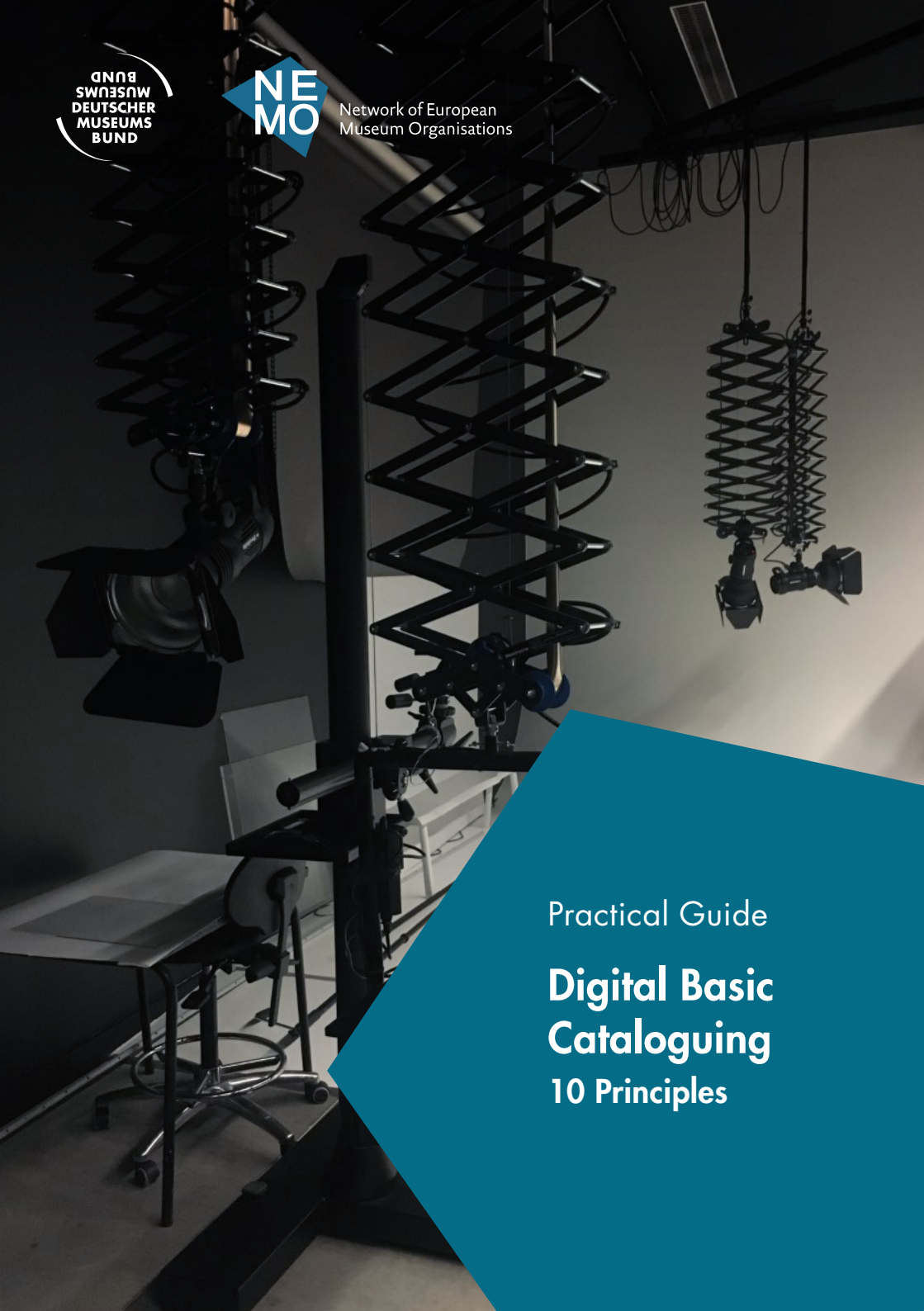


Practical Guide

Digital Basic Cataloguing

10 Principles

CONTENTS

IN USERS WE TRUST

Merete Sanderhoff, curator and senior advisor of digital museum practice at SMK – Statens Museum for Kunst, Copenhagen

What is the first thing you associate with museums? If you're a museum professional like myself, and probably most readers of this book, perhaps the first thing that comes to mind is a particular art experience that swept you off your feet. A unique historical object whose sensual presence made you feel magically connected with times long past.

But for many of the people I meet and talk to in my job, the first thing they associate with museums is the regulation of behaviour. It's the "don't touch the art" signs. The hushing of voices in galleries. The inhibition of people's natural bodily and sensory impulses when they encounter something new and strange, something that appeals to their curiosity but which they are not allowed to get near. The distance that this often creates between people and museum objects. The way it can effectively turn off the light of curiosity and interest.

What is so brilliant and promising about digitising museum collections is how it transforms unique, fragile, irreplaceable works into copies we can touch and study and dive deep into – without breaking the originals. Virtually, of course, but in the hybrid reality we inhabit that can be just as meaningful and empowering.

Digitisation removes the barriers and regulations that we need to uphold to preserve the original heritage objects. However, museums have a longstanding tradition for regulating the copies too. Historically, there were good reasons for licensing analogue reproductions such as ectachromes or plaster casts – physical copies you could lend out but needed returned in order to maintain a full record of your collection. That's not necessary anymore. If you have a digital reproduction, you can share it limitlessly without losing the original file in your storage. If the underlying work is in the public domain, restrictions are obsolete. The reproduction should be in the public domain too – as is clearly stated in the recent EU copyright directive that is being implemented across Europe's national laws.

Still, many museums fret at letting go of the old licensing mechanisms. Of course, if works are protected by copyright they cannot be used unrestricted. But even when it comes to art and heritage that is out of copyright due to age or never was in copyright, many museums maintain a tradition of regulating the natural behavioural impulses of our users – also in the digital realm – to get close to the objects that fascinate them. That's a pity because it's a lost opportunity to build connections of interest, reflection, sensemaking, and new ideas with the public – especially with those who might not think of themselves as museum lovers, but whom we can reach and offer different kinds of value and meaning with open data.

What kinds of value and meaning could that be? Let me name three examples from our work at SMK to make it more tangible. We host a community of young people who work with very diverse user groups to engage with art in alternative ways. One of their projects was to collaborate with users of a drug injection room in central Copenhagen to transform the sterile and anonymous interior into a vibrant space decorated with remixes of public domain art from SMK. Artworks were selected and remixed to reflect the users' unique life stories and hopes for the future. Through the ability to not only look at but touch and rework the digitised artworks, the users expressed themselves authentically and gained ownership of the collection. It's very likely they never set foot in SMK in their life. But with digital open access, people don't have to come to the museum. The museum can come to people, on their terms, in an environment where they feel safe and at ease.

Another community we support is citizen scientists editing Wikipedia. It's the world's biggest online encyclopedia, created by volunteers all over the world, and for millions of people it's the first place they search for information on the web. All content on Wikipedia is open access, which means it can be reused and built upon by others. This also means that open data and public domain images from museums is a godsend for the volunteers. They know that material coming from a heritage institution is fact-checked and reliable, which is a huge benefit to their practice – building the reputation of Wikipedia as a trustworthy and democratic source of information for everyone in the world. SMK's content gets more than 37 million page views a year on Wikipedia, so the benefit of being present there is beyond comparison.

Finally, we support and encourage creative and artistic reuse of our public domain artworks. Our collection bears witness to the creativity of the past, so turning it into a toolbox of digital building blocks and inspiration is a strategic move to support artists of the present and future. The collection is used to create playful children's books, stunning jewellery design, multi-award winning short films, and everything you can imagine in between.

Trust is the secret sauce. It's a conscious effort to change that first impulse to associate museums with restrictions. It requires us to make a leap of faith, but it's the best investment you can make in the future. More than ever, museums today are under pressure to stay relevant and prove that we are a benefit to society, a sound investment in times of multiple competing agendas. I believe in a bottom-up approach here. When people find us inspirational, engaging, important for their ability to develop, learn and grow, and attentive to their diverse voices and needs, our societal relevance is founded on public use value. Then we can hope to have the support of the public, now and in the future.

INTRODUCTION

"The realisation of the online intellectual society and the accompanying 'usability of information', in which knowledge can be transmitted to the point of use with the same ease as electricity, is the higher goal that should inspire our current efforts. Any museum resource information system we propose should therefore be conceived as an integral part of this larger enterprise."[1]

Over the past decade, the digital transformation within the museum sector has led to considerably expanded access to public domain heritage collections. In the various fields of work undertaken by museums – collecting and documenting, researching, preserving and communicating – digitisation has changed many things. If museums take the leap to open their collections digitally, both within but also separate from the museum itself as an established institution, it is important to develop basic criteria and procedures in order to integrate this with a holistic approach in already ongoing processes.

At first sight, publishing these recommendations on the fundamentals of digital information processing in 2022 may seem a bit late. But whereas many museums have developed their digital potential on diverse levels, others have just begun to tackle these topics, especially in the last two years, when the Covid pandemic challenged them to do so.

In order to reach the full potential of digitalisation in museums and to open collections to a broad audience for re-use, the groundwork the basic cataloguing provides the base to develop these levels of digitalisation. Even though this publication primarily addresses the elementary area of basic collecting, the developments described are relevant in many other areas too. In light of ongoing changes, however, they require constant review in terms of content and technology.

Aim of the recommendations

The goal of this publication is not the question of which information should be part of the basic collection – there are already several guidelines on this – but how this

1 Everett Ellin: Information Systems and the Humanities - A New Renaissance, in: Metropolitan Museum of Art, Computers and their Potential Applications in Museums: a conference sponsored by the Metropolitan Museum of Art supported by a grant from the IBM Corporation, April 15, 16, 17, 1968 New York 1968, p. 334.

information should be prepared. This document can be used by all museums and their responsible persons, regardless of institution size and collection type. The recommendations should make it possible to offer scalable solutions, based on personnel and the financial circumstances of an institution.

Automatic data import, export and integration characterise the nature of digital information. The modern Museum Documentation System (MDS) is therefore much more than a reference system used to research collection objects. For example, in appropriately maintained systems, loan agreements can be generated automatically from the available data, including information about the object such as transport and storage conditions, insurance requirements or legal framework conditions. The information a MDS contains is important for the work of the institution in many ways, but it can also be reusable for third parties if processed appropriately. Digital information offers the possibility of passing on this data to external digital applications (such as apps, portals and web presentations).

Usage scenarios:

- Internal processing (condition reports, auditing, loan transactions, exhibition planning and documentation, collection care, public relations)
- Data transfer to external applications for exhibition planning or restoration
- Data transfer to research, other reuse or cultural portals (e.g. the Europeana Collections, DigitaltMuseum in Norway, etc.)
- Data use for museum educational applications
- Social media
- Virtual exhibitions
- User participation

Structure of the publication

The document begins with institutional requirements dealing with the basics for selection and implementation of the recommendations.

Next there is a chapter on general principles for digital data acquisition, outlining the central aspects of acquiring digital information. Controlled vocabularies play an essential element; terminological control was already useful in times of analogue data acquisition in order to ensure standardisation and technical unambiguity of designations. The use and selection of suitable controlled vocabularies, taxonomies or thesauri is of additional relevance in the digital age, as the further processing of linked data is much more effective. One aspect of data prepared according to these recommendations is the possibility of enriching automated data with existing information, drawn from external resources. This aspect is also useful when it comes to multilingual information, which can increase the quality of an institution's own resources and their reach via the internet.

There is also a chapter on the question of provenance, which is growing in importance for collection objects. This applies not only to collection objects from colonial contexts, but also in the case of unresolved questions of provenance.

A separate chapter is devoted to digital representations, because digital images or other digital media objects are also made available as part of the basic digital collection. Directly related to the use and provision of digital information are the requirements for data exports, which are necessary for a wide variety of purposes. There are also associated legal issues, the significance of which will be outlined in this context.

Finally, the chapter on digital sustainability is intended to raise awareness of the permanence of access, discoverability and reuse, and to highlight other aspects of digital information provision.

