valuation	23.3.1. Valuation	Dimensions
Method of acquisition	dcterms:accrualMethod	23.2. Transfer Mode / Method of Acquisition	Event
Date of acquisition	dcterms:available	Not found	Event
Previous owner	dcterms:source	23.5. Owner/Agent	Event
Conservation/treatment history	notMuseumUC:treatmentHistory	15.1.Conservation/Treatment Description	Event
Item legal status	notMuseumUC:legalStatus	23.4. Legal Status	Event
Exhibitions/loans	lido:displayEvent	24.1. Exhibition/Loan Description	Event
Place of origin	dcterms:provenance	4.3. Creation Place/Original Location	Place
Current location	dcterms:spatial	21.2.1. Current Flag	Place
Specific location	lido:repositoryLocation	21.2.4. Gallery/Specific Location	Place
Contains	dcterms:hasPart	20.1. Related Work Label/Identification	Thing
Collection	dcterms:isPartOf	Doesn't found	Thing
Identifier	dcterms:identifier	21.2.3. Repository Numbers	Thing
Title	dcterms:title	3.1. Title Text	Thing
Object type	dcterms:type	1.2. Object/Work Type	Thing
Subject	dcterms:subject	16.2. General Subject Terms	Thing
Classifications	lido:classification	2.1. Classification Term	Thing
Description	dcterms:description	18.1. Descriptive Note Text	Thing
Materials or techniques (display)	dcterms:format	7.1. Materials/Techniques Description	Thing
Materials or techniques (controlled terms)	lido:termsMaterialsTech	7.2. Materials/Techniques Flag	Thing
Language	dcterms:language	8.5. Inscription Language	Thing
Bibliographic references	dcterms:isReferencedBy	27.2.2. Source Full Citation	Thing
Other catalogues	dcterms:relation	27.2.2. Source Full Citation	Thing
Date	dcterms:date	4.2. Creation Date	Time-span
Earliest date	lido:earliestDate	4.2.1.E.arliest Date	Time-span
Latest date	lido:latestDate	4.2.2. Latest Date	Time-span
Styles/periods/groups/movements	dcterms:temporary	5.2. Styles/Periods Indexing Terms	Time-span

Label	Curie	Data value standards	Data content standards
Creator	dcterms:creator	Local value standards VIAF or ULAN	CDWA CCO
Publisher or manufacturer	dcterms:publisher	Local value standards VIAF or ULAN	CDWA CCO
Rights	dcterms:rights	Not controlled vocabulary	CDWA CCO
Cataloger name	notMuseumUC:autoriaRegistre	Not controlled vocabulary	CDWA CCO
Item condition	notMuseumUC:condition	Local Value Standards CDWA recommended terms	CDWA CCO
Remarks	notMuseumUC:remarks	Not controlled vocabulary	Free text
Dimensions	dcterms:extent	Not controlled vocabulary	CDWA CCO
Monetary valuation	notMuseumUC:valuation	Not controlled vocabulary	CDWA CCO
Method of acquisition	dcterms:accrualMethod	Local value standards CDWA AAT	CDWA CCO
Date of acquisition	dcterms:available	Not controlled vocabulary	ISO 8601
Previous owner	dcterms:source	Local value standards VIAF or ULAN	CDWA CCO
Conservation/treatment history	notMuseumUC:treatmentHistory	Not controlled vocabulary	CDWA CCO
Item legal status	notMuseumUC:legalStatus	Local value standards CDWA	CDWA CCO
Exhibitions/loans	lido:displayEvent	Not controlled vocabulary	CDWA CCO
Place of origin	dcterms:provenance	Local value standards TGN, Geonames	CDWA CCO
Current location	dcterms:spatial	Local value standards TGN, Geonames	CDWA CCO
Specific location	lido:repositoryLocation	Not controlled vocabulary	CDWA CCO
Contains	dcterms:hasPart	Not controlled vocabulary	URI
Collection	dcterms:isPartOf	Not controlled vocabulary	URI
Identifier	dcterms:identifier	Not controlled vocabulary	CDWA CCO
Title	dcterms:title	Not controlled vocabulary	CDWA CCO
Object type	dcterms:type	Local value standards AAT	CDWA CCO
Subject	dcterms:subject	Local value standards AAT	CDWA CCO
Classifications	lido:classification	Local value standards Nomenclature 4.0	CDWA CCO
Description	dcterms:description	Not controlled vocabulary	CDWA CCO
