University of Glasgow

National Museums Scotland

Final Report

Surfacing the National Collections: adapting image cataloguing standards to transform access to National Museums Scotland's online collections

Cassandra Kist & Maria Economou
University of Glasgow

Chanté St Clair Inglis, Pam Babes, Angus Kneale & Hannah Norfolk
National Museums Scotland

Funded by the University of Glasgow Knowledge Exchange Fund

July 2023

Have feedback? Share it with us here:
https://forms.gle/UAfRnaGKZ9vmFuFdA

Acknowledgements

We would like to thank the following for sharing their time, expertise, and advice: Gordon Meiklejohn, Fraser Rowan, Helen Green, Lucinda Thompson, and Neil McIntyre from the University of Glasgow and Andrea Cop from National Museums Scotland for their guidance on putting together the project, the funding application process and their support throughout the project; Malcolm Chapman from The Hunterian for thoughtfully sharing his knowledge of collections management systems, (meta)data, and search portals; Our colleague from the University of Hamburg, Research Fellow Quoc-Tan Tran, for being a constant sounding board and pointing us to relevant literature; from the Information Studies PhD researchers, Aimee Fletcher for sharing her experiences in qualitative survey formatting and methods, and Joy Nam for her advice on cognitive mapping techniques; the Information Studies academic staff for their feedback and support; and all the National Museums Scotland staff as well as external workshop, survey, and interview participants who generously gave their time and offered us valuable insights.

Table of Contents

List of figures

1 Project summary, key findings, and recommendations

"Surfacing the National Collections: adapting image cataloguing standards to transform access to National Museums Scotland's online collections" (hence Surfacing the National Collections) was a four-month collaborative knowledge exchange project between the University of Glasgow and the National Museums Scotland (the Museums). This project had the goal of increasing the findability of the Museums' online collections and understanding current and potential user needs. It aimed to develop a framework for applying keywords and captions to images of the collections that fit within staff needs and metadata best practices. It was sparked by a previous 2022 research project[1] which revealed a low access rate to the Museums' online collections; that most visitors to the online collections were new and not returning; and finally, that there was a gap between the terms that the public uses to search the collections and the terms that staff apply to the collections' metadata (Chowdhury et al., 2022).

The Surfacing the National Collections project employed five main methods:

- interviews with the Museums staff;
- a content analysis of user search queries collected over 2020-2021 from the institutions' collections webpage – the Search our Collections (SoC) portal;
- public surveys;
- user interviews involving concept-mapping;
- and a card sort activity.

1.1 Key Findings

This report discusses the findings listed below in greater depth and situates them in the Museums' context and current literature, followed by the recommendations made:

1. General terms pertaining mostly to types of objects, people, and things are commonly used to search and browse the Museums' collections.
2. There is little difference in the keywords/terms suggested for collections' images by survey participants who have 'expert' knowledge of cultural heritage and those with 'general' knowledge.
3. Visual terms such as colour and shape are rarely used by the public in queries of the Search our Collection portal. However, such terms could open alternative entry points to the collections if these different methods of searching are made evident/clear.
4. Four main user groups of the online collections were identified from our surveys: museum/cultural heritage professionals; curious individuals; cultural heritage enthusiasts; and students/academics. These groups are visiting the online collections for work, school, or general interest in a topic and use images primarily for discussing

[1] See the report of this AHRC Towards-a-National-Collection-funded (TaNC) project: Chowdhury, G., Gibson, R., Chowdhury, S., Ross, J., St Clair Inglis, C., Cawston, R., & Ganley, C. (2022). Digital footprints and search pathways: Working with National Collections in Scotland during Covid-19 lockdown to design future. Zenodo. https://doi.org/10.5281/zenodo.6602365

cultural heritage with others, learning, and creative activities. The survey reached users who are mainly Scottish, based in Scotland, and female, aged 25-34 & 35-44.

5. Most users visit the collections to access more than one type of information and are interested in following links across different forms and layers of information/media, and some groups (cultural heritage enthusiast) were interested in sharing/re-using images.
6. Most users in the survey sample undertook more than one search/browsing activity during their visit.
7. Most survey participants are primarily interested in searching/browsing the collections by the theme of underrepresented histories or by the figures and things represented in images.
8. Many of the Museums' staff who were interviewed (particularly those who engage with or respond to user queries) prefer to use the public-facing Search our Collections portal and have encountered successes and challenges that speak to user experiences.
9. Exploring underrepresented histories in user interviews showed that this theme is first associated by participants with different aspects of identity having to do with ethnicity, gender, sexuality, and socioeconomic class. It is then associated with individual/community narratives and experiences such as migration, homelessness, mental health, suffrage, trade groups, and protest.
10. When sorting images of the collections, interview participants were interested in the narratives produced when grouping together different images and how the collections could be organised based on similar functions (e.g., for work, to be decorative or symbolic), and then by date/provenance.

1.2 Recommendations drawn from the findings

1. As shared later in this report, the application of keywords and captions to images should prioritise general terms for type of objects, figures, or things visible in or associated with the collections' images.
2. Secondary priorities may include applying terms connected to individual/community experiences associated with underrepresented histories (see further details in section 7). Additionally, applying visual terms having to do with colour or shape to images of the collections could enable different entry points into the collections – particularly, for creatives and staff who work in design/media.
3. Users desire a more 'helpful' search interface that offers suggestions or drop-downs which can evidence different ways they can search or explore the collections.
4. Incorporate sustainable channels or methods for gathering Museums staff feedback on the Search our Collections portal, and for if/when keywording is implemented, to enable staff to cross-check keywords and problem solve.
5. Ensure high quality images are accessible and that visitors are aware of how they can be used, and that images are linked with other material and media both internally and externally.

6. Involve user groups to develop virtual trails, maps, or canned searches related to underrepresented histories which can improve user experiences of online collections while diversifying knowledge production and sharing.

2 Project Introduction

Cultural heritage institutions, such as National Museums Scotland (the Museums), are mandated to make their collections widely accessible to multiple publics (Museums Association, 2023). Museums are increasingly seeking to facilitate this access through the wide global-reaching capabilities of the web and social media platforms (One Further, 2022). Fulfilling this mandate is met with several challenges. While digitised images, 3D models, and even virtual renderings of collections can offer new experiences to publics online, the proliferation of online collections has also been met with some critiques (Bailey-Ross, 2021) and internally, a lack of staff confidence in the quality of object webpages (One Further, 2022). These critiques and uncertainties can stem from ambiguity regarding who the collections pages are for (Bailey-Ross, 2021; Newbury & Brennan, 2021), their value for different publics (Hopes, 2014), tension between quality/depth and quantity of online collections (Martin, 2021), and how to move online collections pages beyond their roots in catalogue cards to engage users (Newbury & Brennan, 2021).

Moreover, common critiques are also rooted in the observation that the metadata applied to collections/collection images may serve institutional staff practices but not the search methods and interests of users. This creates what has often been referred to in museum studies as the 'semantic gap' – a friction between staff and user search terms (Klavans et al., 2014). Coupled with the intense staff labour/time required for digitisation and applying metadata to images, these factors can lead to ambiguity regarding the value of creating and maintaining large quantities of online collections. However, Covid-19 sparked and brought to the fore diverse uses and forms of value that digitised collections can have for online audiences – including crafting, socialising with others, and for emotional respite, activities which since the pandemic have continued to increase in popularity (The Audience Agency, 2022). To support these practices, digitised collections must be accessible, findable, retrievable, and optimally, reusable. Therefore, the metadata applied to images must be ethical and robust; intended for specific audiences, meeting their vocabulary needs and interests and sensitive to their memory making processes and emotions (Giaccardi, 2006; Martin, 2021); and be standardised so as to help make the images retrievable (Slawsky, 2007).

As such, this project aimed to increase the findability and retrievability of the Museums' collection images through its *Search our Collections* (SoC) portal by addressing the semantic gap. This gap was observed in the Museums' metadata by Chowdhury et al. (2022) in a recent (2020-21) AHRC COVID-19 TaNC project exploring users' search pathways which led to the current project. The change in recording metadata proposed by this project will serve as a foundational step towards enabling the Museums' collection images to be further findable through other search engines such as Wikipedia and Google and eventually, be re-usable in wider digital ecological contexts. The following section provides institutional context for the Surfacing the National Collections project, followed by a literature review,

and a brief description of each research method, ending with the last section presenting the associated key findings and recommendations made deriving from these.

3 National Museums Scotland and project context

3.1 Overview

National Museums Scotland (the Museums) cares for over **12.4 million** objects from diverse subject areas. These are widely shared through research, display, loans, touring programmes; online via the museum's *Search our Collections* (SoC) pages (see figures 3-1 and 3-2); and third-party platforms like Sketchfab. However, less than **2%** of the collections shared via the museum's SoC pages are accessed by anyone. We know audiences are engaging through Wikimedia and social media, but wanted to assess whether and how they are engaging with SoC.

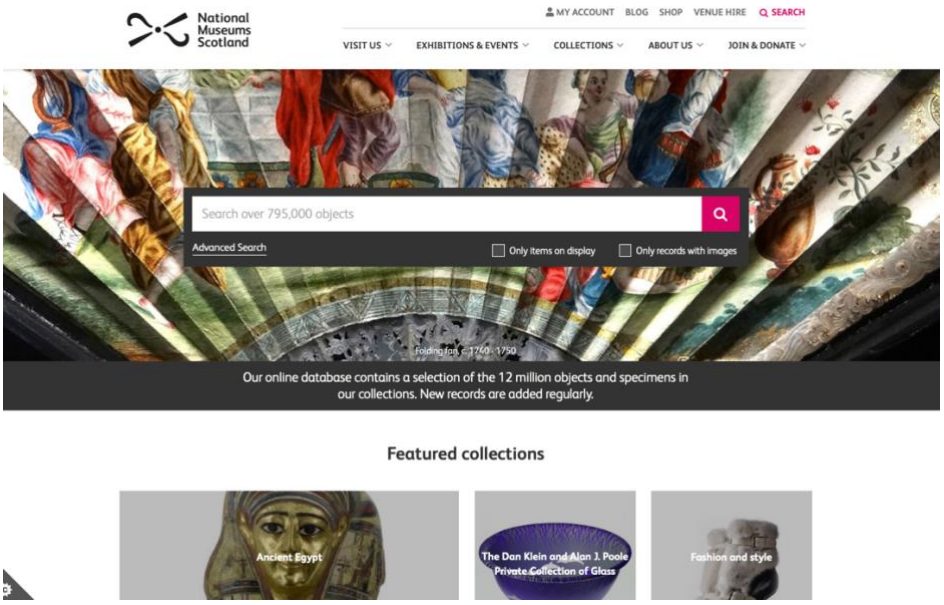


Figure 3-1 Screenshot of the SoC search portal

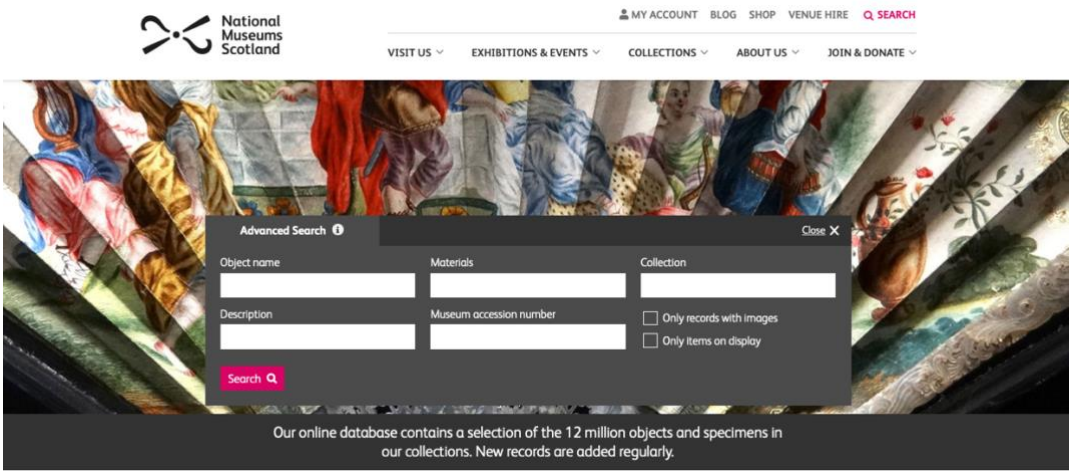


Figure 3-2 Screenshot of the SoC advanced search portal

This 4-month Surfacing the National Collections project was developed to inform image cataloguing standards through an audience-centred approach. The project team aimed to tailor metadata standards to the Museums, based on industry best practices, the motivations and needs of target audiences, and the Museums' internal management needs.