This publication aims to provide museums and their staff with guidance on how to deal with digital data, and to support orientation points when creating a digital strategy. When selecting and adapting software systems, museums should be able to name the aspects that make sense and are adequate for their institution, in order to optimise the data quality of the digital basic recording according to their requirements.

Further information on this topic:

Spectrum. 1: Spectrum 5.0. London: Collections Trust, 2017
Collections Trust
Research Gate
Digitization Policies
Jing Culture & Crypto
Australian Museums and Galleries Association Incorporated (AMaGA)

CHAPTER 1

INSTITUTIONAL REQUIREMENTS

Over the last few decades, a great deal of information and concepts for digital collection documentation has appeared. While there is often still a need for optimisation in less complex basic recording, the existing solutions are not always adequately taken into account in the extended area of scientific documentation. A major problem is the time and technical effort that are often unavailable, regardless of the size of the institution. Adapting the procedure for basic recording to institutional capacity, in terms of personnel and technical and financial aspects, is urgently required.

MUSEUM MISSION STATEMENT AND COLLECTION POLICY

In order to achieve high-quality and appropriate digital basic recording of museum objects, certain institutional requirements are useful. For example, it is helpful to develop a collection policy for working with the collection. This policy should be based on the museum's mission statement, which defines the overriding goals of the institution, including the self-image of the museum, its location in the regional environment and the demands on its own work.

Determining the collection focus, future acquisition strategies, planned collection expansions, deaccession strategy, resource planning, as well as ideas for the use of the collection, should ideally be part of this policy. The museum's personnel, technical, legal and financial framework conditions must also be taken into account. It is crucial to focus on the aspects that can be implemented according to the framework conditions. Unrealistic goals are not helpful and can hinder the successful realisation of the set goals.

On the basis of a realistic policy, the mandatory fields for recording and documentation of collection objects are defined with their subject-specific data classified as essential. These are related to the types of objects in the collection but they should also be adapted to the intended use. Documentation guidelines should define which metadata, in which form, must be entered in the MDS. Writing instructions and additional information on how to proceed in the event of problems can be useful for individual database fields. In practice, it has proven useful to compare help texts in the MDS with the documentation guidelines.

"Ideally, the definition of the collection's focal points should be part of this policy, along with future acquisition strategies, planned collection expansions, deaccession strategy, resource planning and concepts for collection use."

DIGITAL STRATEGY

In addition to the museum mission statement and collection policy, museums should ideally have a written and communicated digital strategy. The digital strategy describes how the goals of the museum mission statement and collection policy are implemented in the digital space. Both the digital preparation of information and digital usage scenarios are defined. This publication is intended to provide helpful suggestions in this regard. Questions regarding the provision of object information on the institution's own website, the transfer of data to portals or via an interface to third parties must be answered here.

"In the ideal design, a digital strategy defines and controls all structures, measures, projects, resources, competencies and valences, but also costs and benefits, that a museum employs in the digital field and leads them into an optimal coexistence."[1]

MUSEUM DOCUMENTATION SOFTWARE (MDS)

There is a wide range of software available for museum documentation. No recommendations can be made at this point, as institutional requirements vary widely due to the range of objects and human and financial resources. Rather, general principles of basic digital recording are discussed here, which can then be implemented in the respective MDS.

RECOMMENDATIONS FOR A DOCUMENTATION POLICY:

- Define the metadata to be captured and the data fields and data groups required for this.
- Determine which data fields are mandatory.
- Determine which fields should use controlled vocabularies.
- Determine how the data is to be entered (e.g. for dates).
- If desired, define which information is to be provided in multiple languages.
- Specify the responsibilities for input, verification, release and export.
- Written guidelines should be created for the onboarding of new employees and for looking up problems.
- It makes sense to adapt the help function of the MDS according to your guidelines, so that the necessary data entry information is available directly in the system.

EDUCATION AND TRAINING OF STAFF

Knowledge about the possibilities and requirements of digital information provision and its possible uses is not available to the same extent in all collections. In addition, new technological developments, as well as cultural policy decisions and new projects, lead to changes. For these reasons, periodic further education of the responsible staff, as well as management level staff, is indispensable. This can be done through participation in conferences and workshops. Museum offices and associations have numerous offers to explore.

1 Gies, Christian: Das digital kompetente Museum – digitale Strategien für Museen, in: Lorenz Pöllmann; Clara Herrmann (Hrsg.), Der digitale Kulturbetrieb. Strategien, Handlungsfelder und Best Practices des digitalen Kulturmanagements Wiesbaden 2019, p. 102.

Further information on this topic

Alex Morrison, Digital Strategy for Museums Guide 2019. Available online at: https://sowl.co/4LsEG, checked on 02.02.2022

CHAPTER 2

BASIC RECOMMENDATIONS FOR DIGITAL INFORMATION PROCESSING

The focus of this publication is not the 'what' but the 'how'. The issue is not the selection of metadata, but the way in which information can be optimally captured in digital systems.

"Museum basements and warehouses groan and their staffs are turned into harassed bookkeepers, condemned to fall further and further behind. (...) The computer, the cool cat of the McLuhan age, seems to be the only way out."[1]

There is a difference between preparing information for use by humans and by digital systems. Not everything that can be easily 'processed' by humans can be processed by computers and vice versa. A sentence such as "the castle was built under Louis XIV" puts us in mind of Baroque architecture, since we know that castle in this context refers to architecture. Furthermore, based on background knowledge, we associate Louis XIV with the absolutist ruler of France.

The computer does not understand any of this. It only sees a sequence of letters and could, if necessary, distinguish this from another sequence of letters. In order for digital applications to be able to process such statements in a meaningful way, they have to be prepared accordingly. This conclusion is fundamental to what follows. The preparation of information in the digital environment differs in several aspects from the analogue way of working. Primary among these are:

- The coding of information elements
- The decomposition of statements into smaller, interrelated units of information

CODING

The possibility of further processing is a special quality of digital information. The encoding of information units in the form of a unique string of characters (identifiers) plays a decisive role here. These character strings are increasingly designed as URIs (Uniform Resource Identifier), which enable a direct link to the source system. If applied consistently, this can realise automated data enrichment and – depending on the resource used – even support multilingualism.

Terminology control, i.e. the application of a fixed catalogue of terms and their designations for different information areas, was already useful in analogue times and led to improved information capture, because unique object designations, location information, etc. could be assigned to the data records. With the possibility of integrating

1 Kenneth C. Lindsay: Computer Input form for art works – Problems and Possibilities, in: Metropolitan Museum of Art, Computers and their potential applications in museums: a conference sponsored by the Metropolitan Museum of Art supported by a grant from the IBM Corporation, April 15, 16, 17, 1968 New York 1968, p. 20.

external vocabularies, such as the Schlagwortnormdatei [subject authority file] (now part of the Gemeinsame Normdatei (GND) [integrated authority file], directly into the MDS via data import, greater uniformity of designations was achieved. For several years, various vocabularies have been accessible directly from the MDS via interfaces, without the need to import them into the system. In the meantime, almost all relevant systems are provided with these (mostly free of charge) interfaces.

In principle, the use of controlled vocabularies is recommended where they are already available in a suitable manner. This approach offers invaluable advantages for the exchange of digital information. In regional or supra-regional cultural portals such as the German Digital Library or Europeana, for example, information relating to external vocabularies can be integrated more easily, as searches in the portal based on vocabulary IDs take effect across institutions.

INFORMATION DECOMPOSITION

Early MDS often used singular data fields such as “artist”, “maker” or “place” to record the individuals or entities involved in the production of the object. This could lead to problems in specifying multiple entries for participants with different types of involvement; for example, in the case of engravings, the entry of the master maker and the engraver. The ability to repeat certain fields or groups of fields and to allow multiple links is an essential part of modern software. Documentation becomes more flexible, accurate and easier for digital systems to process.

REPRESENTATION OF INFORMATION ACCURACY

Information added to the MDS can be uncertain, and this may affect many aspects. For example, the attribution to an artist may not be archival, but may have been made through stylistic analysis. These uncertainties can also occur with regard to title, place of origin, provenance information, etc.

The use of a question mark (for example, “Franz Müller?” or “?Franz Müller”) is only partially suitable for solving this problem as it is unsuitable for automated further processing. If working with controlled vocabulary, retrieved via an interface, the option of adding a question mark is not practicable. Nevertheless, this common problem must be solved to avoid information loss.

One way of marking uncertain information as such is the description field. However, this solution is hardly processable for digital systems and would be ignored during searches or automated further processing. This can be remedied by an additional field in which information on uncertainty is recorded. It would be ideal to store a controlled list for this purpose (for example: uncertain, unexplained, estimated), which could then also be used as a filter in the search.

The improvement of entries for various aspects can be applied to different areas of the metadata. Also indispensable is an adapted access and rights management, by which, especially with the institution's own lists, the uncontrolled use of terms and designations can be prevented. Here the definition of the sequence of capture, data check and release must offer regulations in the workflow policy of the institution.

If possible, use existing metadata recording guidelines. These should include minimum requirements. With minimum information, the object is manageable, recognisable and distinguishable from other objects. Minimum information may include, for example:

- Classification of the object in one or more classifications
- Specification of keywords within the documentation software by the institution
- Material
- Technique
- Purpose of production and use
- Manufacturer or artist (if applicable: role, activity)
- Location (if applicable: place of manufacture, place of use, place of discovery)
- Chronological classification (dating)
- Representation/iconography
- Illustration
- Title (additional)

EXAMPLE: Classification in a systematic

A collection object is usually classified into a professionally accepted, general or in-house systematics. Subject-specific systematics can also be relevant.

If there is general technical acceptance, it makes sense to choose a systematics that is available as a web service on the internet for automated (further) processing in digital systems. This enables web-wide search for objects that have been classified in this systematics.

EXAMPLE: Creation

Persons or entities involved in the process of creating a collection item should be listed with an indication of their role (qualification) in the creation process. If there are multiple participants, they must be listed separately with assignment of their role. This allows a much more differentiated statement on the production process and the persons and entities involved.

EXAMPLE: Title

The title of an object is not always clearly assigned by the artist. Often a descriptive title is used or taken from archival sources. In addition, there may be different titles in different languages and some objects that do not have a title.

Title types (proposal):

- Title from owner
- Title by artist
- Inscription on work
- Historical title
- Descriptive title
- Obsolete title
- Erroneous title

If there is more than one title, mark one title as preferred.

EXAMPLE: Date specifications

Exact dates should be given according to ISO 8601 (or EN 28601:1992).

It is often not possible to give exact dates for the creation of an object, for example. In order to be able to give a chronological classification nonetheless, indications such as "approx.", "around", "about" and other adverbs can be used. Another possibility is to indicate style or cultural epochs ("Baroque", "Early Middle Ages", "Longobard") as a chronological classification.

However in the context of digital information processing, these statements cannot be adequately processed when, for example, searching for the dates of objects between 700 - 1000 BCE. Therefore, it is necessary to give a statement like "around 1760" an earliest and a latest date.

EXAMPLE: LOCATION DATA

Information about places, regions, and countries in the context of data entry risks being ambiguous. For example, a place name such as "Neuhaus" is by no means unique; according to the Thesaurus of Geographic Names there are 49 places world-wide in which "Neuhaus" occurs. It is therefore necessary to work with recognised vocabularies for the specification of a place in order to be able to make an unambiguous assignment here.

In addition it is necessary to note which event the place specification refers to. Here, too, a separate controlled list is useful. There could be the following aspects to choose from:

- Production
- Place of discovery
- Use
- Place of storage/current location (with links to time data if necessary)

EXAMPLE: Material

The specification of the material or materials that the collection object consists of can be simple or complex, depending on the object. Thus, differentiation by the object's parts may be necessary and requires an appropriate preliminary analysis. Precise information on the materials used is particularly helpful for proper storage and any necessary restoration measures.

EXAMPLE: Manufacturing technique

Sometimes material and manufacturing are recorded in the same field. However, in order to be as precise as possible, the information must be recorded separately. Here, too, individual parts of the object are to be recorded separately if necessary.