Materials or techniques (display)	dcterms:format	Not controlled vocabulary	CDWA CCO
Materials or techniques (controlled terms)	lido:termsMaterialsTech	Local value standards AAT	CDWA CCO
Language	dcterms:language	ISO 639.3	CDWA CCO
Bibliographic references	dcterms:isReferencedBy	Not controlled vocabulary	ISO 609, CDWA, CCO or similar
Other catalogues	dcterms:relation	Not controlled vocabulary	URI
Date	dcterms:date	Not controlled vocabulary	CDWA CCO
Earliest date	lido:earliestDate	Not controlled vocabulary	CDWA CCO
Latest date	lido:latestDate	Not controlled vocabulary	CDWA CCO
Styles/periods/groups/movements	dcterms:temporary	Local value standards AAT	CDWA CCO

References

1. Agenjo Bullón, X., Hernández-Carrascal, F.H., Hernández Rubio, S.: El formato LIDO, puente para la integración de bibliotecas y museos en el entorno Linked Open Data y Europeana. In: Terceras Jornadas de Bibliotecas de Museos, pp. 29–46. Ministerio de Educación, Cultura y Deporte, Madrid (2015). http://www.digibis.com/images/PDF/bimus-formato-lido-en-Actas_III_Jornadas-p.pdf
2. Agenjo-Bullón, X., Hernández-Carrascal, F.: Authority records, semantic enrichment and Wikidata. ThinkEPI Yearbook **12**, 361–372 (2018). https://doi.org/10.3145/THINKEPI.2018.61
3. Alam, S.L., Sun, R., Campbell, J.: Helping yourself or others? Motivation dynamics for high performing volunteers in GLAM crowdsourcing. Australasian J. Inf. Syst. (2020). https://doi.org/10.3127/ajis.v24i0.2599
4. Alba, E.: Catalogue and inventory as instruments for the management of cultural heritage. In: Educación y Entorno Territorial de La Universitat de València, pp. 67–93. Universitat de València, Valencia (2014). https://roderic.uv.es/handle/10550/35409. Accessed 29 October 2022.
5. AlberchiFugueras, R., BoadasiRaset, J., CapelliGarriga, E.: Manual d'arxivística i gestió documental. Associació d'Arxivers de Catalunya, Barcelona (2009)
6. AlcarazMartínez, R.: Omeka S as an alternative for the development of digital collections and digital humanities projects. BiD (2022). https://doi.org/10.1344/BiD2022.48.06
7. Alonso Martínez, J.A.: Archivística y Gestió Documental. UOC, Barcelona (2011)
8. Artese, M.T., Gagliardi, I.: Integrating, indexing and querying the tangible and intangible cultural heritage available online: the QueryLab portal. Information **13**(5), 260 (2022). https://doi.org/10.3390/info13050260
9. Baker, C.: Metadata and Minerals: A Library—Museum Pilot Project. Dissertation. In: Digital Initiatives Symposium (2019). https://digital.sandiego.edu/symposium/2019/2019/16/. Accessed 30 Oct 2022.
10. Candela, G., Escobar, P., Carrasco, R.C., et al.: (2022) Evaluating the quality of linked open data in digital libraries. J. Inf. Sci. **48**(1), 21–43 (2022). https://doi.org/10.1177/0165551520930951.Accessed30October
11. Choi, J., Fiorino-Iannace, G.: Information management systems at the metropolitan museum of art. In: Angel, C.M., Fuchs, C. (eds.) Organization, Representation and Description through the Digital Age, pp. 117–129. De Gruyter-Saur, Berlin (2018)
12. CIDOC Documentation Standards Working Group.: Statement of principles of museum documentation. ICOM CIDOC (2012). https://cidoc.mini.icom.museum/wp-content/uploads/sites/6/2020/03/principles6_2.pdf. Accessed 30 Oct 2022.
13. Coronel, C., Morris, S.: Database Systems: Design, Implementation, & Management. Cengage Learning, Boston (2018)
14. Dempsey, L. Scientific, Industrial, and Cultural Heritage: a shared approach: a research framework for digital libraries, museums and archives. Ariane: Web Magazine for Information Professionals 12 January (22), p. 15 (2000). http://www.ariadne.ac.uk/issue/22/dempsey/. Accessed 30 October 2022.