3.2 Intended goals

The resulting innovation of the project, the proposed frameworks for applying keywords and captions to images of the collections, is intended to increase the collections' findability and promote audience-driven collection access. Further, the project intended to help staff integrate in their work the motivations and needs of audiences accessing the collections. This is a foundational well-timed step in the Museums' new strategy (accelerated by Covid-19) to open-up the online collections, diversifying access.

The **primary objectives** of the project encompassed:

1. Increasing the findability and usability of NMS collections through improving image metadata standards, and;
2. Increasing staff's understanding of audience motivations and needs in accessing and using images of the online collections.

The **target outputs** included:

1. A framework for applying key words and captions to the Museums' online collections images (included in section 6 of this report).
2. The application and evaluation of the model to a small sample of images by different members of staff (discussed in section 6).
3. A report on current and potential audience motivations and needs in accessing the collections (discussed in section 5.2 & 5.3).

This project coincided with ongoing multiple changes in the institution including the adoption of a new Digital Asset Management System (DAMS), Piction, and the development of a new web interface. Further, the institution has been working hard to apply a decolonial lens to the collections, critically assessing the type of language and terms used across the metadata in conjunction with curatorial staff. In addition, the Museums' current strategic plan has a commitment to delivering wider and richer access to online collections. This includes a change in open access licensing with plans to share collection assets under more open terms. The Museums also recently (May 2023) started a collaboration with the University of Glasgow and other cultural heritage institutions in a new Innovate UK project exploring Museums in the Metaverse.[2]

3.3 Intended outcomes

Outputs 1 and 2 are intended to enhance the findability of the collections for target audiences and increase their use across networked digital platforms (Wikipedia/social media). Output 3 is intended to support the Museums' transformational next steps in implementing an institutional strategy aimed at increasing engagement with the online collections: it will bolster staff confidence with an audience-centred approach by grounding

[2] https://www.gla.ac.uk/research/az/cspe/news-and-events/headline_933402_en.html

their decision-making in an understanding of audience motivations and needs. The Museums' current strategy includes scaling up collection digitisation programmes, sharing collections via third-party platforms, and moving to open-access licensing models. Therefore, this project will act as a foundational step in a longer-term process of extending the reach of the Museums' collections. An additional step post-project will include evaluating user engagement with the online collections pre- and post- the application of standards. Similarly, an understanding of audience motivations and needs by staff will be measured through staff feedback.

Important broader outcomes of the project include enhancing public education, diversifying perspectives on the collection, and supporting creative partnerships with the institution. Enabling the findability and usability of the Museums' collections will support the diffusion of expert knowledge on a range of subjects and support public education online. Further, the project serves as an important step to enable the use and reuse of the digitised collections, thus allowing audiences to contribute new knowledge and diversifying perspectives on the collections. In turn, as evidenced by the Smithsonian Institution Open Access initiative,[3] an accessible and open access collection can empower users to innovate and create solutions to world problems. (Since opening their collections in 2020, the Smithsonian have had 47.5 million views and 2 million downloads). The findings and frameworks developed can also inform other cultural heritage institutions in the UK for aligning image metadata standards with the needs of online audiences.

To fulfil the goals and outputs, and contribute to the intended outcomes the project had three main stages:

Stage 1) Understanding the Museums and the collections system (interviews with staff)

Stage 2) Understanding who and how audiences' search the collections (content analysis of Search our Collections queries, public surveys, and interviews with users)

Stage 3) Creating and testing frameworks for applying keywords and image captions (discussions and workshop with staff)

4 Literature review on increasing image findability

4.1 Refractive, dynamic, and networked cultural objects

Practitioners and professionals have frequently debated and discussed what an optimal search interface should look like in order to meet the interests and needs of online audiences and support meaningful experiences (Brooke, 2022; Newbury & Brennan, 2021). Interfaces have been discussed from user-centred perspectives that consider user motivations for visiting museums online like those of visiting the physical museum, such as '(1) social-related reasons, (2) recreational, sight-seeing reasons, (3) learning and personal enrichment-related reasons, (4) hobby and professional interest-related reasons, and (5) reverential reasons' (Falk & Dierking, 2016, p. 44). What can sometimes get lost in these

[3] See si.edu/openaccess.

pivotal conversations, however, is how to make the underpinning metadata upon which such interfaces rely, 'user-centred'.

Today, there is a movement away from museum interfaces that enable just the search and retrieval of specific objects, to ones that support more exploratory experiences that encapsulate how visitors can wander, interact, and even create in physical exhibition spaces. Some scholars/practitioners such as Giaccardi (2006), call for collection management systems to be created in ways that are more sensitive to how people stimulate and regulate emotions and memories (p. 39), moving collection management away from its stubborn roots in the traditional catalogue card-file system (Newbury & Brennan, 2021). Others, similarly, call for 'generous' interfaces that enable audiences to browse, lose themselves, and follow their noses (Hall & Walsh, 2021). These interfaces might take on creative map or tree-like qualities, have suggested searches – by colour, by the ever-loved museum cats, certain narratives, and/or rely on more interpretive materials such as stories, games, and videos (Hall, 2018).

To support such interfaces and exploratory experiences, we must consider cultural objects online as wholes composed of multiple pieces, and as not static but constantly changing/dynamic in different online contexts. As describe by Navarrete et al. (2016): 'all individual components—the object, collection, museum, or metadata—are placed side by side at the user's disposal in the information space' (p. 117). Importantly, together, these components, and the perspectives of users in online contexts, change how cultural objects are experienced, how forms of meaning can be made, and opportunities for new social and emotional connections/interaction created (Giaccardi, 2006). Giaccardi (2006), building on Benedetti (2002), describes the interplay between components of cultural objects (such as museum, object, images, metadata, stories) and the resulting different user perceptions and experiences of cultural objects as a form of 'iridescence'.

The question then becomes for the cultural heritage field, how can we sustain this 'iridescence' of digital cultural objects? And how can we be sensitive to and support different publics' emotions and memories when engaging with online collections systems? And for the purpose of this project, how can this sensitivity and dynamism be pursued through the metadata of collection images, as one part of the puzzle of complex cultural objects? From this perspective, the metadata of collections images can add another illumination, another refraction, that complements and works with other components of cultural objects online. One pivotal part to answering these questions, which has been long advocated for and now underpins metadata ethics, is understanding the search interests and needs of museums' different publics (Martin, 2021).

4.2 Understanding audiences

When it comes to digital content creation in museums, often, little consideration is given to detailed audience segmentations/profiles or personas that are defined and catered to in other areas of the museum service such as exhibitions and learning programmes (One Further, 2022). However, it has been repeatedly recognised in the cultural heritage field that metadata and its ethics, is tied to understanding user interest and needs (Dobreva et al., 2011). Previous research has shown a growing interest in understanding users' interests and

needs in searching and exploring cultural heritage collections (Choi & Rasmussen, 2003; Clough et al., 2017; Hopes, 2014; Villaespesa, 2019; Walsh et al., 2020). As a result of this research, it is understood in the cultural heritage field today, that the findability of collections and its interconnection with the search interface is, similarly to visitor experiences in the physical museum, shaped by aspects of user profiles. These user profiles have been categorized and explored by practitioners/academics in several ways, as discussed in the following paragraphs.

Some cultural heritage professionals and researchers, such as Hopes (2014), suggest that shared professional identities, practices and knowledge can influence search motivations, needs, and experiences. For instance, user research in the context of search portals has focused on certain groups such as: digital humanities professionals (Warwick et al., 2008), and students/researchers of American History (Choi & Rasmussen, 2003). How user groups search collections may also be impacted by the domain that users search within (collections categories) e.g., fine art vs natural history (Armitage & Enser, 1997; Chowdhury & Chowdhury, 2011). Comparatively, literature that focuses on users *broadly* suggests that users approach online collections in different ways due to motivations for a particular visit and users' previous experience with technology and cultural heritage (Hollink et al., 2004; Villaespesa, 2019). Diving further into user 'motivations', the main reasons for why broad user groups search/browse the collections have been identified for the Metropolitan Museum of Art (MET) collections as: 'professional researcher, personal interest information-seeker, student researcher, inspiration seeker, casual browser, and visit planner' (Villaespesa, 2019, p. 247).

In turn, researchers have defined the main modes of searching as being either a 'general search' or a 'specific search' which can further impact user experiences and interactions with search portals (Clough et al., 2017) These main modes of searching have been further divided and applied by Skov and Ingwersen (2014) into 4 categories to evaluate information retrieval systems: researching a well-defined topical information need; researching topics via data elements; researching an ill-defined topical information need; and researching a known item via data elements (p. 94). With this idea of interaction and modes of searching in mind, researchers have further segmented users and/or visitation profiles by using web-analytics (Chowdhury et al., 2022; Hale, 2019; Han & Wolfram, 2016; Warwick et al., 2008). This enables patterns to be identified in search queries and search behaviours subsequently categorised to inform tweaks or changes to interfaces and metadata.

Comparatively, Bailey-Ross (2021), undertaking research and an in-depth literature review to inform a new UK-wide collections search engine, identifies and summarises four main ways of categorizing and defining user groups of cultural heritage collections online:

> *'1. motivation (which can change during visits and between visits); 2. level of expertise/role (familiarity with tech/cultural heritage); 3. mode of interaction (browsing, specific search etc.); 4. and the use of web analytics'* (Bailey-Ross, 2021, p. 1).

Other but less prominent ways of understanding users of online collections and their experiences, include: specific use of resources (e.g., for sharing, for learning, for teaching, for art – which overlaps with motivation); access requirements which can connect to expertise but extend to access needs (e.g., alt-text); domains being searched (art vs natural history) (Hollink, 2004); and finally, demographics such as age, gender, and location.

The user research that is part of this study, as will be discussed in section 5.3, attempted to understand users and their experiences using Bailey-Ross' (2021) identified categories, with the addition of demographic questions. As Bailey-Ross points out, echoing Mihelj et al's (2019) research, divisions in access to cultural heritage having to do with socio-economic factors of race, age, and gender tend to be reiterated online. Therefore, considering demographics could help us understand a range of user experiences, interests, and needs beyond what might be considered as core user groups. This was undertaken, keeping in mind Chowdhury et al.'s (2022) suggestion, that 'user and context specific guidelines could be useful in ensuring the aspects considered most important by consumers are indexed, thereby producing more relevant search results' (p. 46). Understanding user needs and interests and how these relate to metadata, was recently added to Spectrum guidelines, indicating its centrality to museum practice: '*How will you enable access to collections data for different users and stakeholders?'* (Collections Trust, 2022, n.p.)

4.3 Categorising and describing collection images for users

The characteristics used to group and describe users, including motivation, level of expertise/role (familiarity with tech/cultural heritage), and mode of interaction (browsing, specific search etc.) are recognised as influencing the terms users use to search or browse. As pointed out by Villaespesia (2019) and Chowdhury et al. (2022), personal motivations for visiting and users' previous experiences with cultural heritage can shape how users search and consequently, users' resulting experiences. In turn, making images accessible to different groups by capturing and describing the appropriate details/level of detail represented in an image and through appropriate terms can be extremely challenging. This has been a central issue in applying metadata to 2D images across cultural heritage institutions (Shatford, 2002).

Fidel (1997), an early pioneer in image retrieval, described how search/retrieval may be influenced by search behaviour. They suggested that image retrieval occurs along a spectrum that conceptualises an image as information to image as an object. Importantly these two poles of the spectrum can encompass more aesthetic aspects of images (colours, shapes) compared to informational details pertaining to specific objects (e.g., date, location etc.). This idea of different types and levels of detail represented by an image is rooted in previous research on describing and categorising the subjects of images, particularly from Panofsky (1972) who initially described three levels of detail apparent in Renaissance paintings. These categories included descriptions that are 'pre-iconographical description' (visual elements), 'iconographical analysis' (symbolism) and 'iconological interpretation' (significance of the subject matter).

Drawing on these three levels, Shatford (1986) expanded the framework to the analysis of a wider breadth of images using accessible words such 'generic of', 'specific of',

'about' and further broke each of these levels down into four categories/facets including: 'who', 'what', 'where', and 'when'. This is commonly referred to as the Panofsky-Shatford model/matrix, pictured below.