EXAMPLE: Keywords

Keywords are assigned to objects to improve search results. These can be sorted according to content and formal criteria.

EXAMPLE: Iconography

For objects that contain one or more pictorial representations, access to this representation is essential for the search. The pictorial components can be accessed via keywords. In many collections from art and craft contexts, however, it is also useful to name the subject of the representation. ICONCLASS, the specialised library classification for art and iconography, was developed for this purpose.

RECOMMENDATIONS FOR DATA COLLECTION

- Use controlled vocabulary for those information areas for which appropriate external vocabularies, thesauri or classifications exist.
- If no suitable controlled vocabularies are available, create your own controlled list in the MDS, which is linked to the data entry field.
- Differentiated statements are made possible by using additional fields. This can be useful when specifying information accuracy or language version.
- When necessary for accurate information collection, repeatable fields or groups of fields are useful.
- If repeatable data fields and groups exist, the entries must be qualifiable; for example, fields with obsolete and descriptive titles. Here, the type of title must also be specified in an additional field.
- If there are fields that are repeated, the preferred entry must be marked.
- It makes sense for the institution to specify in a guideline how entries in the data fields are to be made.

Further information on this topic

Murtha Baca, Cataloging Cultural Objects: A Guide to Describing Cultural Works and their Images, Chicago 2006. Online at: https://vraweb.org/resourcesx/cataloging-cultural-objects/, last checked on 29.11.2021


CONTENT="56 DAYS">
NODE

CHAPTER 3

CONTROLLED VOCABULARIES

Human language is a complex system of communication. For this reason, care in the choice of words is necessary. Technical, terminological agreements in the form of controlled vocabularies have long been common as a mode of communication that achieves a precise, unambiguous word choice. Controlled vocabularies available via the internet enable the networking of information that not only provides terms with multilingual designations, but also enables the semantic processing of information.

What is meant by a controlled vocabulary? In library, information and documentation science, a controlled vocabulary is a list of terms to be used for metadata. It specifies the preferred approach to be used in indexing. Controlled vocabularies are also referred to as authority files.

The use of external controlled vocabularies, as opposed to local thesauri or word lists, enables the creation of cross-institutional, worldwide information networks and search options. Their importance for automated processing should therefore not be underestimated. Just a few years ago, many of these systems were not freely accessible; they were subject to a fee and often only available as file packages that had to be imported into the museum's own software. This is different today. The new systems are also no longer limited one or a few controlled vocabularies. There is now an abundance of choice as to which vocabularies should be used in different institutions.

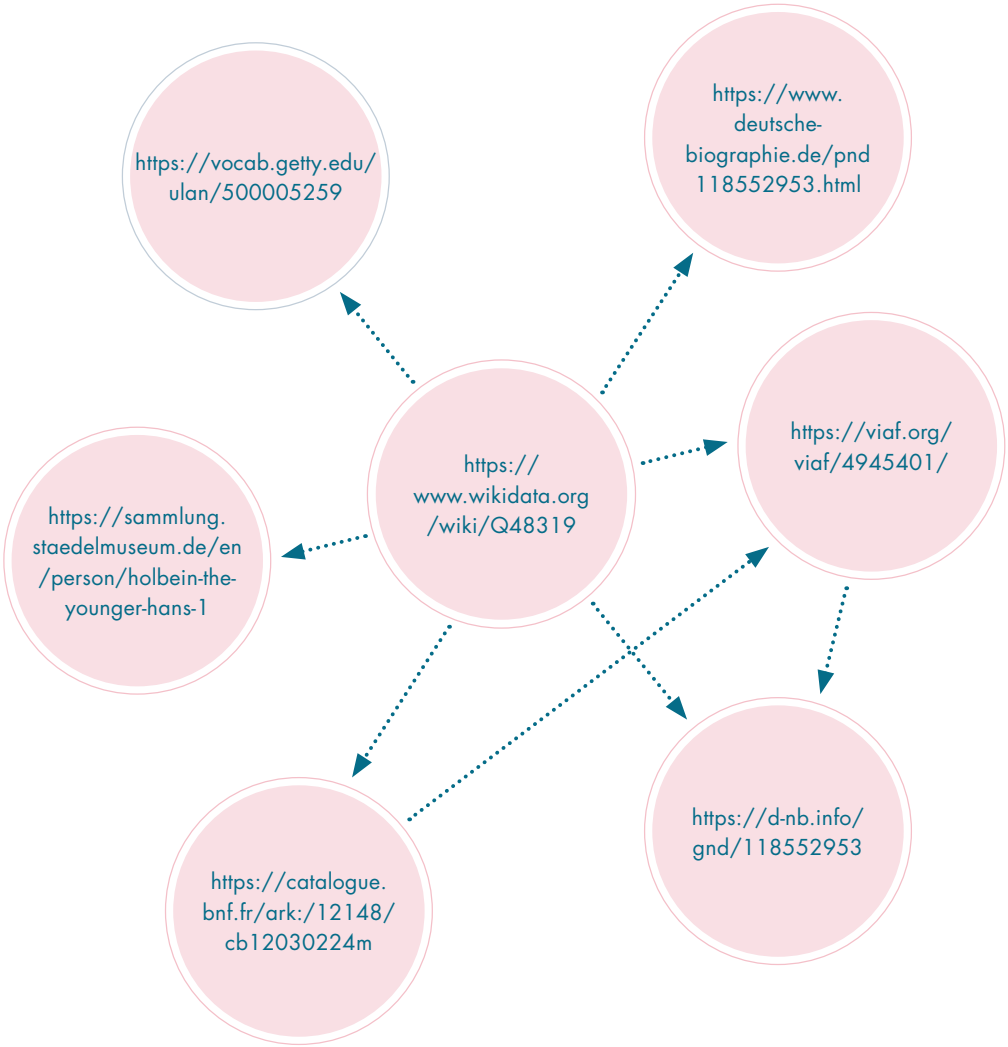
With the Linked Open Data (LOD) Cloud, a global network of freely accessible information resources has been built up for several years, the basis of which are the technologies of the Semantic Web. The advantage of this form of provision, aside from free access, is its automated use through software services. These resources also generate a knowledge network through mutual interconnection. Many controlled vocabularies are already part of the LOD Cloud (Wikidata, thesauri of the Getty Research Institute, GND), but the information resources of museums (Amsterdam Museum, British Museum Collection) are also part of the network.

Controlled Vocabulary: A collection of terms used to standardise the description of things. The terms are uniquely defined within the vocabulary.

Special varieties of controlled vocabulary:

Thesaurus: A thesaurus is an ordered compilation of terms and (their primarily natural language) denotations into a controlled vocabulary suitable for indexing in a documentation area. Equal meanings (synonyms) are defined, multiple meanings (polysemes) are resolved and term relations are documented.

Classification: A classification, typification or systematics is a hierarchically arranged collection of abstract classes (also concepts, types or categories) used for delineation and ordering.



Example: Data networking using the example of information on Hans Holbein the Younger

Modern MDS allow institutions to incorporate the vocabularies that are most appropriate for them, as well as create their own controlled vocabularies.

SELECTION CRITERIA

Many vocabularies cover specific knowledge areas. Therefore, it is not always easy to determine which information resources are most appropriate for the institution and to integrate them into the basic collection.

RECOMMENDATIONS FOR VOCABULARY SELECTION:

- Optimal coverage of the knowledge area.
  Are all the necessary terms that the institution needs available in the external resource?
- Is there free online access and a freely available API?
  Not all controlled vocabularies available online are freely accessible. Some are only viewable through access fees. Others allow free research, but there is a fee to use the API.
- Are the controlled vocabularies multilingual?
  Multilingual controlled vocabularies provide the terms in multiple languages.
- Can I be sure that this information will be available and maintained in the long term?
  For the use of controlled vocabularies for automated processing, assurance of resource provision and maintenance is elementary.
- Is the information linked to other controlled vocabularies?
  This provides the opportunity for further research and information retrieval.
- Are the vocabularies part of the LOD Cloud?
  Has the information here been created and stored using Semantic Web technologies?

Application Programming Interface (API): A programme part that is made available by a software system to other programmes for connection to the system.

EXAMPLES OF CONTROLLED VOCABULARIES

Art & Architecture Thesaurus® (AAT)

The Art & Architecture Thesaurus® (AAT) of the Getty Research Institute in Los Angeles was started in the late 1970s and has been continuously maintained and expanded ever since. It is a polyhierarchical and multilingual thesaurus for indexing art and cultural history holdings. It contains technical terms not only for object designations, but also for the physical description of objects, such as colour, materials and production techniques, abstract terms, activities and roles, and the assignment of styles and periods. There is no licensing cost associated with using the API.
Link: AAT at the Getty Research Institute

Virtual International Authority File (VIAF)

The VIAF combines several personal name files into a single service operated by the Online Computer Library Center (OCLC). By combining over 25 data sources from around the world, a very comprehensive database was created. There is no licensing cost associated with using the API.
Link: VIAF-Website at OCLC

ICONCLASS

Iconclass (also spelled: ICONCLASS) is a classification concept for capturing and indexing image content and was developed for scientific work in the fields of art history and iconography. It is operated by the Dutch Institute for Art Historical Documentation. In December 2021, the current system was replaced by a new version, which was also multilingual (English, French, German, etc.).
Link: ICONCLASS

Union List of Artist Names (ULAN)

ULAN contains names, relationships, and biographical information about makers and other individuals and entities necessary for documenting, collecting and locating information about art, architecture and other material culture. The thesaurus includes artists, architects, firms and studios, both named and anonymous. There is no licensing cost to use the API.
Link: ULAN at Getty Research Institute

Thesaurus of Geographical Names (TGN)

The Getty Research Institute's Thesaurus of Geographical Names is a structured vocabulary containing names, descriptions and other metadata for existing and historic cities, empires, archaeological sites and physical features important to the study of art and architecture. While most entries in the TGN contain coordinates, these are approximate and for reference purposes only.
There is no licensing cost associated with using the API.
Link: TGN at Getty Research Institute

Wikidata

Wikidata serves as a central repository for the structured data of Wikimedia sister projects such as Wikipedia, Wikivoyage, Wiktionary, Wikisource and others. It is easily possible to create one's own entries. Wikidata content is available under a free license, can be exported in standard formats, and can be linked to other open datasets on the Linked Data Web. There are no license fees associated with using the API.
Link: Wikidata

General Encyclopedia of Artists of the World

The database Artists of the World is the digital successor of the standard works Thieme-Becker and Vollmer. According to analogue predecessors, the AKL contains a directory of visual artists from all over the world and from all times, from antiquity to the present. Included are not only painters, sculptors and graphic artists, but also architects, designers, and many more. The use of the encyclopedia is subject to a fee.
Link: AKL

OTHER VOCABULARIES

Materials Thesaurus

Thesaurus for graphical materials.
Link: TGM

CAMEO - Conservation & Art Material Encyclopedia Online

Vocabulary of terms related to materials and techniques used in the fields of art and historic preservation.
Link: CAMEO

Unesco Thesaurus

The UNESCO Thesaurus is a controlled and structured list of terms in the fields of education, culture, natural sciences, social sciences, humanities and communication.
Link: Unesco Thesaurus

Graphic vocabulary

Controlled vocabulary of the Graphic Arts Working Group.
Link: graphics vocabulary in XTree

Creation of institutionally controlled lists

Suitable controlled vocabularies are not available for all aspects. This may be due to a lack of specialisation of existing systems for the particular tasks of an institution. However, it may also be desirable to limit the number of terms and labels that can be used in the software. As a rule, it is possible to create one's own word lists or thesauri and to integrate them into the data entry work.

CHAPTER 4

AUTOMATED DATA ENHANCEMENT AND FURTHER PROCESSING

The main advantages of digital information provision are that it automates the enrichment of object information from external sources and supports the use of the recorded data in other applications. With appropriate processing of the information as part of basic recording, a wide range of usage scenarios are possible.

Digital information processing offers a range of new possibilities for data enhancement and data reuse. As part of object recording and object documentation, information can be retrieved and displayed from external resources or even incorporated into the user's own recording system.