15. Dublin Core Metadata Initiative (DCMI): DCMI: Dublin CoreTM User Guide. DCMI, [s.l.] (2019) https://www.dublincore.org/resources/userguide/. Accessed 30 Oct 2022
16. Farneth, D.: How can we achieve GLAM? Understanding and overcoming the challenges to integrating metadata across museums, archives, and libraries: part 2. Catal Classif Q **54**(5–6), 292–304 (2016). https://doi.org/10.1080/01639374.2016.1192078
17. GarcíaFernández, I.M.: University museums in Europe. Challenges and initiatives. Cuadernos de Arte de La Universidad de Granada **49**, 11–32 (2018). https://doi.org/10.30827/caug.v49i0.7750
18. Gilliland, A.J.: Setting the stage. In: Baca, M. (ed.) Introduction to Metadata, chap 2, 3rd edn., pp. 1–19. Getty Research Institute, Los Angeles (2016)
19. González Mello, R.: Information, data and metadata for cultural heritage conservation. Intervention **9**(17), 6–20 (2018)
20. ICA.: ISAD(G) General International Standard Archival Description: adopted by the Committee on Descriptive Standards, Stockholm, Sweden, 19–22 Sept 1999. ICA, Ottawa (2000). https://www.ica.org/sites/default/files/CBPS_2000_Guidelines_ISAD%28G%29_Second-edition_EN.pdf. Accessed 30 Oct 2022.
21. Inskeep, E.: National and Regional Tourism Planning: Methodologies and Case Studies. Routledge, London (1994)
22. ISO.: ISO 15489-1:2016-Information and documentation-Records management-Part 1: Concepts and principles. ISO, Geneva (2016). https://www.iso.org/standard/62542.html. Accessed 30 Oct 2022.
23. ISO.: ISO/IEC DIS 11179–30(en), Information technology—Metadata registries (MDR)—Part 30: Basic attributes of metadata (Draft). ISO, Geneva (2022). https://www.iso.org/obp/ui/#iso:std:iso-iec:11179:-30:dis:ed-1:v1:en. Accessed 30 Oct 2022.
24. Kozac, Z.R.: Promoting the past, preserving the future: British university heritage collections and identity marketing. PhD Thesis, University of St Andrews (2007)
25. Marín Torres, M.T.: University museums in Spain: towards greater visualisation and dissemination. Cuadernos de Arte de La Universidad de Granada **49**, 89–108 (2018). https://doi.org/10.30827/caug.v49i0.7754
26. Miller, S.J.: Metadata for Digital Collections: A How-to-Do-it Manual. Facet Publishing, London (2022)
27. Mitchell, E.: Metadata developments in libraries and other cultural heritage institutions. Lib. Technol. Rep. **14**(5), 5–10 (2013)
28. Morón de Castro MF (2018) Art collections as an educational and research resource. Cabás: Revista del Centro de Recursos, Interpretación y Estudios en materia educa-tiva (CRIEME) de la Consejería de Edu-cación, Cultura y Deporte del Gobierno de Cantabria (España) 20 (December 2018), pp. 104–124. http://revista.muesca.es/articulos20/442-museos-colecciones. Accessed 30 Oct 2022.
29. Muñoz Cruz, V., Fernandez Cacho, S., Arenillas, J.A., et al.: Introducción a la documentación del patrimonio cultural. Consejería de Cultura de la Junta de Andalucía, Seville (2017)
30. Nykänen, P.: First steps in global advocacy: some perspectives on the formation of UMAC, an international committee of ICOM. Univ. Mus. Coll. J. **10**, 10–21 (2018)
31. Oppel, A.J.: Databases a beginner's guide. McGraw-Hill, New York (2009)
32. Pratesi, G., Franza, A.: Mineralogical, petrological and planetological heritage: the (Italian) story so far. Rendiconti Lincei: Scienze Fisiche e Naturali (2021). https://doi.org/10.1007/s12210-020-00970-2
33. Roberts, A.: Inventories and documentation. In: Boyland, P. (ed.) Running a Museum: A Practical Handbook, pp. 31–51. UNESCO; ICOM, Paris (2007)
34. Salse, M., Jornet-Benito, N., Guallar, J.: University heritage from a GLAM perspective Analysis of European university websites. General J. Inf. Doc. (2021). https://doi.org/10.5209/RGID.77215
35. Salse, M., Guallar-Delgado, J., Jornet-Benito, N., et al.: GLAM metadata in museums and university collections: a state-of-the-art (Spain and other European countries). In: Global Knowledge, Memory and Communication (ahead-of-print). https://doi.org/10.1108/GKMC-06-2022-0133. Accessed 30 Oct 2022.