	Iconography (Specifics)	Pre-iconography (Generics)	Iconology (Abstracts)
Who?	individually named person, group, thing **(S1)**	kind of person or thing **(G1)**	mythical or fictitious being **(A1)**
What?	individually named event, action **(S2)**	kind of event, action, condition **(G2)**	emotion or abstraction **(A2)**
Where?	individually named geographical location **(S3)**	kind of place: geographical, architectural **(G3)**	place symbolised **(A3)**
When?	linear time: date or period **(S4)**	cyclical time: season, time of day **(G4)**	emotion, abstraction symbolised by time **(A4)**

Figure 4-1 'Panofsky-Shatford model/facet matrix' (Armitage and Enser 1997, p. 290)

In turn, others such as Jorgensen (1998), have analysed descriptions of images to suggest new descriptive categories along similar lines, including: 'perceptual', 'interpretive', and 'reactive' (p. 168). While all frameworks share similarities, the Panofsky-Shatford model has been used multiple times in both the analysis of user queries and the application of metadata to images in the cultural heritage field. Enser and Armitage (1997) for instance, successfully applied the Panofsky-Shatford Matrix to the analysis of archive image queries. They conclude that:

> *'the dual applicability of the schema to the characterisation both of images and the queries which address those images offers some encouragement to the view that the embedding of such a schema within the user interface might offer a pathway towards the effective processing of unmediated transactions within a visual information retrieval system'* (p. 287-288).

Klavans, LaPlante and Golbeck (2014) tweaked this matrix further in a way that is useful for our project. They analysed user-contributed tags (also called social tags) applied to 100 images of artworks supplied by the Steve Museum Project.[4] In their analysis of social tags, they added two additional facets to the Panofsky-Shatford matrix: this included 'visual elements' - such as colour or shape visible in imagery, and an 'other' category to allow for additional tags that may fall outside the other facets such as methods used to create the image (Figure 4-2). While only a small percentage of the social tags analysed in their project fell into these additional categories, the presence of these categories makes space for considering further descriptions/descriptive words that may be relevant to users.

[4] For greater description of the Steve Museum project see: https://www.steve.museum/

TABLE 2. Subject categorization matrix (Armitage & Enser, 1997; Shatford, 1986).

		(G) Generic (Preiconography)	(S) Specific (Iconography)	(A) Abstract (Iconology)
1	Who?	G1: Kind of person or thing *(woman, dog, sunlight)*	S1: Individually named person, group, thing *(Mona Lisa, Daddy the pit bull, American)*	A1: Mythical or fictitious beings *(Clio the Greek muse, unicorns)*
2	What?	G2: Kind of event, action, condition *(football game, war, buying and selling, death)*	S2: Individually named event, action *(2009 Super Bowl, Battle of the Bulge, JFK's funeral)*	A2: Emotion or abstraction *(a woman buying fruit representing commerce or a funeral representing sorrow)*
3	Where?	G3: Kind of place: geographical, architectural *(city, jungle, inside)*	S3: Individually named geographical location *(New York, Mars)*	A3: Place symbolized *(Mount Olympus, paradise)*
4	When?	G4: Cyclical time: season, time of day *(spring, night)*	S4: Linear time: date or period *(June 1885, Renaissance)*	A4: Emotion, abstraction symbolized by time *(spring symbolizing youth or fertility)*
V	Visual elements	V: Shapes, forms, and colors *(square, line, red)*		
O	Other	O: Items that do not fit into the above categories		

JOURNAL OF THE ASSOCIATION FOR INFORMATION SCIENCE AND TECHNOLOGY—January 2014 7
DOI: 10.1002/asi

Figure 4-2 Subject categorization matrix reworked by Klavans et al. (2014), pg. 7.

These three levels of description (generic, specific, and abstract) pertaining to subject matter of imagery, as described by Hourihane (2002), are often prioritised by cultural heritage classification systems. Shatford (2002) suggests for instance, that 'of-ness' encompassed by the 'generic' and 'specific' levels falls into the 'Description' or 'Identification' categories of the classification system: Categories for the Descriptions of Works of Art (CDWA). While she suggests that 'aboutness' pertaining to the 'abstract' level can fall into interpretation which is a subcategory of Subject Matter in CDWA (Shatford, 2002). Therefore, the above framework is central to this study due to its potential dual application in understanding user queries and in turn, applying relevant captions and keywords to images that are catered towards users' interests and needs.

4.4 Inclusion/exclusion and vocabulary

When it comes to understanding user search behaviour, attention is often placed on the *semantic gap* (Klavans et al., 2014) – the gap between terms used by the public and those used by professionals/specialists of cultural heritage institutions. This gap may be due to different communication models: Dahlgren (2022) suggests users' active subjective meaning-making based on their context may contradict the perceived transmission model/objectives of cataloguers. This coincides with and has been connected to a conflict between expert/public terminologies applied to collections' metadata (Klavans et al., 2014). Scholars, practitioners, and communities have frequently highlighted that terminologies used by cultural heritage institutions might not suit the interests/needs of users, and further, that terminologies may be contradictory and even offensive to how individuals and groups refer to themselves or their cultural heritage (Martin, 2021). Such terms are historically rooted in legacies of colonialism, violence, and racism that underpin the museum institution, and in turn, are perpetuated by sharing outdated collections information (Turner, 2016, 2020). There is thus, recognition in the cultural heritage field that applying metadata is a non-neutral/interpretive practice and that these practices

should incorporate user/community-generated vocabularies and be transparent regarding outdated terms (Martin, 2021; Vaughan, 2018).

Museums have tried to rectify the discrepancies between institutional and user terminologies in several ways. A core approach, initially popular in the early 2010s was crowdsourcing and social tagging. Some of these practices are still taking place today - see for instance, Art UK's recent and ongoing social tagging initiative.[5] Crowdsourcing projects for tags and metadata, while appealing, do not guarantee that the resulting metadata will meet diverse audiences' interest and/or needs. As Dahlgren and Hansson (2020) argue, a larger crowd does not necessarily correlate with overall diversity. However, community consultation, participation, and the publication/advocation of user-, community-, and group-generated terminologies and subject lists, can provide essential changes to cultural heritage systems that can make a difference to different publics and community groups.

As advocated by Martin (2021), metadata professionals have a responsibility to accept their practice as non-neutral and to actively 'use their judgment to note in the cataloguing records the prejudices, biases, or misleading nature of materials' (p. 14). This is similarly supported by Spectrum procedures which suggest rather than erasing harmful and/or outdated terms, to 'make a transparent record of any changes and/or provide appropriate context' (Collections Trust, 2022a). To counter legacies of outdated, harmful, and inaccurate terminologies, there are now various vocabularies/glossaries dedicated to updating terms associated with LGBTQ+ histories and collections, colonialism, and indigenous cultures/peoples. For instance, regarding LGBTQ+ terms, there is the V&A's LGBTQ+ terminology, the Queer Metadata Collective project (North America focused), and the Homosaurus (International LGBTQ+ Linked Data Vocabulary). Carissa Chew's Inclusive Terminology project attempts to build multiple crowdsourced glossaries. Local Contexts is an online platform and set of tags/labels that supports Indigenous communities to implement protocols and conditions for access and use to their cultural heritage circulating digitally.

Additional resources include the Anti-racist Educator Glossary created by educators in Scotland, and a working glossary to describe the impacts of colonialism by the Curatorial Research Centre. As mentioned in the National Museums Scotland context (section 3), the institution has been engaging with decolonial discourse and adjusting terminologies and access to certain collections images. The glossaries listed here, in keeping with the Museums' commitment to decolonialisation and inclusivity could further inform the keywords and captions applied to images of the Museums' collections.

4.5 Museums staff time and depth/richness of metadata

One central challenge to updating metadata in ways that increase the findability and retrievability of the collection images, while being sensitive to the terminologies, emotions and meaning making of users, is the limits of staff time. As is discussed in a developing document on the ethics of metadata professionals by communities of practice,[6] staff must

[5] https://artuk.org/for-collections/artwork-tagging
[6] See: https://docs.google.com/document/u/0/d/1IBz7nXQPfr3U1P6Xiar9cLAkzoNX_P9fq7eHvzfSlZ0/edit?pli=1

make decisions regarding the level of depth, detail, and categories captured based not only on users' interests and needs but also staff time and resources.

As Martin (2021) reflects, shrinking budgets and backlogs might require cataloguers to agree on the detail and specificity that will be described. One tool that is being incorporated into museum catalogue practices which harbours the potential to relieve some of this staff labour is artificial intelligence (AI). AI in the form of computer vision can analyse and group together similar objects based on a range of features. AI can identify aesthetic features of images, faces, landmarks and more 'offering curatorial staff novel ways of analysing, researching, and describing museum collections' (Villaespesa & French, 2019, p. 103).

Recently, Villaespesa and Crider (2021) tested different forms of AI for its abilities to tag the Metropolitan Museum of Art's collections and how these compared to human generated tags. They compared these to the Getty's Art and Architecture Thesaurus (AAT) and the categories of terms outlined in the Shatford-Panofsky matrix (discussed in section 4.3) for describing subjects of images. They conclude that,

> *'computer-generated subject tags fall into pre-iconography, and specifically the factual category, that is, describing what the actual entities, attributes and actions featured in the image are. There is no information about the iconography and iconology of the artwork. Similarly, looking at these results from Shatford's classification, these tags are categorized under the "of-ness" but lack information about the "about-ness" of the artwork'* (p. 958-959).

In turn, they suggest pre-iconographic tags or 'general' tags could be fulfilled by certain computer vision, leaving the specifics and more abstract/iconological facets of image subject description to museum cataloguers (Villaespesa & Crider, 2021).

However, just as cataloguers may tag and describe images in ways that could be inaccurate and/or offensive, so too can computer vision, as algorithms that underpin AI are similarly created by people with biases (implicit and otherwise). For instance, Ciecko (2020) discusses museum projects using computer/machine vision and describes some of the biases that were made evident by its poor ability to label non-western art. He suggests: 'while machine vision may unlock new potentials for the cultural sector, it is essential to scrutinize the ways that machine vision can perpetuate biases, conflate non-Western cultures, and generate confusion' (par. 1). While computer vision could be useful for National Museums Scotland's staff and for carrying out the aims and results of this project, sensitive collections having to do with challenging topics whether about oppression or colonialism, may be better described and tagged by cataloguers.

5 Research methods and findings

For this project we drew on five main methods to investigate current and potential user interests and needs, and the needs of staff in relation to the collections image metadata. This included: interviews with the Museums staff, content analysis of user search queries collected over 2020-2021 from the institutions' Search our Collections webpage, user

surveys, and user interviews involving a concept-mapping and card sort activity. The ethical processes of the research project were reviewed and approved by The University of Glasgow's College of Arts Ethics Committee.[7] The following section 5 details each method and associated key findings in turn. It is followed by the prototyped frameworks in section 6 and the recommendations for practice and conclusions in section 7.

5.1 National Museums Scotland's staff perspectives on collections metadata

5.1.1 Method

We undertook 13 individual/group interviews involving 22 members of staff particularly those who use images of the collections for engaging with the public and/or who contribute to the collections' metadata. Interviewees were based in the following departments/roles:

- Collections management
- Digital media
- Marketing and communications
- Exhibition design
- Library services
- Photography
- Curatorial
- Learning and engagement

The interviews followed a set of semi-structured questions pertaining to the following two themes/goals, tailored to staff roles:

> **Goal 1: understand current uses of image metadata in staff's everyday work.**
> This goal encompassed understanding staff roles and their current use and challenges encountered when using the current collections' metadata/search interfaces. For staff working on metadata of collections, questions also addressed methods of describing, tagging, and indexing.
>
> **Goal 2: investigate how staff pursue user access.**
> This goal encompassed what is/has been done to make the online collections accessible for users, perceived challenges for audiences in accessing images of the collections, key goals for future user access and possibilities of enabling this access through keywords/captions.

During these interviews, staff discussed their experiences with the Search our Collections (SoC) interface and/or Axiell (the Museums' collections information system) when creating programs, events, exhibitions, content, or communicating with the public, and reflected on their own search challenges. The resulting interviews touched upon five main topics discussed further below, including methods of searching/interacting with the SoC portal, the affordances of the search portal, the ability to cater to users (interests and needs), and suggested ways to approach applying keywords and captions to images of the collections.