EXAMPLE 1:

By linking geographic information with controlled vocabularies, it is possible to source geographic coordinates from an external resource (GND, TGN, GeoNames), avoid confusion with locations that have the same name, and use them for visual presentation in maps.

Objects from the collection of the Stiftung Stadtmuseum Berlin

With this approach, it is not necessary to determine the geographic coordinates manually and enter them into the recording system.

Multilingualism is another aspect that can be simplified using similar resources.

EXAMPLE 2:

On the website of the International Computer Game Collection, only the German titles of the objects are entered into the system. A link to Wikidata allows the titles in French, Spanish, Dutch, Korean and Japanese to be automatically read out and displayed on the website when the object page is called up.

Multilingual title display on the website of the International Computer Game Collection

RECOMMENDATIONS:

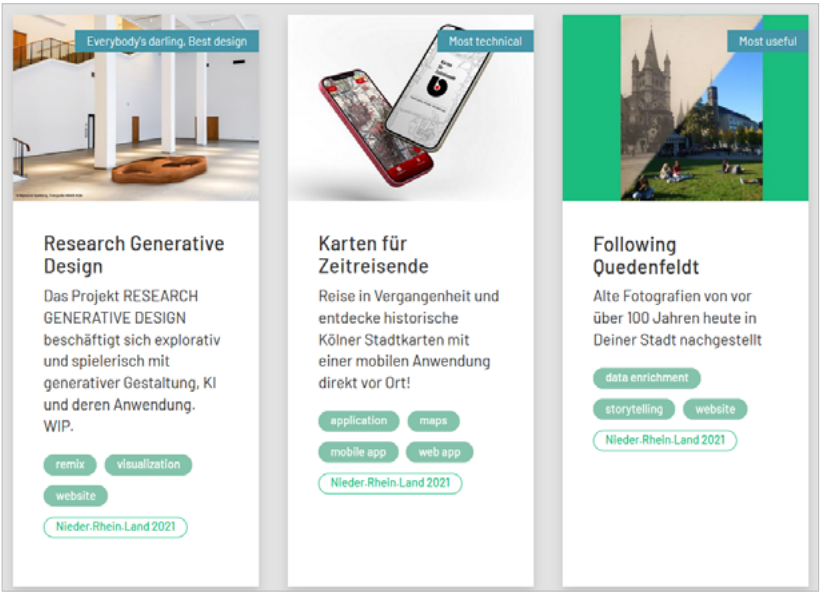
- Check which of the data fields in basic data entry are suitable for importing external data.
- Select the appropriate data source (controlled vocabulary) that contains the information these fields require but do not capture the fields themselves.
- Decide whether this information should be retrieved on demand or stored permanently in the MDS.

FURTHER PROCESSING/REUSE OF DATA

By sharing their own data, for example via the API of the German Digital Library or Europeana, museums can enable their data to be made available for third-party applications.

Wikimedia Germany, together with the German Digital Library, the Digitisation Research and Competence Centre (DigiS), and the Open Knowledge Foundation Germany, are founding partners of the Coding da Vinci Hackathon. Since 2014, hackathons have been held at various locations in Germany and other countries, where computer scientists and cultural institutions work together to explore ways to create new applications from the data available on the internet.

On the website you find many example applications (such as websites or mobile apps) for the reuse of information from different collections. As a rule, the starting point was an institution's data from the German Digital Library, which was enriched by external resources or used in new contexts.



Examples from the Coding da Vinci Hackathon

RECOMMENDATIONS:

- An institution should be aware that the information provided should conform to the principles of digital data preparation so that it can be further processed by other users.
- Subsequent use must be made possible by appropriate licensing.

Further information on this topic

Website „Coding da Vinci“

CHAPTER 5

DIGITAL MEDIA OBJECTS

"A picture is worth a thousand words." This statement is not always true. Use of words and object visualisation is optimal. For this reason, it is common practice to assign one or more images, videos or audio documents to the metadata of a collection object. Furthermore, in the digital age, born digital items are increasingly finding their way into collections. Whether original media objects or media objects generated in the course of acquisition, it is imperative that technical and legal aspects are also taken into account for their use in the context of basic acquisition.

In this context, a distinction must be made between media objects created of an analogue object for representational purposes, and born digital objects, which do not have an analogue original. Examples of this are video games and digital art.

In addition to digital photographs, digital media objects can also come in the form of audio or video files as well as 3D object data. With born digital objects, digital derivatives can likewise be created as part of the basic acquisition. For example, a multimedia application or a browser application can be represented by screenshots.

In connection with the digital basic acquisition of objects, one or more digital presentations are usually provided. This is invaluable for identifying the object. In addition, if the representations are updated or supplemented regularly, the current state of preservation can be documented.

Usually the digital representation is created by the institution itself or on its behalf. This allows the institution to choose the file format and technical requirements for its use. Existing technical guidelines, such as the Federal Agencies Digital Guidelines Initiative (FADGI) or Metamorphose, can be used to get good quality files when digitising museum objects[1]. The most extensively supported data formats possible, which are also usable in the longer term, should be chosen. Options for licenses for subsequent use must also be chosen by the institution. Licensing as open data is advantageous in terms of long-term usability.

Open Data: Open data refers to data that can be used freely by third parties but can have restrictions such as naming the author and Share-alike.[2]

File format parameters: When specifying a file format it must be taken into account that specifications are also made for the various parameters of the format (e.g. data compression, colour space, character encoding). The aim of this procedure is to create uniform format versions.

1 https://www.digitizationguidelines.gov/guidelines/digitize-technical.html
2 Share-alike requires copies or adaptations of the work to be released under the same or similar license as the original

DIGITAL PHOTOGRAPHS OR SCANS

The use of photographic images of objects is common practice in collection documentation. Nowadays, the introduction of digital photography has made this readily available for a variety of uses in different areas such as damage control, lending and press relations. Flatware such as graphics, documents, drawings, etc., can also be scanned digitally, thus contributing to the preservation of the original.

RECOMMENDATIONS FOR ILLUSTRATIONS

- As part of the basic recording, at least one digital representation should be used, ideally supplemented by detailed images or different perspectives for three-dimensional objects.
- It is important to choose the appropriate file format. Open formats such as JPEG/JPEG2000 or PNG are well suited for uncomplicated and long-term use.
- It is important to ensure that the necessary rights are secured when a digital representation is created by third parties (e.g. in the case of digitisation projects).
- It makes sense to assign licensing information that defines the options for subsequent use. Open data would be desirable (e.g. using a Creative Commons license).
- A reduced-size photo is optimal for publication on the web. The original should be saved as an unedited master file in TIFF format.
- Relevant information (name of author, object designation, inventory number) is entered in the file header (file header) using the IPTC or XMP metadata schemes. This allows the most important information to be saved in the file itself.

DOCUMENTS

If the collection item is a multi-page archival document (e.g. a deed, contract, notebook or publication), one digital photograph or scan is often not sufficient.

RECOMMENDATIONS FOR DOCUMENTS

- If the text in the document is to be made available, this can in many cases be done by automatic text recognition (Optical Character Recognition - OCR).
- It is important to select the appropriate file format. PDF/A is suitable for uncomplicated and long-term use and is also ideal for securing long-term use.
- Use of metadata schemes such as METS/MODS can be helpful for complex objects, as their use in special viewers (e.g. DFG-Viewer) offers additional options.
- Ensure that the necessary rights are secured in the event of creation by third parties (e.g. in the case of digitisation projects).
- It makes sense to assign licensing information that defines the options for subsequent use. Open data would be desirable (e.g. using a Creative Commons license).
- It is advisable to save master files (individual scans) separately in TIFF format.

AUDIO AND VIDEO FILES

Depending on the object type, representations may also be relevant in the form of audio or video files. For example, the digital representation of a record may consist not only of a photo of the record sleeve but also of an audio file with a recording of the playback result.

RECOMMENDATIONS FOR AUDIO AND VIDEO FILES

- The selection of a common file format is necessary to enable playback on as many systems as possible via the browser:
  - Audio: WAV (files may be large as data compression is missing), alternatively MP3
  - Video: MPEG-4 (with h264 codec) or webm (HTML5) with VP8 or VP9 codec.
- Licensing information is important, taking into account artist and publisher rights, ideally as open data (e.g. under a Creative Commons license).
- It is advisable to save the master files in the original file format for later use.

3D REPRESENTATIONS

There are different methods to generate a three-dimensional representation of an analogue object. 3D data can be generated with different technical systems of different levels of accuracy (for example, laser scanning or structure-from-motion). This can represent the real object as well as an augmented or reconstructed version.

Unfortunately the multitude of different proprietary file formats, which are not compatible with each other, does not simplify the choice of format with regard to long-term use.

RECOMMENDATIONS FOR 3D REPRESENTATIONS:

- The choice of a suitable file format for display in the browser is necessary. The X3D or VRML formats are currently recommended.
- Because the creation of a 3D scan is time-consuming, the following aspects should be taken into account:
  - Documentation of the scanning process (technology used, settings)
  - Backup of the raw data and colour information
  - Backup of the master file in the original file format.

DIGITAL ASSET MANAGEMENT SYSTEMS (DAMS)

DAMS are standalone applications (usually server-based) that interface with the MDS. They are designed for the management, storage and output of digital content, especially media files (graphics, video, audios, PDFs, etc.). They enable convenient management of media by simplifying import, export and enrichment with technical metadata. They are also able to control versioning and automatically provide media in suitable file formats (data migration) for different purposes.

Although software systems usually do have tools for managing media data, the scope of their services is often very limited compared to DAMS, and in part they are still tailored to the management of analogue representation (negatives, prints). There are additional costs to acquiring a DAMS, but integration into documentation software is supported by almost all providers.

When the institution owns larger quantities of digital representation, or should corresponding quantities arise in the future due to upcoming digitisation measures, it is advisable to consider acquiring a DAMS.

Further information on this topic

Ern Bieman (Hrsg.), Capture your Collections: A Guide for Managers who are Planning and Implementing Digitization Projects, Ottawa, ON 2020

https://wiki.de.dariah.eu/pages/viewpage.action?pageId=38080370, last checked on 29.11.2021

https://www.loc.gov/preservation/resources/rfs/TOC.html, last checked on 29.11.2021

CHAPTER 6

MULTILINGUALISM

The use of digital collection systems brings new possibilities in the area of multilingualism. Making information available on the web, for example via an online collection database, opens it up to a multilingual audience. The advantages of making information available in multiple languages for different purposes of use are obvious. Bearing in mind certain prerequisites, for example the use of controlled multilingual vocabulary, digital technology can be helpful support in this task.

By publishing information about objects and collections online, a primarily multilingual audience is addressed. It can therefore be advantageous to make the information available in several languages. The effort required to make information available in multiple languages depends on the institutional objective and the resources available.

Even if multilingual input is mostly of little importance for internal use, it may be beneficial for external use of information on the basic collection, i.e. in the context of an institution's collection presentation on its own website or for passing on information to national or international portals. By referring to multilingual controlled vocabularies, designations can be retrieved in other languages and thus be usable by a wider audience.

EXAMPLE: Material designation of "copper"
Link: Art and Architecture Thesaurus

The following language variants are available:

copper (metal)	(English)
Cu (copper)	(English)
koper	(Dutch)
roodkoper	(Dutch)
cuivre (metal)	(French)
Kupfer	(German)
rame	(Italian)
cuprum	(Latin)
cobre	(Portuguese)
cobre	(Spanish)
koppar	(Swedish)

Many institutions use MDS to create their own classifications or field contents with stored checklists. If available, designations in other languages can be imported automatically. If this is not provided, the entries should already be recorded in the desired language(s) in the MDS when they are created.