36. Santana Falcon, I., Cuevas Garcia, J., Viñuales Melendez, O.: Instrumentos de conocimiento para la gestión del patrimonio cultural. In: Introducción a la Documentación del Patrimonio Cultural. Consejería de Cultura de la Junta de Andalucía, Sevilla, pp. 41–63 (2017)

37. Triplett, J.: Increasing Interoperability through the Transformation and the Consolidation of Image Collections' Metadata. In: Angel, C.M., Fuchs, C. (eds.) Organization, Representation and Description through the Digital Age, chap 15, pp. 228–240. De Gruyter-Saur, Berlin (2018)
38. UNESCO: Convention for the Safeguarding of the Intangible Cultural Heritage 2003. UNESCO. Int. J. Cult. Property **12**(4), 447–458 (2005). https://doi.org/10.1017/S0940739105050277
39. UNESCO: Heritage. In: Culture Indicators for Development, pp. 109–140. UNESCO, Paris (2014)
40. Universitat de Barcelona, Vicerectorat d'Arts, Cultura i Patrimoni: Jornada sobre patrimoni cultural de la Universitat de Barcelona: implementació del pla integral. 26 November 2019. Universitat de Barcelona. Vicerectorat d'Arts, Cultura i Patrimoni, Barcelona (2019)
41. Universitat de Barcelona, Vicerectorat d'Arts, Cultura i Patrimoni: Patrimoni cultural de la Universitat de Barcelona: recull de dades. Report, Universitat de Barcelona. Tech. rep. (2019)
42. University Museums Group.: Impact and engagement: University Museums for the 21st Century. University Museums Group, UK (2013). http://umac.icom.museum/wp-content/uploads/2017/05/UMG-UMIS_ADVOCACY-single.pdf. Accessed 30 Oct 2022.
43. Waldron, Z., Reser, G., Webster, M.N., et al. (2018) Cataloging and metadata practices survey report. Archives 45(1). https://online.vraweb.org/index.php/vrab/article/view/24 Accessed 30 Oct 2022.
44. Wijesundara, C., Sugimoto, S.: Metadata model for organizing digital archives of tangible and intangible cultural heritage and linking cultural heritage information in digital space. Libres **28**(2), 58–80 (2022). https://doi.org/10.32655/LIBRES.2018.2.2
45. Wikipedia: Template: Infobox Mineral. Wikipedia, the Free Encyclopedia (2021). https://en.wikipedia.org/wiki/Template:Infobox_mineral. Accessed 30 Oct 2022.
46. Wilkinson, M.D., Dumontier, M., Aalbersberg, I., et al.: The FAIR guiding principles for scientific data management and stewardship. Sci. Data **3**(1), 1–9 (2016). https://doi.org/10.1038/sdata.2016.18
47. Yerbury, D.: The Cinderella Collections: An Australian Fairy Story. In: Managing University Museums. Organization for Economic Cooperation and Development, Paris, pp 55–68 (2001). https://researchers.mq.edu.au/en/publications/the-cinderella-collections-an-australian-fairy-story. Accessed 30 Oct 2022.
48. Yvey, L., Emanuel, M.: Large Scale with a Small Staff and even Smaller Budget: updating metadata to reflect revised best Practices. In: Organization, Representation and Description through the Digital Age. De Gruyter Saur, chap. 16, pp. 241–254 (2018).
49. Zeng, M.L., Qin, J.: Metadata, 3rd edn. Facet Publishing, London (2022)
50. Zorich. D., Waibel, G., Erway, R.: Beyond the Silos of the LAMs Collaboration Among Libraries, Archives and Museums. OCLC Research, Dublin Ohio (2008). https://www.oclc.org/content/dam/research/publications/library/2008/2008-05.pdf. Accessed 30 Oct 2022.

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