[7] Ethics Application 100220077

5.1.2 Search/interaction/use

In reflecting on their own experiences and those of visitors, staff advocated for the ability to browse and follow one's nose through the online collections. In relation to this recommendation, staff suggested the need to go beyond subject categories to enable more ethereal, nebulous, or abstract searches, or be able to follow thematic threads through the collections. Other related suggestions included to interlink the online collections with different forms of information and media across the website (e.g., images, video, sound, 3D digitisations, blog posts). Beyond general browsing, staff also emphasised the importance of being able to undertake specific searches (with something in mind) or to narrow down a broader search through various facets – such as exhibition, provenance, dates etc. As described by one staff member,

> *'I guess there will be other types of users who know exactly what it is they want to pinpoint, but my view in that is that we need to give people much more ability to browse things and lose themselves and follow their nose, because I think that would be much more engaging for a lot of people' (Jo, Collections Data Manager)*

> *'Apart from anything else there's just far too much white space [on the collections webpages] and you're either coming to a dead-end or end up going off on a tangent, it's quite nice going somewhere else - but not necessarily somewhere else that you have no idea, no intention of going' (Neil, Lead Photographer).*

These methods of searching are intertwined with the affordances of the collections systems, the search interface, and the API as discussed in the following theme.

5.1.3 System affordances

Closely entangled with different methods of searching or browsing, staff's comments also critically reflected on the affordances of the SoC portal. This included critiques of the portal as being 'unhelpful', with limited entry points into the collections (for example, not recognising misspellings or text chains), and a lack of clarity regarding image rights pertaining to use and reuse. Conversely, staff also reflected on what a 'helpful' system might look like – offering suggestions in terms of related words, objects, or images, and to provide alternative entry points into the collections for users who do not have a specific search in mind. Some examples staff shared included 'Google Luck Dip', the Science Museum's Random Object Generator, Meow Met, and displaying trending searches visually in word clouds or nets. As a few staff reflected:

> *'That's the problem - it [the SoC portal] says "explore our collections" and you go into it and you're like "I don't know how to have a casual browse of this"' (Staff member, Learning Enabler).*

> *'Do you know Google Lucky Dip? Is that what it's called? Where you just hit it and it'll bring up a random webpage you've no idea what you'll get. That might be quite nice! Like an object lucky dip' (Staff member, Learning Enabler).*

5.1.4 Catering to user interests and needs

As the staff made evident in the interviews, there is limited knowledge of online users of the Search our Collections portal, and a desire to know more - including the types of images searched for and the search terms used. Taking into account the networked and competitive nature of online education/entertainment, some staff reflected on the need to reach out and go beyond the Museums' own digital estate, enhancing the Museums' networked nature.

> *'From our point of view, we don't really know how people get our images and how they go looking for them. I guess to find that out would certainly be very useful because it could inform the next steps we take. It might also help us to inform where we might start with the metadata of the images themselves, if we knew how people were expecting to find them' (Angus, Collections Systems Manager).*

> *'I think it's about thinking about who our users actually are and who we want them to be because I think we have a lot of academics that come to search our collections … but then we'll get emails, saying I want to see dinosaurs with my son, and I don't think they've even had a go at the collections because it doesn't feel like it's for them' (Hannah, Collections Data Specialist).*

> *'I would expect us to then look to create content for specific audiences but think beyond our own digital estate as well. And think, "if that's not the audience that we're connecting with - where do we go to reach those audiences?" Do we create content to be published on other sites that already have that audience and bring them in that way? That seems like an obvious thing to me' (Staff member, Interim Head of Digital Media).*

Reflecting on their own experiences and knowledge of other collections, staff emphasised the importance of having many high-res photos with varying levels of detail and styles available through the Search our Collections portal. Some staff reflected that they themselves and users might be more interested in artsy photos that are 'striking' as opposed to 'archival' photos.

Staff also frequently spoke about the need to understand the terms users use in searching the collections, to enable user tagging and suggestions, and to apply accessible terms to the collections metadata.

> *'When you're writing labels, maybe the Search our Collections should be written with a seven-year-old in mind, what kind of language would they be using in case they want to find something more specific - but for the most part, think about what your average 7- or 8-year-old is going to be able to know about' (Staff member, Learning Enabler).*

Further, they reflected on the need to meet user interests that are shaped by current events or broader themes. This included topics such as colonialism and climate change, and associated tags, as further discussed in the following sections on staff's suggestions for keywords and captions.

'But thinking about our other work on colonial history and legacies and where it might be relevant to tag things with those kinds of keywords which are not as temporary or transient - that will still make sense and still be relevant in a collection's context in 10 or 20 years' (Staff member, Interim Head of Digital Media).

5.1.5 Good keywords should…

When discussing keywords with staff and their relation to enhancing the findability and retrievability of the collections, many staff discussed describing the collections images in ways that move beyond the Museums' main subject categorisation. This included describing the visual aesthetics of the image, such as the emotion associated or style of the image, textures of the object, and what's visible in 'general terms', rather than specific or 'factual' details of the object (provenance, donor/creator, dates). As staff reflected, image metadata should be:

'descriptive from an aesthetic point of view as opposed to descriptive from a factual point of view' (Karen, Exhibition designer).

And that images could be described by different aesthetic features:

'Colour and emotions, we've looked at doing those things - when you think about the Marketing and Communications team, they have loads of images and they may be looking for images of people that are smiling, or where there's a child, or a building or a rainbow' (Pam, Collections Data and Digitisation Manager).

'Surely you can put 'weird' in that one or something like that, so that when we search it, it just comes up with some really funny weird stuff' (Russell, Digital Media Manager).

'I would probably think about the fabric of what you're looking at. So, what the materials, the feel of something - it's not just the visual but the touchy feely of something. What is it like to hold?' (Neil, Lead Photographer).

'I'm not sure if the V&A does this, although they do kind of tag images in a way that you can click on a designer - and that will bring up everything by them or certain keywords which is actually really handy for the Picture Library. So, it's not just who the photographer is, or the designer but you can search by the model or what's actually in the image itself' (Carys, Assistant Curator of Modern and Contemporary Design).

This also included associating the image with topically relevant events or themes such as climate change, colonialisation/decolonialisation, words associated with holidays, and school/curriculum terms also used across the Museums' blog:

'A silly example is when we are planning content around particular days of the year, so you know things around Christmas, or St. Andrews day, or Burns Night, we create content around not all significant days but some where we have a connection to those stories. Then it might be a case of looking into the collection in certain areas to try and

pull out interesting images that might be able to form the basis of posts in that sense' (Staff member, Interim Head of Digital Media).

'We look for a lot of dates throughout the year, that we can then engage with on social media - like pancake day, looking for images related to pancakes. If there was an option, if a curator or whoever is managing the image to say this is quite a fun piece, it's cat-related, or pancake-related - terms like that attached to the objects could be useful for identifying objects that we can then use in our social media posts but that's for the digital team …' (Staff member, Marketing and communications).

Despite the need for the collections to be topically relevant, staff emphasised the need to choose relevant but not temporary terms. Alternatively, some staff critiqued current metadata of the collections themselves, as lacking information pertaining to the 'why' of a collection object or image, and its general significance.

5.1.6 Good captions should…

Similar to keywords, staff suggested that captions should focus on aesthetics, taking inspiration from alt-text. In turn, staff reflected on what is currently included in the metadata and what might be unnecessary for the captions of collections' images. This included 'unnecessary details' such as specific place, production, and object type.

'I think the description [image caption] could do what the object description doesn't which is explaining what the image is and how the object has been captured' (Hannah, Collections Data Specialist).

'I suppose I'm more interested in what the alt-text might say, which is more a descriptive idea of what the image is showing as opposed to the technical - which is a specific object that's used in this way, that's not going to mention that it has an image on the cover, or it has this kind of colouring' (Karen, Exhibition Designer).

Some staff also advocated for more participatory work to develop image descriptions and evaluate descriptions of objects, including working with visually impaired groups.

When analysing the information from these interviews with staff and in discussions with the collections team, we found the Panofsky-Shatford matrix (further explained in section 4.3) helpful for visualising staff's comments and pinpointing the image descriptors missing and desired for the collection images' metadata. In particular, what could be referenced as 'missing' from the metadata includes more general terms for type of things visible in the image from an aesthetic point of view (person, vase, ring) rather than 'specifics' (for example, specific person, specific place, date ranges etc.). Further, in reference to browsing and 'following one's nose', although system affordances are out with the scope of this project, the underpinning metadata such as keywords associated with broader themes, emotions, or more visual colours and shapes may enable users' ability to follow different threads whether narrative- or more visually based.

5.2 Search our Collections queries analysis

5.2.1 Method

For this project we drew on a pre-existing anonymised dataset from the AHRC TaNC research project reported in 'Digital footprints and search pathways …' by Chowdury et al. (2022), which required extracting search queries from the Search our Collections (SoC) portal and documenting the frequencies of these queries from 2020-2021. We used this data to achieve two goals:

> **Goal 1:** understand the types of terms searched (general/specific – who, what, where, when, or visual).
> **Goal 2:** understand which collections and objects are being searched for.

To achieve these goals, we used a sample of this large dataset of user queries and undertook a deductive content analysis based on the Panofsky-Shatford framework tweaked by Klavans et al. (2014) (figure 5-2) for defining level of detail/description of subject matter portrayed by images and applied this to user queries. Like Klavan's et al. (2014), we applied only one perceived main code or category to each user query. We subsequently coded the queries for common general subject matter.

Overview
Total unique search queries for 2020-2021: 70,874
Total search queries including repetitions for 2020-2021: 570,990

Total unique queries in our sample: 1,779
Total queries in our sample including repetitions: 168,970

The sample encompassed search queries with:[8]

- frequencies from 90 - 1909 (search queries with higher repetitions)
- frequencies from 15 - 30 (search queries with lower repetitions)

Tweaks to the subject description matrix (figure 5-2)

- In category S2 (individually named people, group, or thing) – named objects and individual people were uncommon in SoC queries, but it was useful to use this 'specifics/detailed' category to place species' names, technical/complicated terms, and culturally specific terms for objects, and very specific materials.
- In category A1, spiritual objects and beings were grouped together, as they could be connected to more abstract ideas (see key word list, Appendix C)
- In category G4, we grouped words for holidays, seasons, and general time periods (e.g., medieval) as these are missing from the Museums' collections metadata, placing only specific numeric dates/date ranges in S4.

5.2.2 Key findings

In the sample, one-word queries of general or non-specific terms are most common (e.g., dress, clothing, ring, sword) (see Appendix C for examples). In some frequently searched

[8] A sample of the most popular search queries and less popular search queries were taken in order to understand broadly the *types* of terms used on the SoC webpage – not only the most popular objects searched.

object categories, such as photography, materials, and minerals, more technical and specific language was used (e.g., ambrotype). General terms most frequently fell into describing general 'who' – types of people, objects, or things (G1) (circled in figure 5-2) and much fewer 'what' - actions/events (G2), 'where' – places (G3), 'visual elements' - colours etc. and 'other' - such as photography style. There was a lack of abstract terms overall, including ideals/symbols, and emotions, however, there were many queries for spiritual or religious objects and figures, and objects related to death including the concept of 'death' itself. Overall, the most popular or frequently searched collections could be grouped together under 'clothing', 'accessories', 'dishware', 'materials', 'tools/technologies', 'collections subjects', 'art', 'containers', 'animals/insects' and 'instruments (figure 5-1). The interest in spirituality and religious figures may have been due to the data being collected in the context of Covid-19 and, the interest in clothing may be a result of staff using the SoC portal to prep for an upcoming exhibition – 'Beyond the Little Black Dress'.

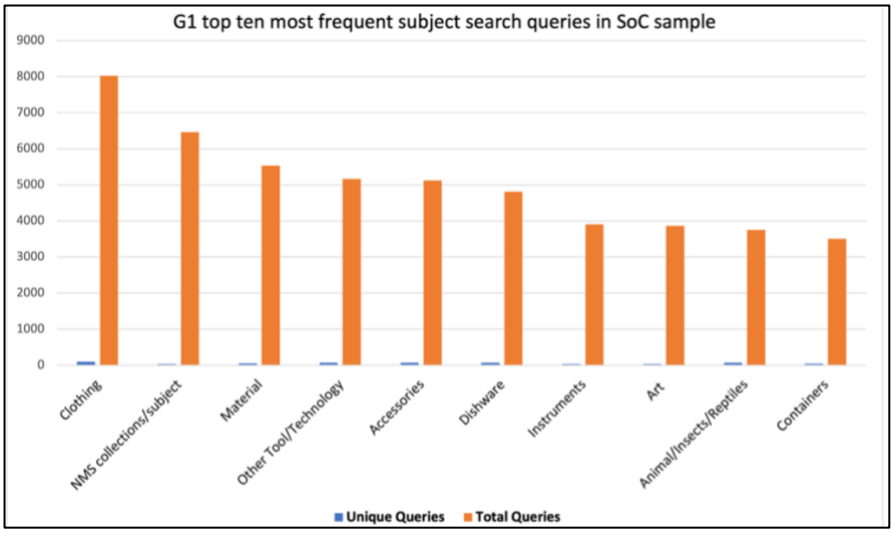


Figure 5-1 Top general 'who' (G1) subject search queries

TABLE 2. Subject categorization matrix (Armitage & Enser, 1997; Shatford, 1986).