EXAMPLES of multilingualism through references in controlled lists:
Internal systematics Art and Architecture Thesaurus

The following language variants are available:

prunted beakers	(English, plural)
prunted beaker	(English, singular)
beakers, prunted	(English, plural)
krautstrunk	(German, singular)
noppenbekers	(Dutch, plural)
noppenbeker	(Dutch, singular)

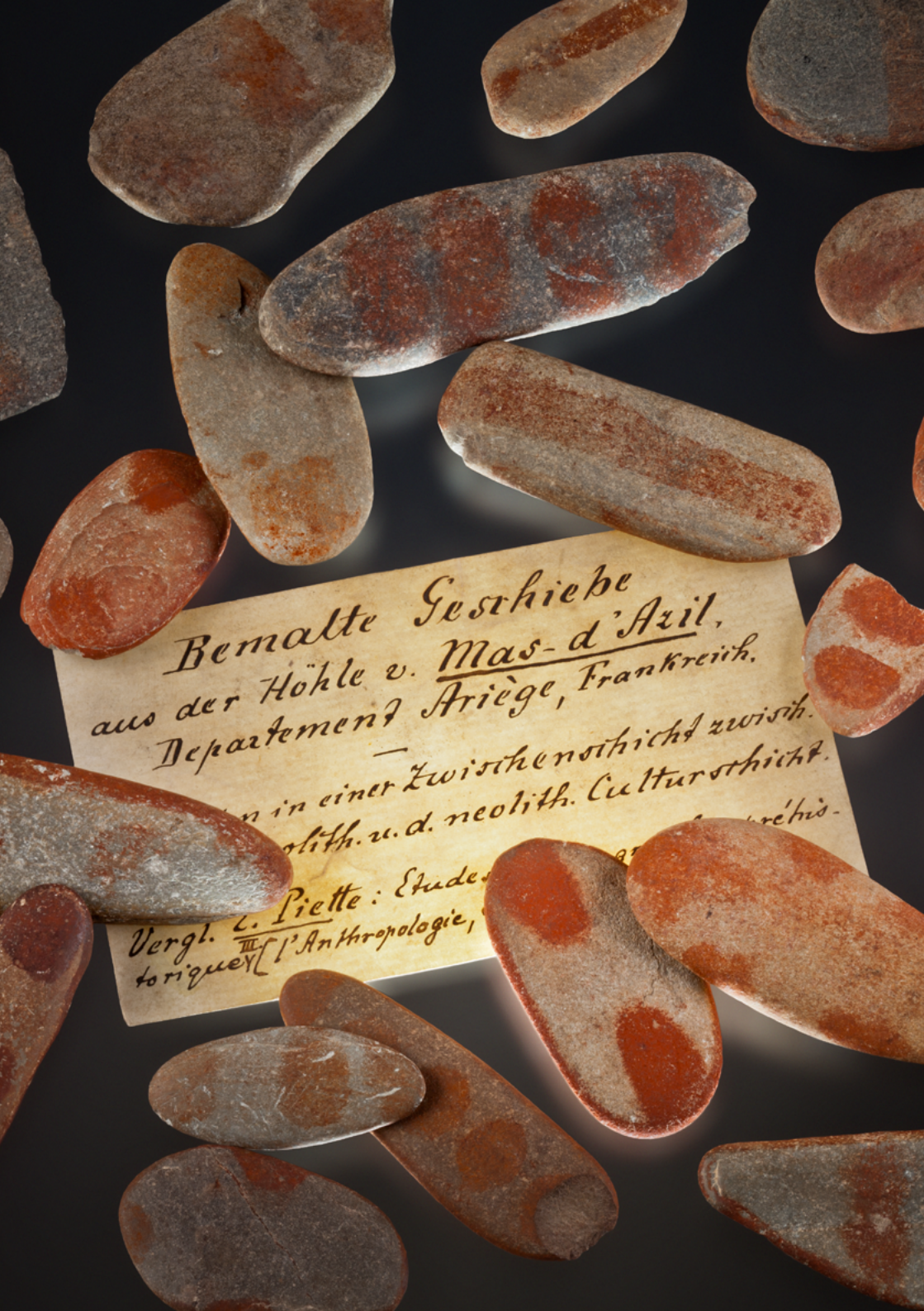
In some case, multilingual object descriptions that have to be entered into the MDS itself during data entry are more complex. As a rule, systems already offer the possibility of integrating different text versions. However, since this cannot be done automatically, manual translation is necessary as part of the data entry process and requires increased personnel effort.

If the provision of multilingual data is desired, it makes sense to integrate this directly into the data entry workflow and record it in the corresponding documentation guidelines for employees.

RECOMMENDATIONS FOR MULTILINGUALISM

- The institution should generally determine whether multilingualism should be offered.
- The institution should determine the information units that must be recorded in multiple languages.
- When using internal word lists or external vocabularies, multilingualism can be implemented more effectively than with simple text fields.
- In case of data fields that are linked to external, multilingual vocabularies, the designations can be retrieved automatically in other languages if the MDS has been prepared for this.
- Storing individual texts, for example in the object description, in multiple languages requires an increased data entry effort. Therefore it must be weighed up whether this increased effort is possible within the framework of available resources.
- Once created, foreign-language descriptions or exhibition information on the object should be stored in the software as part of the exhibition documentation and can be called up for further use.

Further information on this topic
https://pro.europeana.eu/post/europeana-dsi-4-multilingual-strategy


Bemalte Geschiebe
aus der Höhle v. Mas-d'Azil,
Departement Ariège, Frankreich.
—
in einer Zwischenschicht zwisch.
lith. u. d. neolith. Culturschicht.
Vergl. E. Piette: Etude
torique[III] l'Anthropologie,

CHAPTER 7

DATA EXPORT AND INTERFACES

A significant advantage of digital information is the possibility of its transfer and further processing. Information can be exported in a variety of ways. Which technical solution is the most suitable in each individual case depends on various prerequisites. The following section describes common processes for the respective requirements.

"Digital content is only reusable if the format in which it is stored meets the standards required by new software and delivery mechanisms. Fortunately, the electronic museum community is aware of the importance of standards..."[1]

A key feature of digital information systems is the ability to pass on data and use it in other systems. Unlike index cards or inventory books, data transfer can be automated in the digital environment. This can be necessary in various situations. Data transfer is required for hosting an object database on an institution's own website or participating in cultural portals, for example, as well as transferring the educational work of the museum to mobile apps, participating in research portals or migrating to a new MDS.

Over the last two decades, a large number of portals on museum-related topics have emerged. The dissemination of information on collection holdings has therefore become an important task of museum work. Cultural portals such as Europeana or the German Digital Library, as well as regional portals or sector portals, improve the general findability of collection items. They enable research beyond institutional boundaries and are therefore an essential building block for the presentation of our cultural heritage on the web. It is in the interest of museums to present (selected) collection objects to the broadest possible public. Optimal provision of data for these portals and other application scenarios is undoubtedly a task that must be taken into account in the basic digital collection.

There are several ways to reuse existing information from the MDS. There are different technical procedures for the transport of data volumes. When exporting by means of data packages, data from the recording system are transferred and transmitted in specially designated file formats or made available on special servers; for example, servers based on the Open Archives Initiative. When using interfaces (such as APIs), data retrieval is performed by the target system or application. Suitable preparation of the data in terms of form and content, i.e. agreements on data format and the meaning of data fields, is relevant for both approaches. On one hand, this prevents information loss and misallocations, and on the other, it minimises the financial and human resources required for data import. As a rule, generic formats are not suitable for general data provision because they involve high efforts for data interpretation and data migration.

For the appropriate form of data provision and data labelling, each institution must determine the optimal solution for itself.

1 Suzanne Keene, Digital collections. Museums and the information age Abingdon, Oxfordshire 1998, p. 17.

DATA EXPORTS

Which data about which objects should be exported?

The institution providing the data determines which data will be made available for export. The following considerations must be taken into account: existing collection depth and data quality; legal restrictions; and objectives from the museum's mission statement and collection policy.

Which data can be imported from the target system?

The requirements of the target system must be considered when preparing the export according to the intended use. For example, the choice of content for a cultural portal may be profoundly different to that for a subject portal or a research data platform. After clarifying the transmission path and narrowing down the content, it is also necessary to clarify how the requirements of the target system are designed with regard to data import. What information can be processed by the target system? Which metadata formats, standard data concepts, ontologies and standards are accepted or can be served?

What methods of data transfer are possible?

In many cases, existing data can be easily exported as a CSV file or in a programme-specific XML file (see below). With this procedure, the data provider first determines the content selection based on the aspects just mentioned. As a rule, this method requires data preparation (data mapping) from the internal data format of the MDS to a standard data format. For this purpose, the harvesting format LIDO was developed for the transfer of museum data. The use of standard formats has the advantage that definitions for the content, structure and contexts are defined in the form of an ontology. Ideally, further processing of the data can be omitted. Experience has shown that the need for individual agreements on the data format minimises if data importers and data exporters adhere to these standards.

Simple data export via CSV

Almost all data acquisition systems allow exporting data as a CSV file. As a rule, one determines the fields to be exported and receives a simple text file.

```
Abbildung.Bildnummer;Inventarnummer;Sammlung.Name;Hersteller;Titel;Datierung.von;Objektbezeichnung
3 image 13.JPG;3;Barry Humphries Collection;Humphries, Barry;"""Jag"" Costume";1998;Costume
image 4.JPG;4;;Truscott, John;Mechanical horse;1985;Puppet
100000.JPG;6;,Bach, J.S.,Poster for 'A Package Deal',2004-01-00,Work Of Art
123000.jpg;12;;Djorlam, Leonie;Basket;2010;Work Of Art
```

Example of a CSV export

The first line contains the data field names, the following lines each contain a data record. The individual fields are separated by semicolons, the data records by line breaks.

However, the performance of a CSV file is limited, for example, when exporting the links to controlled vocabularies. In addition, the character selection is limited with CSV, which not infrequently causes problems due to special and control characters. Import also sometimes requires complicated data restructuring to meet the requirements of the target system.

Data export via generic XML

Many MDS have an export function to a programme-specific (generic) XML. It is usually advantageous that this XML can be used to export all database information without loss. However, similar to the CSV export, it is often necessary to convert the data structure before passing it on to meet the requirements of the target database. In addition, the use of generic data formats requires a high level of knowledge about the indexing and input rules of the data-owning institution.

XML (Extensible Markup Language) is a markup language for representing hierarchically structured data in the format of a text file that is readable by both humans and machines. An XML document is a text file that is similar to a HTML document but uses custom tags (markers) to define objects and data within each object.

```
<?xml version="1.0" encoding="utf-8"?>
<adlibXML>
  <recordList><record priref="3" creation="2020-05-26T13:30:04" modification="2021-08-20T11:34:42">
    <media.reference>3 image 13.JPG</media.reference>
    <object_number>3</object_number>
    <creator>Humphries, Barry</creator>
    <creator>Goodwin, Bill</creator>
    <creator>Atkins, Nancy</creator>
    <title>"Jag" Costume</title>
    <dating.date.start>1998</dating.date.start>
    <object_name>Costume</object_name>
  </record><record priref="4" creation="2020-05-26T13:30:05" modification="2021-08-04T15:05:54">
    <media.reference>image 4.JPG</media.reference>
    <media.reference>image 5.JPG</media.reference>
    <object_number>4</object_number>
    <creator>Truscott, John</creator>
    <title>Mechanical horse</title>
    <dating.date.start>1985</dating.date.start>
    <object_name>Puppet</object_name>
  </record><record priref="6" creation="2020-05-26T13:30:05" modification="2021-08-18T17:44:12">
    <media.reference>100000.JPG</media.reference>
    <object_number>6</object_number>
    <creator>Bach, J.S.</creator>
    <title>Poster for 'A Package Deal'</title>
    <dating.date.start>2004-01-08</dating.date.start>
    <object_name>Work Of Art</object_name>
  </record><record priref="12" creation="2020-05-26T13:30:07" modification="2021-08-04T12:49:28">
    <media.reference>123000.jpg</media.reference>
    <object_number>12</object_number>
    <creator>Djorlam, Leonie</creator>
    <title>Basket</title>
    <dating.date.start>2010</dating.date.start>
    <object_name>Work Of Art</object_name>
  </record></recordList>
</adlibXML>
```

Example of an export in generic XML (© Axiell)

COMPLEX EXPORT BASED ON METADATA STANDARDS

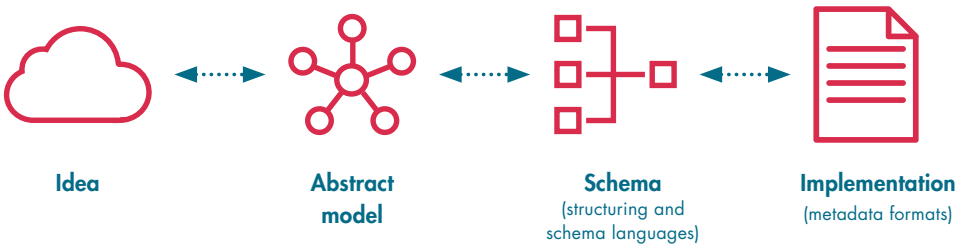
In order to minimise individual definition agreements, metadata formats that have been specially designed for this purpose are available. Their concepts have comprehensive definitions and contextual information and uniform structural specifications, and they identify what data format specification the information must be mapped in.