		(G) Generic (Preiconography)	(S) Specific (Iconography)	(A) Abstract (Iconology)
1	Who?	G1: Kind of person or thing *(woman, dog, sunlight)*	S1: Individually named person, group, thing *(Mona Lisa, Daddy the pit bull, American)*	A1: Mythical or fictitious beings *(Clio the Greek muse, unicorns)*
2	What?	G2: Kind of event, action, condition *(football game, war, buying and selling, death)*	S2: Individually named event, action *(2009 Super Bowl, Battle of the Bulge, JFK's funeral)*	A2: Emotion or abstraction *(a woman buying fruit representing commerce or a funeral representing sorrow)*
3	Where?	G3: Kind of place: geographical, architectural *(city, jungle, inside)*	S3: Individually named geographical location *(New York, Mars)*	A3: Place symbolized *(Mount Olympus, paradise)*
4	When?	G4: Cyclical time: season, time of day *(spring, night)*	S4: Linear time: date or period *(June 1885, Renaissance)*	A4: Emotion, abstraction symbolized by time *(spring symbolizing youth or fertility)*
V	Visual elements	V: Shapes, forms, and colors *(square, line, red)*		
O	Other	O: Items that do not fit into the above categories		

JOURNAL OF THE ASSOCIATION FOR INFORMATION SCIENCE AND TECHNOLOGY—January 2014 7
DOI: 10.1002/asi

Figure 5-2 Subject categorization matrix reworked by Klavans et al., p. 7, 2014 with G1 circled

5.3 User/non-user survey

5.3.1 Method

To understand not only the terms used in search queries but also who is using the collections, their motivations, and methods of interaction/search, we distributed two surveys - one through social media/our project team's networks and one as a pop-up on the Museums' Search our Collections portal for a period of around 7 weeks (3 April – 26 May 2023). Both surveys were similar, with the exception that the survey distributed over social media allowed people who had never used the collections portal to participate and explored their interests in cultural heritage/collection images (since the surveys were similar, only a copy of the webpage survey is included in Appendix A).

The surveys addressed the following questions/goals:

> **Goal 1:** understand who is using the collections, and who filled out the survey.
> **Goal 2**: understand how users search the collections (terms and method).
> **Goal 3:** understand how non-users might want to search and use images of the collections.
> **Goal 4:** understand what themes participants are interested in searching the collections by.

To address these goals the survey consisted of several key parts including: questions exploring users' motivations for visiting the collections online; users' 'roles'; their knowledge of cultural heritage; methods of searching/browsing; an activity involving keywording four images of the collections (see survey in Appendix A); an opportunity to provide suggestions for the online collections' portal; and an optional demographics section.

5.3.2 Key findings

- 132 survey participants
- 81 had used the online collections/51 had not
- Mainly museum/heritage professionals followed by curious individuals responded to the survey (figure 5-3)
- Users are mainly visiting for work followed by general interests in a topic, followed by for school/schoolwork (figure 5-4)
- Users of the SoC portal are mainly based in Scotland, are Scottish, female, aged 25-34 & 35-44 (figures 5-5, 5-6, 5-7)[9]
- In the keyword activity, general terms for people, objects and things were most popular, followed by specific/technical words for objects, followed by general terms for action or event (see *Types of keywords applied to images* for further details)
- No significant difference in terms/types of terms used between people who describe themselves as users with 'general' knowledge of cultural heritage vs those with 'expert' knowledge of cultural heritage (figure 5-10 & 5-11)

[9] While this is likely a core demographic of the online SoC, our surveys were biased by our distribution methods (though our own and the Museums' social media and museum mailing lists, reflecting our own networks), and some core groups such as grandparents/parents/teachers, and artists/creatives, were potentially missed.

- Four main user groups of the online collections are identifiable from our survey sample (see next page)
- Most popular theme to explore images of the collections by is ‘under-represented histories’ (figure 5-8)
- Non-users of the collections were primarily interested in using collections images for discussing cultural heritage with others, followed by learning about a particular subject (figure 5-9)

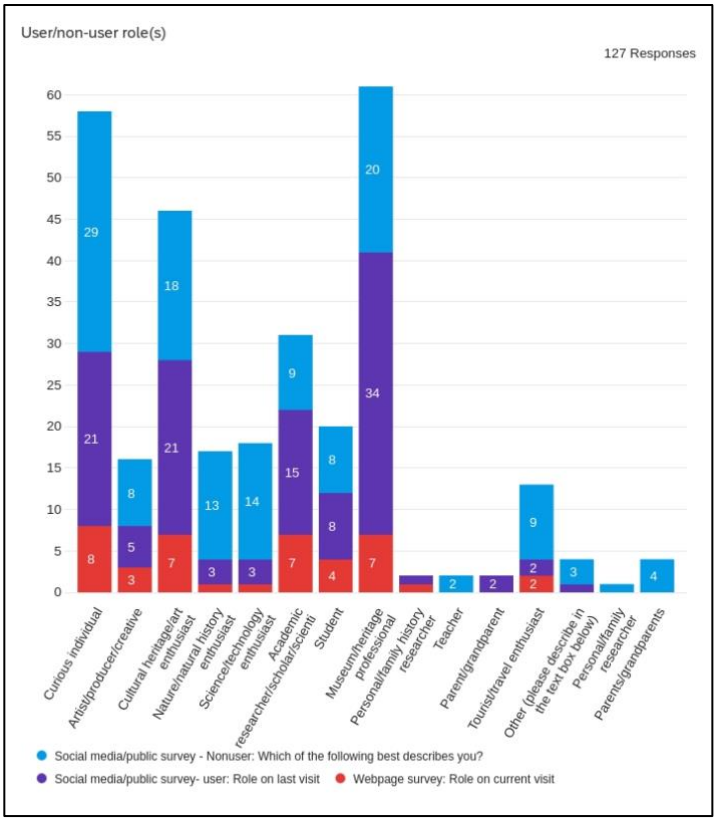


Figure 5-3 Chart depicting self-defined roles of survey participants

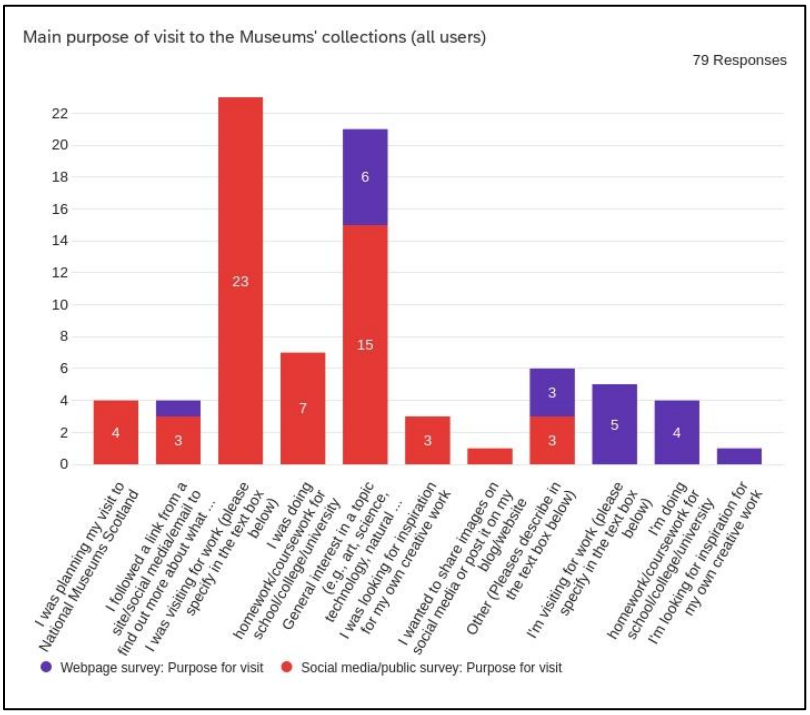


Figure 5-4 Chart depicting users' motivations for visiting the online collections