With the introduction of electronic documentation systems, the need for data exchange and data transfer also emerged. Within the framework of these processes, libraries first began to develop data formats, e.g. for the transfer of title and personal data records, which made it possible to use unique notations in different institutions. The primary goal here was to avoid individual or institutional discrepancies in the recording of book titles, keywords and authors.

These data models were not suitable for the use of individualised information in the context of in-depth indexing and mapping of contexts, as is often the case in the field of museums. Especially not if collection data from several institutions were to be represented in one context. As a rule, there are not simply duplicate objects in two museums, but two objects that are in context to each other due to a relation (place, time, person, etc.). The desire to pass on these contexts and make them available in general and subject-specific cultural portals made the process of data transfer very challenging for data providers, as well as for portal operators.

Initially, museumdat, a harvesting format for providing core data from museum holdings, was published as an exchange format based on CDWAlite (Categories for the Description of Works of Art Lite) and CIDOC-CRM (CIDOC Conceptual Reference Model), and was able to cover a wide character space through the use of XML. Internationally, museumdat was further developed into the LIDO (Lightweight Information Describing Objects) Harvesting Format in 2010. The version LIDO 1.1 has since been introduced. LIDO has established itself as the standard for data exchange in the museum sector.

Metadata standards: The implementation of the complete concept requires an abstract model if the knowledge domain, a structuring language and corresponding schema for this language, from which a metadata format is finally formed for implementation.



LIDO (Lightweight Information Describing Objects)

LIDO is an XML-based schema for the exchange of metadata of museum and collection objects. The CIDOC-CRM serves as the ontological basis, and care has also been taken to integrate the Spectrum standards. LIDO has established itself as a standard in many areas, for example in the transfer of data to all cultural portals and subject portals (DDB, Europeana, Digital Dürer, etc.). Since LIDO contains so-called events and not only transmits pure text information but can also be uniquely identified in many places by means of concept IDs and URIs, it is particularly suitable for mapping context information.

LIDO is divided into a descriptive part and an administrative part. In the descriptive part, information on object class, object identification, object history and relationships between objects is recorded. The administrative part contains information about rights, record identifiers and resource links. Using LIDO for data sharing has several advantages. Based on CIDOC-CRM, LIDO allows a highly flexible adaptation to the structure of the data system of the exporting MDS and enables the user to map all information relevant for the export. For the import system, this information can be transferred without loss after prior creation of a one-time import customisation.

LIDO minimum requirements: There are only three mandatory sections in LIDO, plus a LIDO record identifier and mandatory metadata language information.

- Object classifications
  - Object/work/type (mandatory)
- Object identifications
  - Title/name (mandatory)
- Administrative metadata
  - Data set (mandatory)

```
<?xml version="1.0" encoding="UTF-8"?>
<lido:lidoWrap xmlns:lido="http://www.lido-schema.org" xmlns:xsi=
"http://www.w3.org/2001/XMLSchema-instance" xmlns:gml="http://www.opengis.net/gml"
xmlns:schemaLocation="http://schemas.opengis.net/gml/3.1.1/base/feature.xsd" xmlns:xalan=
"http://xml.apache.org/xalan" xsi:schemaLocation="http://www.lido-schema.org
http://www.lido-schema.org/schema/v1.0/lido-v1.0.xsd">
  <lido:lido>
    <lido:lidoRecID lido:source="http://www.museum-digital.de/berlin" lido:type="local">
    DE-MUS-911113/1209</lido:lidoRecID>
    <lido:category/>
    <lido:descriptiveMetadata xml:lang="de">
      <lido:objectClassificationWrap>
        <lido:objectWorkTypeWrap>
          <lido:objectWorkType>
            <lido:term xml:lang="de">Computerspiel</lido:term>
          </lido:objectWorkType>
        </lido:objectWorkTypeWrap>
        <lido:classificationWrap>
          <lido:classification>
            <lido:conceptID lido:source="local" lido:type="collection_id">40</lido:conceptID>
            <lido:term xml:lang="de" lido:addedSearchTerm="no">Computerspiele</lido:term>
          </lido:classification>
        </lido:classificationWrap>
      </lido:objectClassificationWrap>
      <lido:objectIdentificationWrap>
        <lido:titleWrap>
          <lido:titleSet>
            <lido:appellationValue lido:pref="preferred">The Secret of Monkey Island
            </lido:appellationValue>
          </lido:titleSet>
```

Example 1: Extract from a LIDO document

Many MDS now have a LIDO export function, which may have to be adapted to the specific institutional modifications of the MDS.

Dublin Core (DC)

Dublin Core (DC) consists of 15 core elements for metadata description and a number of proprietary vocabularies. The goal in developing DC was to create a metadata format for describing digital resources that was as easy to use as possible. Since its first release in 1995, DC has become a widely used standard. An important use scenario was and is the use of DC to provide meta information for internet search engines.

DC can be used in a simple variant (15 elements) as well as in an extended variant. Nevertheless, when using DC, it should be kept in mind that it is a data set with a limited amount of information. DC can be notated in HTML, XHTML as well as XML. DC is still relevant for data dissemination because it is a standardised metadata format.

```
<?xml version="1.0" encoding="UTF-8"?>
<metadata
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
<dc:title>Venus und Amor als Honigdieb</dc:title>
<dc:creator>Lucas Cranach (der Ältere)</dc:creator>
<dc:subject>Venus, Armor</dc:subject>
<dc:publisher>Gemäldegalerie der Staatlichen Museen zu Berlin - Preußischer
Kulturbesitz</dc:publisher>
<dc:date>nach 1537</dc:date>
<dc:type>Image</dc:type>
<dc:identifier>Ident.Nr. 1190</dc:identifier>
```

Example 1: DC in XHTML (excerpt)

```
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head profile="http://dublincore.org/documents/2008/08/04/dc-html/">
<meta name="DC.Title" content="Venus und Amor als Honigdieb" />
<meta name="DC.Creator" content="Lucas Cranach (der Ältere)" />
<meta name="DC.Subject" content="Venus, Armor" />
<meta name="DC.Publisher" content="Gemäldegalerie der Staatlichen Museen zu
Berlin - Preußischer Kulturbesitz" />
<meta name="DC.Date" content="nach 1537" />
<meta name="DC.Type" content="Image" />
<meta name="DC.Identifier" content="Ident.Nr. 1190" />
</head>
```

Example 2: DC in XML

DATA PROVISION VIA INTERFACES

As an alternative to the procedure of exporting data by generating export packages, data can also be read directly from the institution's system by third parties.

It is important to note at this point that the data can be provided in different data formats. Interfaces never define the format of the transmitted data. As a rule, these interfaces deliver generic formats unless additional standard formats have been defined.

APPLICATION PROGRAMMING INTERFACE (API)

APIs are used to exchange information between different systems. With the increasing use of Software-as-a-Service (SaaS), where the application programmes are no longer installed on the computers or internal servers of the institution but are made available as a cloud service by the software provider, APIs are gaining in importance. More and more applications are being offered as SaaS in the museum sector, especially as MDS. Now that the data no longer resides on the institution's servers, the importance of interfaces for data exchange is growing. For example, if an institution wants to present

its collection in an online database and has a corresponding web presence designed for this purpose, this interface of the acquisition system can provide the corresponding data. Complete data sets but also single pieces of information can be provided for a specific request (e.g. "give me the IDs of all objects of the artist AB").

Authentication and authorisation measures can be used to restrict external access, i.e. only third-party systems accepted by the data provider are granted controlled access to specified areas. This access can take place via free and low-threshold registration, but it can also be subject to a charge. The provider alone decides on this.

DATA HARVESTING

In addition to the option of sending data packages to portals or other target systems using one of the aforementioned export formats, some portal providers also offer the option of data harvesting. This means that the data provider makes the information available on a suitable platform and the data user, the portal, is allowed to 'pick up' the information provided. The data provider controls access. The OAI-PMH (Open Archive Initiative - Protocol for Metadata Harvesting), among others, was developed for this purpose. This is not a special format, but a protocol with predefined steps. Different metadata formats can be used (minimum standard is Dublin Core, other formats can be offered, e.g. LIDO).

Although this method has so far been used mainly in the library sector, more widespread use in the museum sector is to be expected. The advantage of this approach is that it requires less effort on the part of the data-providing institution, since it is only necessary to determine which data should be made available for the harvesting process. The effort required for the creation of exports and the entire export management is thus eliminated.

CONCLUSION

Various approaches are available for an institution's data sharing path. The decision on which approach to take depends on a number of factors:

- For what purposes should the forwarding of data be done?
- Should the data be sent as packets or collected from the target system?
- Is it possible to use a metadata standard? Are the appropriate export options available in the capture software?
- Are financial and human resources available for the chosen path and is the effort justified?
- Which data formats can be provided? Which are required?

Only by clarifying these questions is it possible to decide on the right path for data transfer.

RECOMMENDATIONS FOR DATA SHARING:

- The creation of unique identifiers by the institution is advised.
- Determining the purposes for which data is to be forwarded is a prerequisite for efficient and adequate data handling.
- The decision as to how the data should be made available depends on the technical possibilities. Data delivery or harvesting via OAI or an API is also possible.
- The content for the target systems must be defined, i.e. which data is required for the export by the target system or desired by the data provider.
- The institution decides which data will be passed on or released and under which conditions.
- The institution must determine who is responsible for controlling the data quality and releasing the data.
- It may be necessary to adapt or supplement the MDS for the desired method of data transfer.

Further information on this topic

Dublin Core Metadata Element Set (DCMES) Version 1.1 2007. http://www.dublincore.org/documents/dces/, checked on 29.11.2021.

CHAPTER 8

PROVENANCE

Sensitivity to the nature and circumstances under which objects entered museum collections has grown considerably in recent years. The provenance of objects is therefore an important subject area that museums are committed to researching and documenting as part of the recording process. In this context, too, existing standards of digital information processing should be observed.

In recent years, the question of the provenance of museum objects has come into greater public focus for legal and/or ethical reasons, and with it the need to provide information on the acquisition of objects and, in certain cases, the history of their origin.

For all objects in museums, at the very least, the currently available information on acquisition by the institution should be recorded in the system. For a number of objects, documenting an origin history that is as complete as possible is of particular relevance. This applies to objects from colonial contexts, as well as to objects from unjustified expropriations or dubious acquisitions during for example, the Third Reich, or acquisitions from questionable sources (theft, expropriation, etc.). The same applies to cultural property that has been stolen in the course of war or other conflicts. Demands for detailed provenance information, however, cannot be limited to specific object genres or acquisition periods. It must therefore be the task of collection management to proceed sensitively and prudently in this regard. In this context, it is essential to enter new findings into the system immediately, for example by evaluating archival records.

The goal should be a complete recording of provenance. Unfortunately, this is not yet available to a sufficient extent for many objects. Either the available information is uncertain or unverified, or there are gaps in the evidence. Likewise additional scholarly research in archives, which would provide a detailed account of an object's provenance, may still be lacking. Despite these potential obstacles, the institution should address this task and provide the information available at the current time.

The museum documentation system must provide appropriate forms for recording provenance information and allow for the mapping of digital documents (e.g. literature, archival records). To enable cross-institutional searches, for example in portals for provenance research, the use of controlled vocabulary, e.g. for persons and institution names or location information, is urgently required. Furthermore the use of a uniform vocabulary for change of ownership is advised.

RECOMMENDATIONS:

Basic recording of the acquisition process by the institution should include the following:

- Data of acquisition (purchase, deposit, gift or loan)
- Where was the object located at the time of acquisition?
- Reason for acquisition
- Purchase price
- Parties involved in the acquisition process (seller, owner, auction house, person responsible for the institution)
- The digital provision of the corresponding documents would be advantageous

Additional provenance information:

- Period of ownership or time of change of ownership
- Name of the owner (with biographical data), owner-related place(s) of residence and/or abode at the time of ownership
- Type of acquisition with reference to the place of acquisition and the previous owner(s)
- Supporting documents (source references, e.g. archival record, auction catalogues, correspondence, etc.)

Further information on this topic

The Romance of Science: Essay in Honour of Trevor M. Levre, https://springer.com/chapter/10.1007/978-3-319-58436-2_14

Philipp Schorch: Sensitive Heritage: Ethnographic Museums, Provenance Research and the Potentials of Restitution, 2020, Museum & Society, https://epub.ub.uni-muenchen.de/72086/1/3459-9133-1-PB.pdf

David Pantalouy: Collectors, Displays and Replicas in Context: What We Can Learn from Provenance research in Science Museums, 2017

Arthur Tompkins: Provenance Research Today: Principles, Practices, Problems, 2021

Jane Milsoc: Collecting and Provenance: A Multidisiplinary Approach, 2019

Förster, L., Edenheiser, I., Fründt,S. and Hartmann H.(eds.): Provenienzforschung zu etnografischen Samlungen der Kolonialzeit. Positionen in der aktuellen Debatte, Berlin: Humboldt-Universität, https://edoc.hu-berlin.de/handle/18452/19768

CHAPTER 9

LEGAL ISSUES

With the use of digital media, legal issues, especially those concerning copyright, but also personal rights and data protection, are becoming increasingly important. This concerns the inventory and documentation, but especially the presentation of collections - particularly when it goes beyond traditional exhibition operations via digital media and especially the internet.

COPYRIGHT

Copyright is of great importance. It is based on the principle that creators of literary, scientific and artistic works have the right to decide how their works are used. Fundamental to this is that it must be a work, an "own intellectual creation" of the creator. This is not the case with facts, ideas, concepts, systems or methods as such; it is always necessary that a manifestation occurs. To make this concrete: a method as such is not protected by copyright, but a text about this method is. This means that the author can decide on the use of this text - but not whether the method is applied.

Works must stand out from purely everyday things (for example, a shopping list). However, the requirements for protection are not very high. For example, even short texts usually have the required level, if it is the expression of a personal, creative utterance. In contrast, this is usually not the case with formalised utility texts such as invoices.

If more than one author is involved in the creation of a work, each co-author is equally a rights holder and joint decisions must be made to authorise the use of the work.

NEIGHBOURING RIGHTS

In addition, there are the so-called neighbouring rights, which grant protection comparable to authors rights. They arise where such protection is granted by the legislature in connection with creative work. However, it is not necessary that a "personal intellectual creation" exists. For example, in favour of performing artists such as musicians or actors, in the case of producers of phonograms or films.

In some European countries, even simple, non-original photographs that are not works in the sense of personal intellectual creations are protected by neighbouring rights. However, this is not the case in all countries. And the duration of protection is shorter.

Since protection by neighbouring rights is largely equivalent to that by copyright, the most important difference in practice is the different duration of protection. Works are protected by copyright until 70 years after the death of the author. For materials protected by neighbouring rights, the term of protection takes effect as soon as the work is created or published and is 15, 20, 25, 50 or 70 years depending on the type of material and jurisdiction.

USES

Copyright generally protects the copying, distribution and the communication to the public of a work. Such uses of protected works and protected materials may only be made if the rights holder expressly consents or if there is a statutory provision permitting such use. In particular, "copying" is a protected use - hence the term "copyright". Since almost every use in the digital world involves copying data, almost every use in the digital world is also affected by copyright. Showing, publicly exhibiting, performing or reciting, filming works or editing are also activities based on protected uses. The mere enjoyment of a work (looking at a picture, listening to music), on the other hand, does not constitute a use and is not relevant under copyright law.

It is important to note that copyright protection exists in parallel with ownership rights of the physical object. For example, the purchase of a painting does not transfer copyright, i.e. the right to reproduce the painting, publish it in a book or present it online. These rights must be explicitly transferred or licensed and are not automatically transferred with the purchase of the artwork.

The author is the copyright holder by default. No registration is required for this. This person can transfer rights of use to another person, who then decides on the permissibility of further uses. The decisive factor is therefore who the rights holder is, i.e. who has the right to allow further uses.

Change in the legal situation: The law, and copyright law in particular, is subject to constant change and has been reformed several times over the last years - in some cases with far-reaching consequences for the work of museums. Fundamental is the Directive on the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society (InfoSoc Directive) adopted in 2001. It was a reaction to the changed situation due to digitisation, which, however, was not yet far advanced at that time. In 2014, an European directive was issued that allowed the online presentation of so-called "orphan works" under certain conditions. Finally, in 2019, the European Directive on Copyright in the Digital Single Market (DSM Directive) was adopted. This also contained far-reaching changes for museums. On the one hand, it became possible to make "out-of-print works" accessible online, on the other hand, it was clarified that no new property rights arise from reproductions of works in the public domain. As technology and society continue to develop, it is to be expected that there will be further reforms in the law.

LIMITATION AND EXCEPTIONS OF COPYRIGHT

For museums, there are a number of regulations according to which uses of works protected by copyright are permitted even without the consent of the rights holder.

For example, the InfoSoc Directive of 2001 provides that museums open to the public (just like archives, libraries or educational institutions), which do not pursue any economic or commercial purposes, may be permitted to make in some cases reproductions of works protected by copyright. However, there is no general legal permission to present collections on the internet. How these exceptions from copyright protection are regulated differs in the various countries.

The DSM Directive of 2019 has clarified once again in art. 8 that it is permitted in all member states of the European Union to make reproductions of protected objects that are permanently in the collections of cultural heritage institutions. This applies regardless of the format or medium, as long as it is necessary for the preservation of the materials.

ORPHAN WORKS

Especially in the case of older works, it is often not possible to identify or locate the rights holder. In order for museums to be able to present such "orphan works" online nonetheless, the Orphan Works Directive of 2012 allows the use of these works if a diligent but unsuccessful search for the author has been carried out. This diligent search must also be documented. Should the author reappear later, they can demand that the work no longer remain on the internet, and even claim appropriate remuneration for the online presentation carried out by the museum. The Directive does not apply to photographs, although in the case of historical photographs in particular, the rights holder is often unknown. Due to the narrow scope of application, the high requirements for careful searching and the risk of later claims for remuneration, this possibility of online presentation of orphan works by museums has hardly been used and is not suitable for the mass digitisation of large collections.

In the orphan works database of the European Intellectual Property Organisation (EUIPO), objects registered as "Orphan Works" can be viewed.

OUT-OF-COMMERCE WORKS

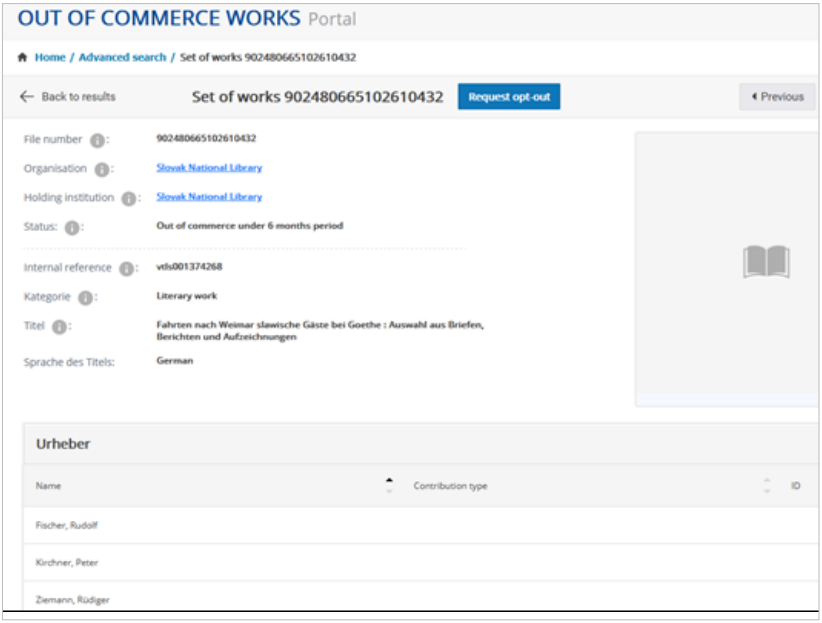
Art. 8 of the DSM Directive provides that museums, just like other cultural heritage institutions, may make available online out-of-commerce works of all kinds from their collections. Out-of-commerce works are works that are not available through the usual distribution channels. Since museums mostly have works in their collec-

tions that cannot (any longer) be "bought in a shop" or obtained through the usual distribution channels, this new regulation enables the (copyright-protected) holdings to be made available online to a large extent.

If representative collecting societies exist, licence agreements on the use of the out-of-commerce works must be licensed through them. This also applies to the works of rightholders who are not themselves organised in a collecting society. It varies from country to country which collecting societies exist for different sectors and whether these collecting societies are also representative.

However, the use of out-of-commerce-works by cultural heritage institutions is also permissible if there are no representative collecting societies. Then the use is even free of charge on the basis of a legal permission.

Required conditions for showing a work online are that it has been registered in a database of the European Union Intellectual Property Office set up for this purpose and that no rightsholder has objected for six months. Six months after registration, the museum can put the works online. Even after a museum has put the work online, the rights holder may object, in which case the work must be taken offline again. The objects already registered can also be viewed in this database:



The official database of Out-of-Commerce-works of euIPO.

EXTENDED COLLECTIVE LICENSING

The DSM Directive has also laid the foundation throughout the European Union for the establishment of systems of extended collective licensing. In the Scandinavian countries, there is also a long tradition of collective licensing in the field of cultural heritage. Extended collective licencing enables cultural heritage institutions to conclude general contracts with collecting societies for entire collections. However, use may not take place if the author concerned objects to it; they have the possibility to "opt out" for their works.

REPRODUCTIONS OF PUBLIC DOMAIN WORKS

According to Art. 14 of the DSM Directive, reproductions of public domain visual works are not protected by neighbouring rights; they are also in the public domain.

In the past, some museums only allowed the use of public domain works from their collections if they had expressly authorised it. The justification they gave was that they had rights of use to the reproductions of the works, namely the ancillary neigbouring rights of the photographer. Reproduction photography (or even the reproduction of three-dimensional opjects) did not create new works, but such reproductions were nevertheless protected by neighbouring rights.

Also in response to this behaviour by some museums, the 2019 European Union Directive on Copyright in the Digital Single Market stated that reproductions of public domain visual works were also in the public domain. The restriction on the re-use of public domain works thus no longer has any basis.

PERSONAL RIGHTS AND DATA PROTECTION IN PHOTOGRAPHS

As soon as photos of people are digitised or even are made accessible online, the General Data Protection Regulation (GDPR), which has been binding since 2018, is relevant. This is because a photograph of a person is also "personal data". It applies to personal data that their processing is only lawful under certain conditions. It is well known that use is lawful if the data subject has consented. However, there are a number of other reasons why usage may be lawful. Probably the most important case is the exercise of legitimate interests according to art. 6 (1) No. f GDPR. This requires a balancing of the interest of a museum to show a photograph of a person - for example, as a testimony of an important historical event - and the interest of the data subject not to be shown.

Many states have enacted specific laws on the use and publication of photographs, according to which the conditions are regulated in more detail. However, where this has not been done, a balancing of interests must take place according to the general provisions of the GDPR.