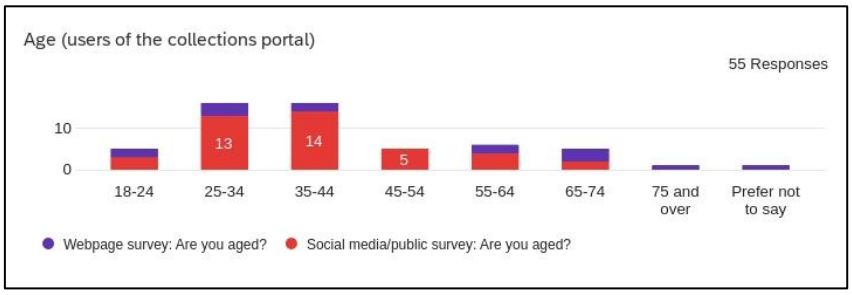


Figure 5-5 Chart depicting the age of users of the online collections

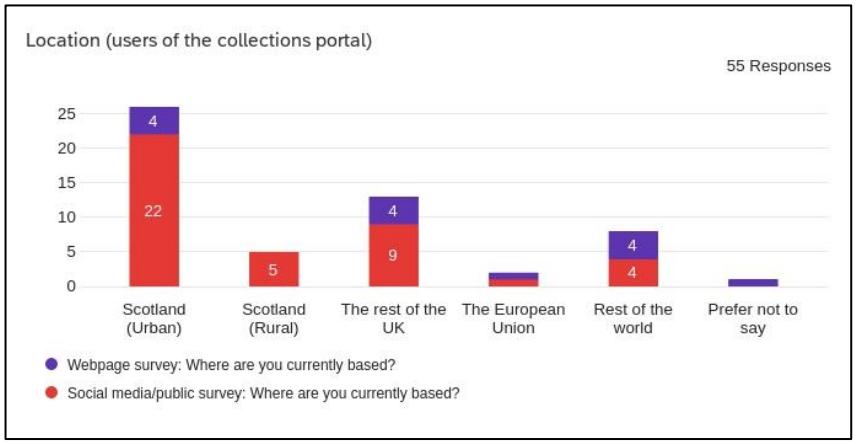


Figure 5-6 Chart depicting the location of users of the online collections

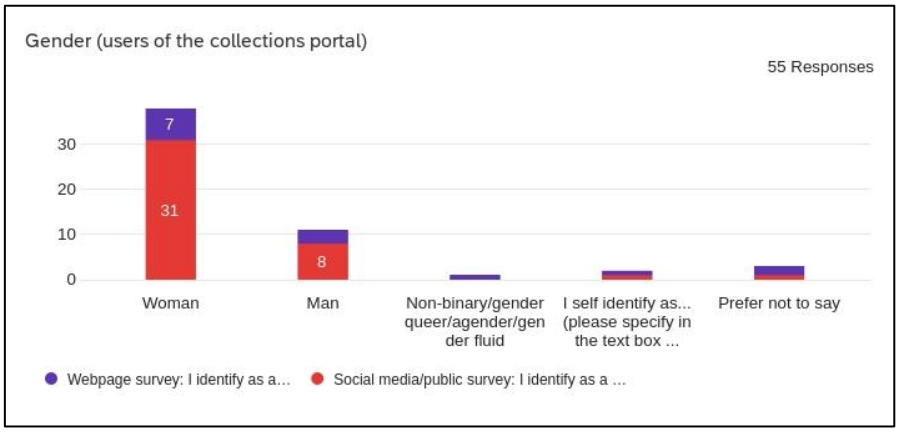


Figure 5-7 chart depicting gender of users of the online collections

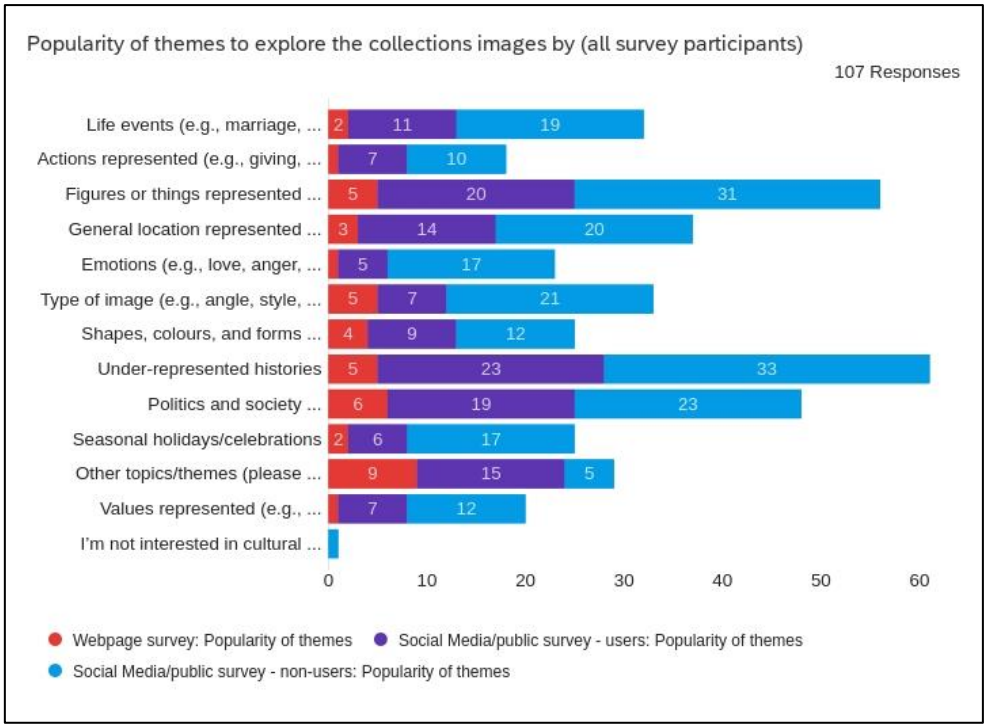


Figure 5-8 Chart depicting the popularity of themes by survey participants

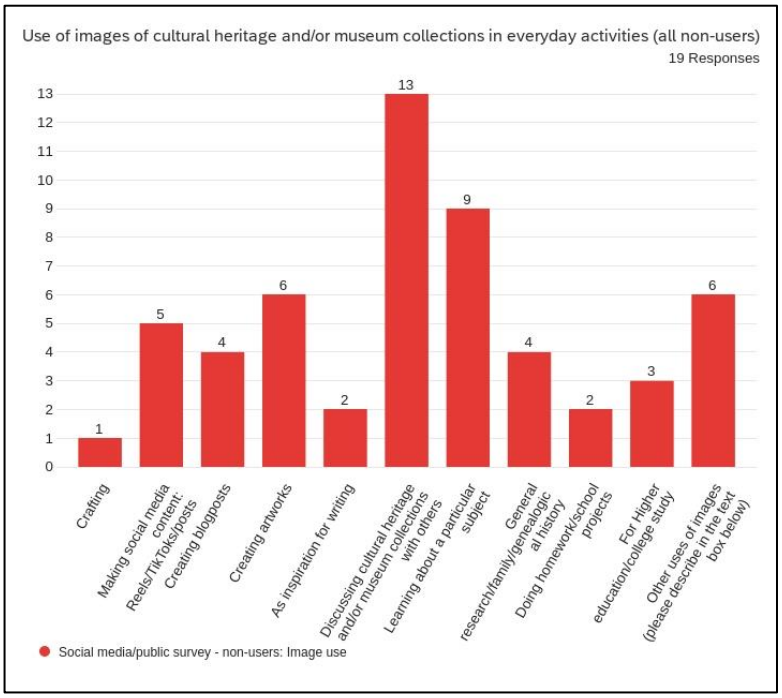


Figure 5-9 Chart depicting how non-users use images of cultural heritage in everyday life

There are four main user groups identifiable from the surveys: these groups are composed of users who have similar interests, levels of knowledge of cultural heritage, motivations, and methods for searching the collections, and use images in similar ways (See Appendix E for additional charts related to these groupings).

1. Museum/cultural heritage professionals

- Mainly visiting for work, followed by general interests
- Mainly have specialised knowledge
- Mainly searching by specific object/object type and advanced searching
- Mainly using images to discuss cultural heritage with others, followed by learning

2. Curious individuals

- Mainly visiting for general interests, followed by work
- Make up the majority of general knowledge users
- Mainly search by specific object/object type, followed by advance search, and searching/browsing general topic or subject
- Mainly using images to learn about a particular subject, followed by discussing culture/art with others and making social media content

3. Cultural heritage/art enthusiasts

- Mainly visiting for general interests followed by work
- Mainly have general/little knowledge followed closely by specialist knowledge
- Mainly searching by specific object/object type and advanced search, followed by browsing featured collections and by general topics
- Mainly use images for creating social media content, followed by creating art and discussing arts/culture with others

4. Students/academics

- Mainly visiting for school or work
- Mainly have specialised knowledge
- Combo of mainly searching and advanced search, and a little bit of browsing by general topics
- Mainly use cultural heritage images for discussing with others, creating art or social media content

Summarising user search methods and image use

Overall, most users in this sample had some idea of what they were looking for, using the search box as opposed to canned searches and did more than one 'search activity' - making visiting the online collections complex. This may be due to initial searches being unsuccessful, or to the fact that undertaking multiple activities can be an essential part of creating satisfying user experiences. The inability to find the objects/information that survey participants were looking for is evidenced by some users copying and pasting the same terms from the general search into the advanced search box. Users were commonly looking for more than one type of information: often images, but also other content like links to other resources, additional context (e.g., who wore it? Any relations to social class?), the objects' provenance and associated dates which may be due to users mainly visiting for work or school. Some participants wanted to search for objects from a particular location or culture, from a specific exhibition, or associated with a specific person (e.g., Mary Queen of Scots). A few just wanted to have a look – to be 'inspired' and to see something that 'caught their eye'. The main use of images was to discuss cultural heritage with others, to learn, and to undertake more creative activities such as creating social media/blog content and art – insinuating the importance of having collection images that are reusable/shareable.

Types of keywords applied to images

As part of the survey, participants were invited to apply keywords to four images of the collections (see survey, Appendix A). The keywords applied to the images were also coded and categorised based on Klavans et al.'s (2014) matrix (with adjustments as described in section 5.2.1) and their frequencies compared. It is interesting that there was little difference in terms used by different user groups, by different levels of knowledge, or by users with different 'roles'. For example, see comparative word clouds below with few differences between 'general' cultural heritage knowledge users and 'expert' knowledge users.

Figure 5-10 Keywords applied by general knowledge participants to Task 1 (mariachi figure)

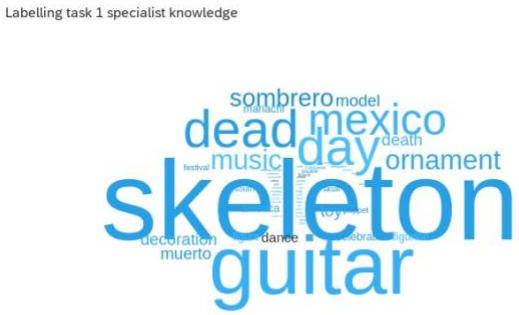


Figure 5-11 keywords applied by specialist knowledge participants to Task 1 (mariachi figure)

However, the types of terms applied did differ significantly between the images, based on what users picked out as being most prominent. See the two pie-graphs outlining the quantity of types of terms used between the figure and vase using Klavans et al.'s (2014) tweaked matrix for image description (view the remaining word clouds in Appendix E).

Task 1: Applying keywords

Title of object and link to online catalogue: **Figure**
Object description: Miniature guitar-playing mariachi figure, with spring-wire neck, plastic, made for the Mexican Day of the Dead festival: North America, Mexico, Mexico City, 1999

Figure 5-12 A mariachi figure from the Museums' online collections

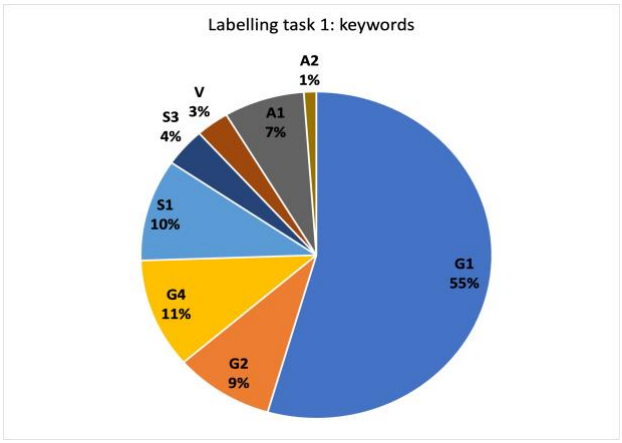


Figure 5-13 Percentage of types of terms applied in labelling Task 1

Task 2: Applying keywords

Title of object and link to online catalogue: **Vase**
Object description: Coloured and glazed porcelain vase decorated with 'Love's Wheel of Fortune' in pâte-sur-pâte and on neck and base in Pompeiian style with coloured slips and gold; English, Stoke-on-Trent, by Minton and Company, decorated by Marc Louis Solon, 1876 – 1877.

Figure 5-15 Vase from the Museums' online collections

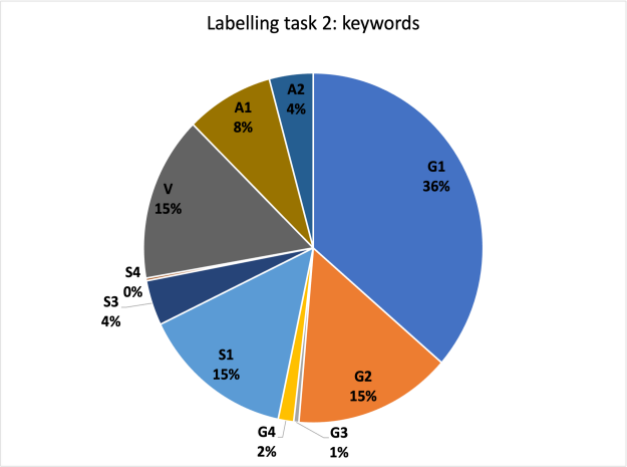


Figure 5-14 Percentage of types of terms applied in labelling Task 2

- **Overall, G1 or general terms for 'who' person, place, or thing** was the most common category of terms applied in all the labelling activities.
- This was frequently followed by **S1 – encompassing a specific place, person, people, or thing**. This included specific materials, dates, cultural terms.
- After this, **G2 General event, action, and condition** was prominent: with users putting down words such as celebration, festival, and death.

Additional categories of keyword terms that were present but less evident:

- **(G4 – General when)** There were some references to holidays and seasonal events (e.g., Halloween, Day of the Dead), and periods such as Victorian or Medieval.
- **(A1 Abstract/mythical/religious people, objects, things & A2 Abstract events, actions, emotions, concepts)** Mythical people, places, things, spiritual beings, emotions, and abstract ideas were most common in labelling task 2 (vase, Figure 5-14) which had a relief of cherubs and a wheel of love.
- **(V – Visual terms)** Visual terms were most common in labelling task 2 (vase, Figure 5-15) and labelling Task 4 (agate, see Appendix A), with lots of tags associated with the aesthetics of the rock.

These observations align with the findings from the Search our Collections (SoC) queries, except that overall, visual terms such as colour and more abstract terms were very minimal in the SoC queries and played although small, a more prominent role in the survey sample.

5.4 User interviews (concept mapping and card-sort)

5.4.1 Method

Following the surveys, 10 interviews were undertaken with mainly survey participants and a few participants from the researchers' own networks, encompassing:

- 3 Museum professionals
- 1 Librarian/Archivist who is also a fiction writer
- 1 PhD student
- 1 Post-doctoral researcher
- 4 Cultural heritage enthusiasts/Curious individuals

The interviews had two main goals and two associated activities/methods:

Goal 1. To explore what 'underrepresented histories' means to different people, possibilities for canned searches and associated terms for keywording.