LICENSING BY MUSEUMS, CREATIVE COMMONS

Museums create digital surrogates and new information through photos, scans, 3D models or scientific descriptive texts that can be made available to the public through digital platforms. This can be published as part of an institutional website or collections databases.

In order for website visitors and other users to understand the extent to which they can use this information for their own research, journalistic or creative activities, teaching, private or commercial purposes, it is useful to publish it together with information about the re-use possibilities. In addition, if you pass on digitised material to an aggregator, information about the copyright status is usually part of the mandatory metadata to be provided. This applies, for example, to Europeana. Since one of the tasks of museums is to keep awareness of culture alive and to bring cultural heritage to new life, many museums use standard licences that allow everyone to use works that are still protected by copyright. Creative Commons licences have become a standard not only in Europe but internationally. These licences allow anyone to use copyrighted works under certain conditions. Eurepeana, the most important portal for cultural heritage in Europe, also provides for these licences as a standard.

With these licences, which are structured as a modular system, various conditions for the use of works can be established and combined. For example, use can be made subject to the condition that it may not be used commercially. Or that no modifications or derivatives of a work may be published. Or that derivatives of a work must be published under the same free licence as the original work. If the use of a work is already permitted by law, this takes precedence over the licences.

However, it is important to note that Creative Commons licences cannot be used by an institution that does not have the necessary rights. For example, if the copyright of collection is held by another person (e.g. the author) and the museum only has permission to publish the materials online, but not to allow re-use.

In addition to the Creative Commons licences, there is the possibility of labelling public domain content as such. The Public Domain Mark was developed for this purpose. And the Creative Commons Zero release declaration, on the other hand, ensures that copyright-protected content can be used as if it were in the public domain.

In addition, rights statements have been developed with which content that is not placed under a free licence can be marked. This makes perfect sense for copyright-protected content. The use of statements that restrict the use of public domain works, on the other hand, is questionable and also legally dubious in view of the strengthening of public domain through the DSM Directive.

The licence or rights statement to be chosen depends on the institution's attitude to re-use, including its open access policy, but also on the existing legal requirements. For example, digital photographs from third parties may already contain restrictions on use based on contracts. In terms of digital sustainability, it would be desirable to have the broadest possible reuse options.

Further information on this topic

European Union, Copyright Copyright in the EU: How to get copyright protection - Your Europe (europa.eu).

European Union, New Directive on Copyright and Related Rights in the Digital Single Market. https://op.europa.eu/en/publication-detail/-/publication/288dad28-9d3d-11e9-9d01-01aa75ed71a1/language-en/format-PDF/source-164638297

European Union, Your guide to IP in Europe: https://op.europa.eu/en/publication-detail/-/publication/ddf8fb93-ec0e-11e9-9c4e-01aa75ed71a1/language-en/format-PDF/source-164620483

CHAPTER 10

DIGITAL SUSTAINABILITY

The expenditure for digital basic recording should not be underestimated in terms of personnel, organisational and technical resources as a central task of museum work. For this reason, it makes sense to design the results of basic acquisition in such a way that they can be used for as long as possible. However, long-term archiving is only one element in the catalogue of requirements for digital sustainability.

The problems and issues of sustainable use have become a socio-politically important, if not explosive, topic in recent years in the context of climate protection. Sustainability as a requirement should enable us not to consume more raw materials than can be reproduced by nature.

This topic is also highly relevant for the museum landscape. All aspects of museum work can contribute to sustainability. Exhibitions, loan traffic, energy management in-house, museum cafés and restaurants should be examined in their work under different aspects of sustainability. In addition to this, however, the concept of digital sustainability has emerged as an objective that museums and their collections should consider.

Here, the question of permanent access, findability, and reusability of digital information is at the centre of considerations. Linked to this is the idea that resources can be conserved if information and information collections, once developed, are available in the long term. Digital resources are considered sustainable if they are permanently available and their subsequent use is made possible. This relates to legal regulations on access and further processing, as well as technical issues relating to the provision of information – in other words, a comprehensive package of possible measures.

The definition of the individual elements of digital sustainability are not uniform in the literature, but certain subject areas can be regarded as generally valid. Digital goods should be reusable and changeable. This also includes possible commercial use. To make them technically reusable, these goods should be designed in such a way that open formats and open standards are used. From a legal perspective, the use of free licenses is also necessary. To be available in the long term, they must be discoverable, accessible and permanently available.

"The concept of digital sustainability (...) states that knowledge is also important for the continued existence of humanity and should therefore be treated as a resource worth protecting. Although digital knowledge in the form of data and software could be used and reproduced at will, restrictive licenses or technical barriers hinder its potential for society." (Stürmer (2017), p. 9)

For research data, the so-called FAIR principles have become established in recent years. They consist of four principles:

- Findability (Findable)
- Accessibility
- Interoperability
- Reusability (Reusable)

Several characteristics of digital sustainability can be found here, although the focus is basically on the needs of scientific research and not on the level of society as a whole, as is the case with the principles of digital sustainability.

Further information on this topic

Stürmer, Matthias u. a. (2016): Digital Sustainability: Basic Conditions for Sustainable Digital Artifacts and Their Ecosystems. Online at: https://link.springer.com/content/pdf/10.1007/s11625-016-0412-2.pdf, last checked on 23.08.2021

Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, I. Jsbrand Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie et al. (2016): The FAIR Guiding Principles for Scientific Data Management and Stewardship. In: Sci Data 3 (1), p. 16-18

ASPECTS OF DIGITAL SUSTAINABILITY FOR DIGITAL BASIC RECORDING

The framework of this publication is the implementation of basic digital capture. Numerous aspects of digital sustainability, as listed above, remain outside this focus at present. With regard to digital sustainability, however, there are topics whose attention is worth mentioning in this context.

1. Sustainable software

Sustainability, i.e. in this case the permanent usability of software, is a component of digital sustainability. Open source software is characterised by the fact that its source code is publicly accessible and can be freely viewed, modified and used by third parties. By its very nature, proprietary software does not have this flexibility.

Nevertheless, even when using proprietary applications, measures can be introduced to ensure at least the loss-free transfer of the information already recorded when the software is changed.

RECOMMENDATIONS:

- It is necessary for the institution to have access to complete data backups, preferably in a generic XML. This should include not only the complete metadata for digital information, but also system-internal thesauri and word lists as well as digital representations.
- Documentation of the functional processes of input forms outlining individual steps should be available. Institution-specific modifications to an MDS should be documented by the manufacturer and made available to the institution.
- Export to a common file format (especially XML) should be possible at any time. An export in LIDO format is not sufficient here.

2. Permalinks

Correct citation is an essential element of scientific work. This makes it possible to locate the sources and citations used so that they can be consulted by third parties. If the source is a printed publication (book, journal, newspaper) this is usually only limited by its availability at the location. Freely accessible resources online can in principle be viewed from any location, but often have a different problem: the URL of the resource has changed and calling it up generates an error message.

Basically, there is no technical system that automatically makes a link permanent. All concepts for permanent availability require the attention and care of the providing institution. Although there is technical support in the form of higher-level reference lists (link resolvers), such as DOI (Digital Object Identifier), URN (Uniform Resource Name), PURL (Persistent Uniform Resource Locator) and others, permanent maintenance and care are also necessary for these.

Nevertheless, care should be taken to ensure that the URL is usable in the long term. What is important in all approaches is that the institution holding the digital information is aware of this problem and is committed to maintaining the permanence of the links.

Further information on this topic

Jens Klump/Robert Huber, 20 Years of Persistent Identifiers – Which Systems are Here to Stay?, in: Data Science Journal 16 (2017), last checked on 29.11.2021.

3. Versioning

The citability of sources also opens up another problem area. Resources on the internet are often characterised by the fact that they do not reach a final status. Thus, information about collection objects can be changed again and again. This may consist of correcting erroneous or scientifically outdated information, but it may also concern the expansion of documentation. The versioning of links makes it possible to transform the individual states of the resource into URLs, comparable to the different editions (and contents) of a publication. Only through versioning is citation in the sense of scientific work possible. In case of significant changes in the content of the information, new URLs should be created for citation.

To make this versioning as easy as possible for the user, the software used (MDS or Content Management System (CMS)) must be enabled to create versioned links on demand.

In Wikipedia it is possible to call a permalink for the current version. In the current version, the page on Johann Wolfgang von Goethe was accessed at the following URL:

In the menu, a separate permalink is offered in deviation from this URL:

This follows a different structure:

The page itself is supplemented by the following entry:

4. Use of metadata standards

Both content and formal aspects are relevant when selecting suitable metadata standards. The aim is to ensure findability on the basis of metadata standards, for example through standardised information on material, manufacturer or dates of origin.

In the section on -> Data exchange, metadata standards for data exchange were already presented (e.g. LIDO, Dublin Core). Dublin Core is a minimum standard for object metadata. However, there are also more complex examples such as Cataloguing Cultural Objects (CCO), Categories for the Description of Works of Art (CDWA) and CDWALite. In addition, the Spectrum standard for museum documentation is a resource that contains specifications for recording individual metadata elements and groups. The data field catalogue of the German Museums Association and the recommendations of the German Museums Association on documentation can also be used as standards.

Further information on this topic

Forside - KulturIT

5. Licensing for reusability

Providers of web resources should be aware of the fact that there is a question of reuse of the digital information they make available. For this reason, it makes sense to inform website visitors about the ways in which these resources can be reused.
See -> Legal issues

Further information on this topic

Kreutzer, Till (2011): Open Content Lizenzen. Ein Leitfaden für die Praxis. [Elektronische Ressource]. Bonn: Dt. UNESCO-Komm (Bildung, Wissenschaft, Kultur, Kommunikation). Online at: https://irights.info/wp-content/uploads/userfiles/DUK_opencontent_FINAL.pdf, last checked on 29.11.2021.

6. Semantic processing

Developed by T. Berners-Lee et al. the concept of the Semantic Web is a reaction to the still prevailing 'atomisation' of information on the World Wide Web. Semantic processing of information is absolutely necessary in the context of digital sustainability in order to relate information of all kinds and thus enrich it semantically. Technically, this information is based on the use of certain standards such as Resource Description Framework (RDF), Uniform Resource Identifier (URI), ontologies (OWL and RDFS) and the query language SPARQL. This can automate the exchange and usability of data between machines.

Further information on this topic

G. Antoniou/Frank van Harmelen, A Semantic Web Primer (Cooperative Information Systems), The MIT Press, Cambridge, Mass., 2008

7. Digital long-term archiving

The main difference between digital sustainability and digital preservation is that the former is primarily concerned with the production conditions of digital objects, while the latter focuses on long-term preservation and securing the use of existing digital objects.

However, with a view to the Open Archive Information System (OAIS), which is now established as a standard, numerous overlaps become apparent. The requirements contained there go far beyond the aspect of pure data protection (backups). Procedures are also implemented in OAIS to ensure the long-term use of digital information.

PICTURE CREDITS

Cover: © Sylvia Willkomm
p. 10: © Dorothee Haffner, cc-by-sa 4.0
p. 22: © Pixels Hunter | stock.adobe.com
p. 42: © Landesamt für Archäologie Sachsen, Foto: Rengert Elburg
p. 61: © HTWK Leipzig, Caroline Schmunck

IMPRINT

Practical Guide
Digital Basic Cataloguing
10 Principles

1st version 2022
ISBN: 978-3-9822232-7-8

Publisher: Network of European Museum Organisations
and German Museums Association

Author: Winfried Bergmeyer, Paul Klimpel (Legal issues),
Member of NEMO Working Group Digitalisation
and Intellectual Property Rights:
Anna Benedek, Ismo Malinen, Mareen Maaß,
Ariadna Matas, Liv Ramskjaer, Sylvia Willkomm.

Proofreading: Geraldine Kendall Adams

Design: MATTHIES WEBER & SCHNEGG, Berlin

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Supported by:

KULTUR STIFTUNG DER LÄNDER

Co-funded by the European Union

Network of European
Museum Organisations

NEMO - The Network of European Museum Organisations
Karl-Marx-Platz 12
12043 Berlin
Germany
www.ne-mo.org