> **Method:** Participants were invited to 'mind-map' using Miro or pen and paper in response to the question 'what does underrepresented histories mean to you?'. This is a method commonly used to understand users' thoughts, experiences, and preferences (Priestner, 2018) and enable access to the 'world' of the user (University of Glasgow, n.d.). Participants were encouraged to switch sticky note or line colour every two minutes to enable easier observations of what concepts/terms came to mind first and to verbally walk the interviewer through their thought process.

Goal 2. Understand how users might group together collections into broader categories that go beyond current collections categories/departments.

> **Method:** Participants were invited to sort images of the collections (a total of 16 – from 8 most popular collection categories in the Search our Collections queries sample) using Miro or physical printouts. After the initial sort, users were asked to describe their thought process and groupings, and then, asked if they could subdivide their groups further. This was chosen as a method to understand how people might organise the collections into different categories and then hierarchically organise these groupings.

5.4.2 Key findings

Concept Mapping

When mapping out 'underrepresented histories' the first thing that participants commonly thought of were different facets of identity such as:

- Ethnicity/race
- Socio-economic status
- Gender and LGBTQ communities
- Age: youth and kids

- People with disabilities
- Travellers
- Religion
- Non-western perspectives
- Post-human perspective (materials and environments)

This is demonstrated by the first lines/sticky notes in the maps by Kimber and Amelia shared below: in purple lines, and yellow sticky notes respectively (figure 5-16 & 5-17).[10]

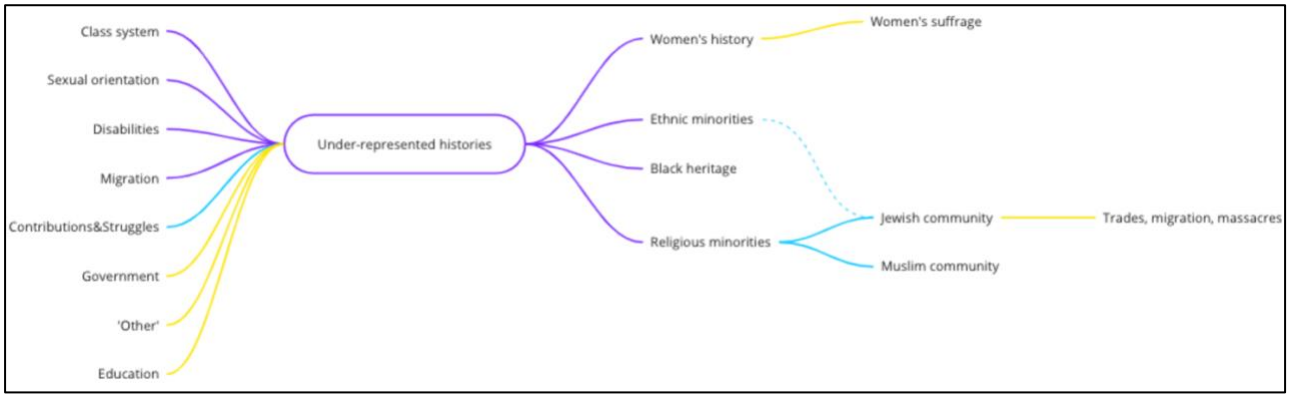


Figure 5-16 Kimber's concept map[11]

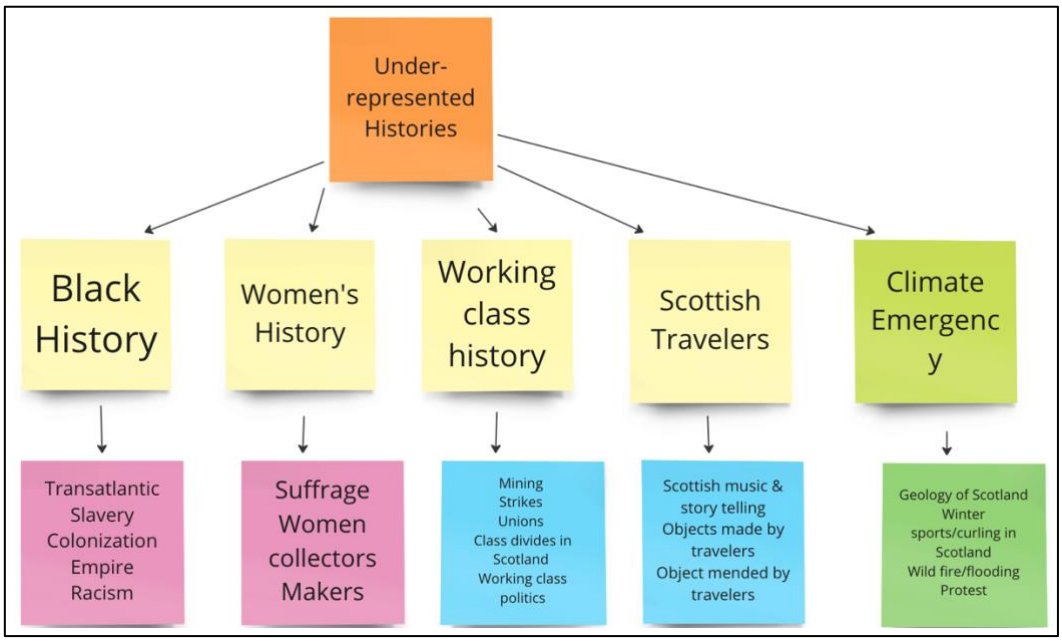


Figure 5-17 Amelia's concept map

Subsequently, some participants noted down the particular or specific experiences of people that may be underrepresented, such as:

[10] The different size or styles of text visible in the concept maps are not representative of level of importance but rather a result of the formatting of Miro tool used during the interviews for concept mapping.
[11] Interviewee's names have been replaced by a pseudonym either chosen by the researchers or the participant to ensure their anonymity.

- Migration
- Suffrage
- Experiences in clans
- Making, creating, collecting
- Contributions and struggles
- Wildfire and flooding
- Work, trades, domestic work
- Mental health
- Poverty, homelessness
- Racism
- Strikes
- Affective and embodied experiences

This is evident in the previous concept maps (e.g., of user interviewee Amelia above), but also in Isabel's map below (figure 5-18) through the participant's references to a 'pulse on the particular', 'experiences', and 'affective and embodied experiences.'

Some participants further explored the 'why' of underrepresentation:

- Western/heteronormative being seen as 'universal'
- Power dynamics as a result of colonialism, war, and segregation
- Lack of different knowledge creators in institutions
- Lack of diverse perspectives in historical accounts and/or the accounts are framed in a way that re-subjugates/is violent
- Dependent on material culture and oral/intangible cultural heritage - how that's been maintained, included, or excluded
- Barriers to visiting and participating with and in cultural heritage institutions
- Small objects that are rich with individual experiences may be hidden or viewed as 'plain'

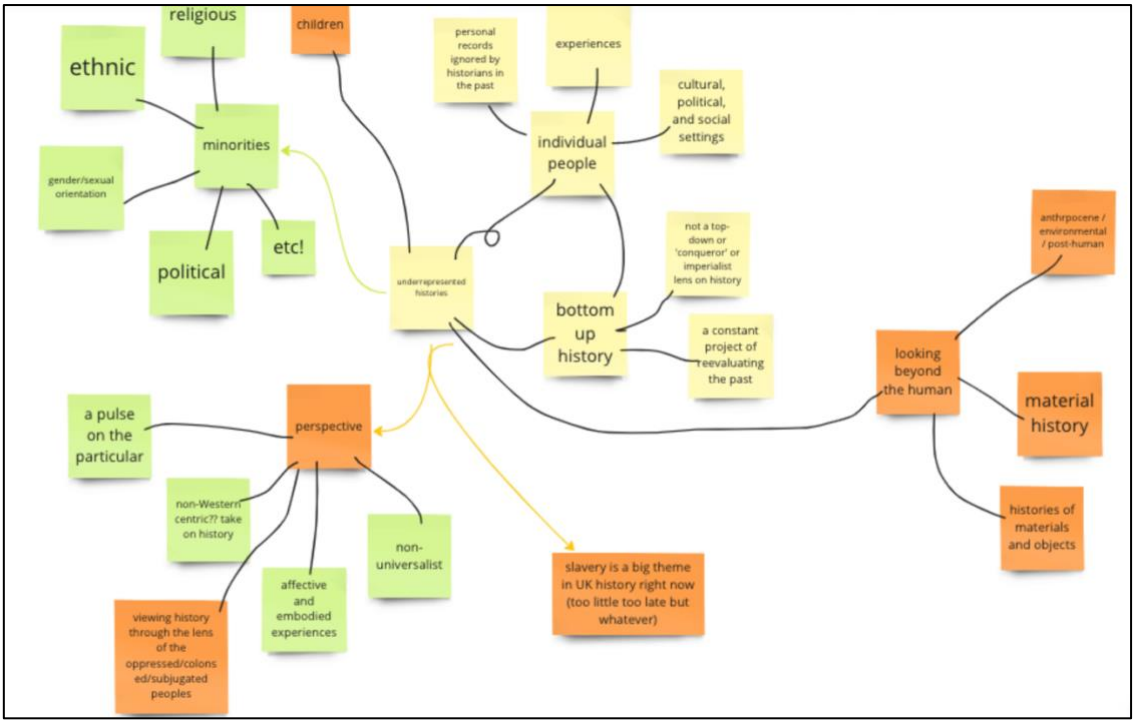


Figure 5-18 Isabel's concept map

Some participants explored the future:

- Diverse histories can connect to a wide array of epistemologies and ways of being, which can inform future policies and community development
- Ability and willingness to care for the environment as connected to place/understanding of place through cultural heritage
- Awareness of other peoples' experiences can garner understanding, challenge prejudices and racism
- Sharing multiple histories can enable people with different experiences to feel seen/like the space of the museum is for them
- Governments and education can help challenge underrepresentation
- Participation of different groups can help in the ongoing (never ending) re-evaluation of history and new possibilities/visions for the future

For instance, the final sticky note added to this map by user interviewee Jordan (figure 5-19), emphasises the impact of under-representation on future community building and policy making.

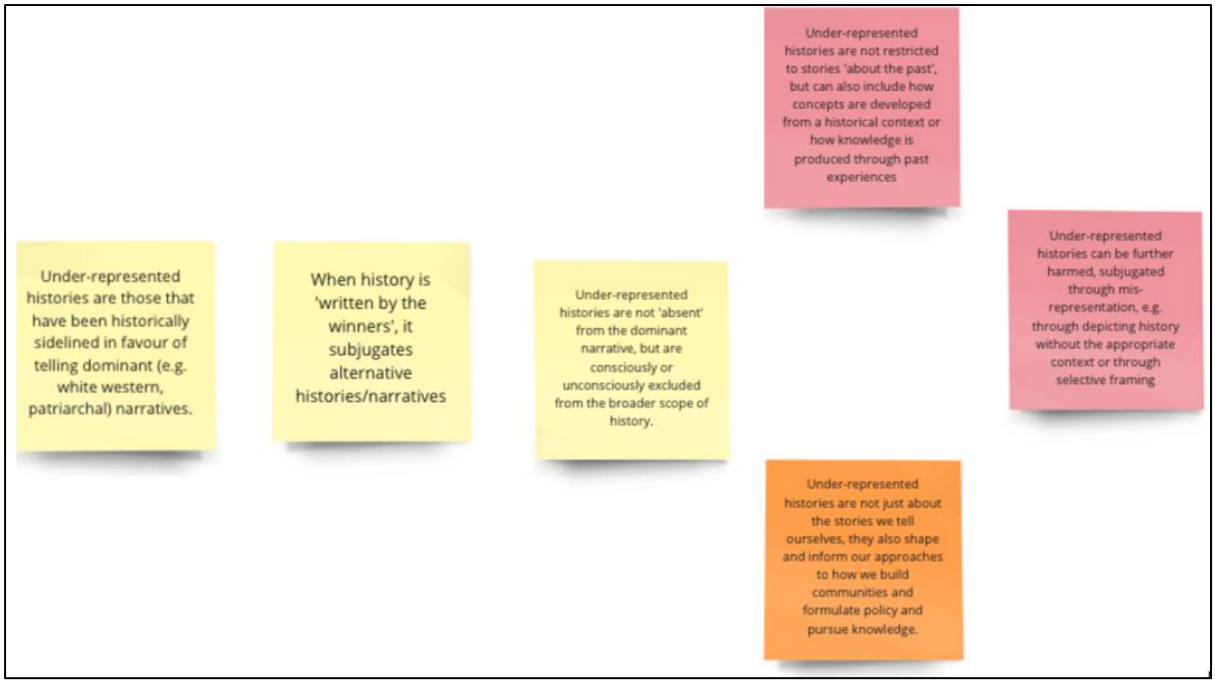


Figure 5-19 Jordan's concept map

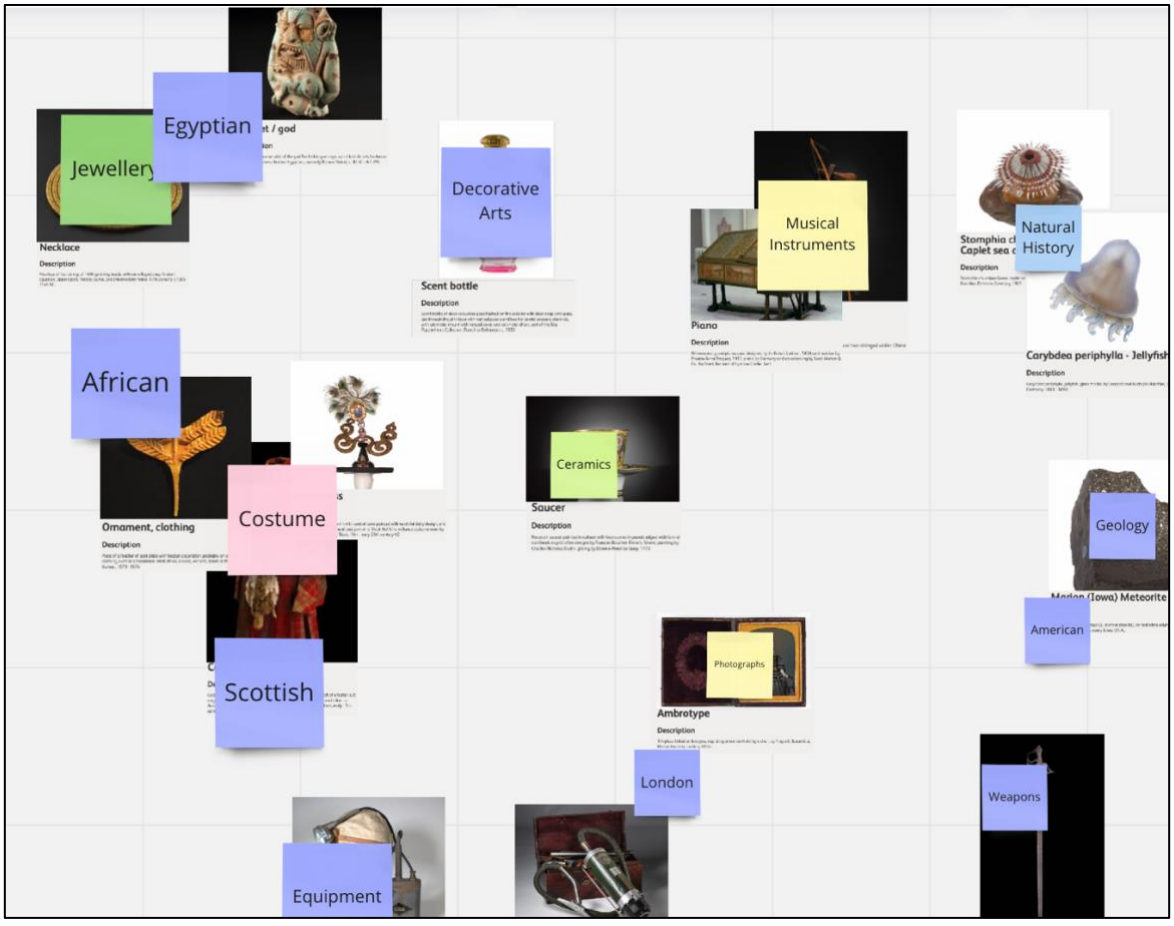


Figure 5-20 Angie's card sort activity

Card sorting activity

Sorting images of the collections revealed:

- Objects are layered or dynamic prisms depending on personal experiences, perspective, and communicative intention
- Museum professionals reflected that they found it hard to not group images by their own collections categories e.g., decorative arts or geology
- Some participants organised objects based on broader subjects comparable to those of the museum (fashion/textiles, geology, natural history, musical instruments). For instances, see image of Angie's card sort from our interview (figure 5-20).
- Several groupings suggest a clear divide between natural and human-made objects, and a common division between three groups of objects:
  1. 'functional objects' → e.g., tools for work (vacuum/miner's hat), objects for entertainment (instruments/teacup);
  2. 'decorative objects' → e.g., pieces for 'show' (scent bottle, jewellery);
  3. 'symbolic objects' → e.g., of death and religion (mortuary sword and amulet).

  One example of this, is the card sort pictured below from Isabel's interview whose categories evidence these divisions (figure 5-21).

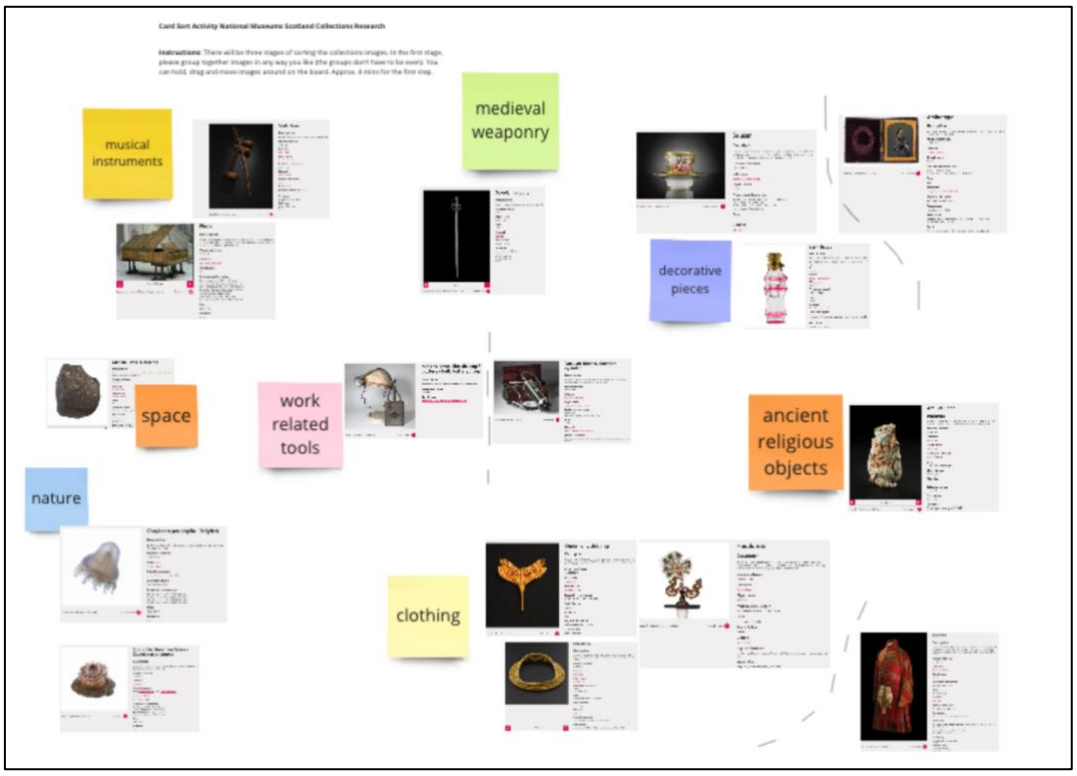


Figure 5-21 Isabel's card sort activity

- Others emphasised the conversation/narrative that happens by grouping objects together (e.g., a few participants commented that the meteorite and miners' hat could tell a story about extraction (figure 5-22), some objects were connected to women's histories (figure 5-22 & 5-23), everyday life (figure 5-24) and industrialization (figure 5-25)).

- Some participants were interested in provenance and periods, particularly when further subdividing groups

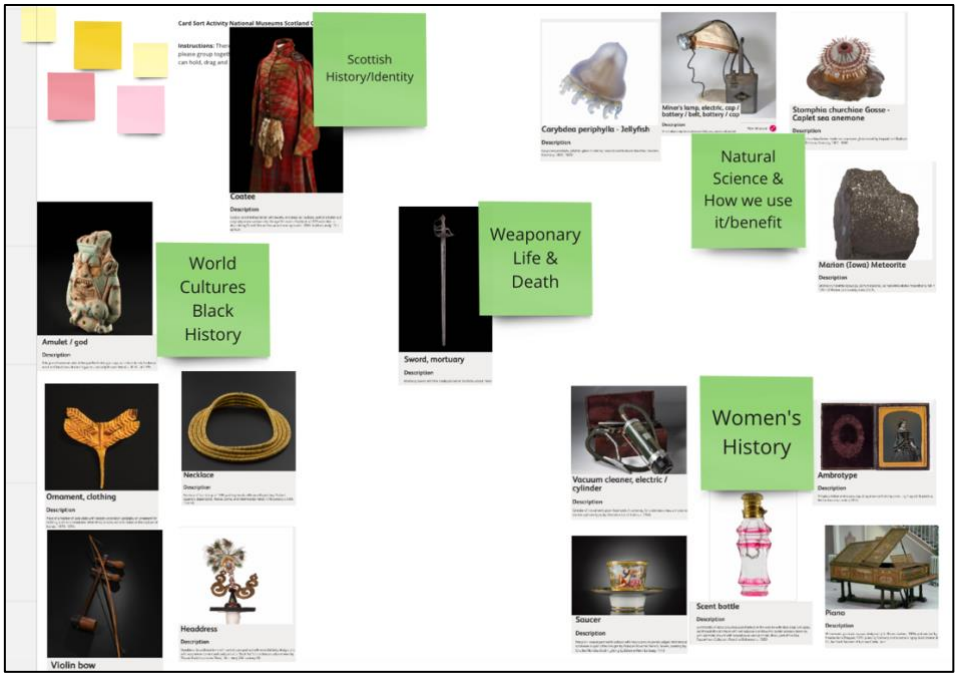


Figure 5-22 Amelia's card sort activity

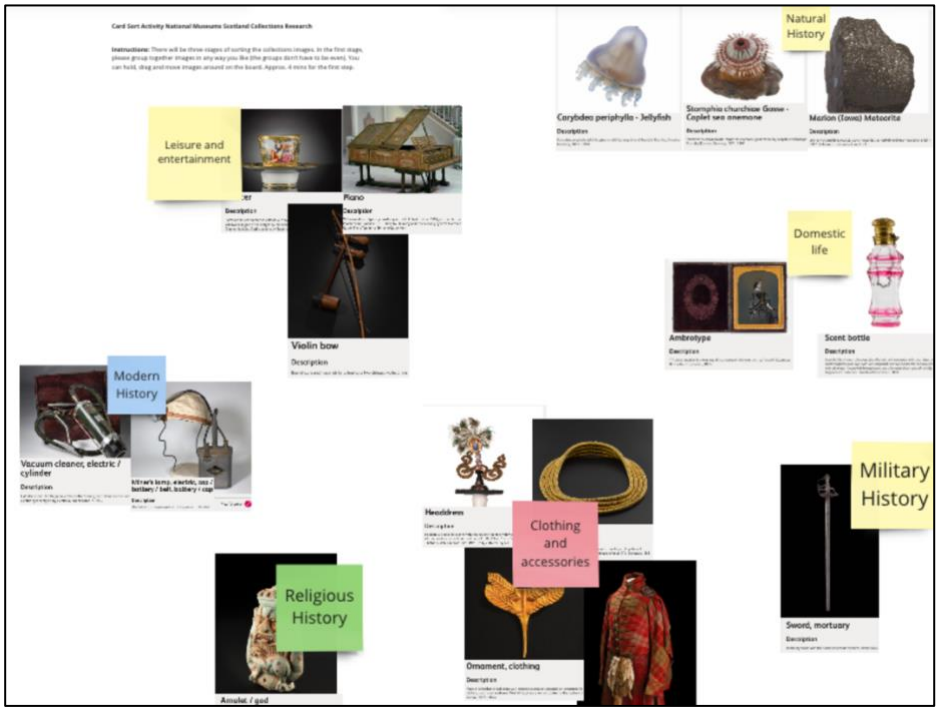


Figure 5-23 Kimber's card sort activity

Some narratives/conversations identified between objects included:

- Fab women/women's history
- King George
- Black history

- Work histories
- Extraction and mining
- Domestic life
- Everyday life
- Industrialization
- Technological progression

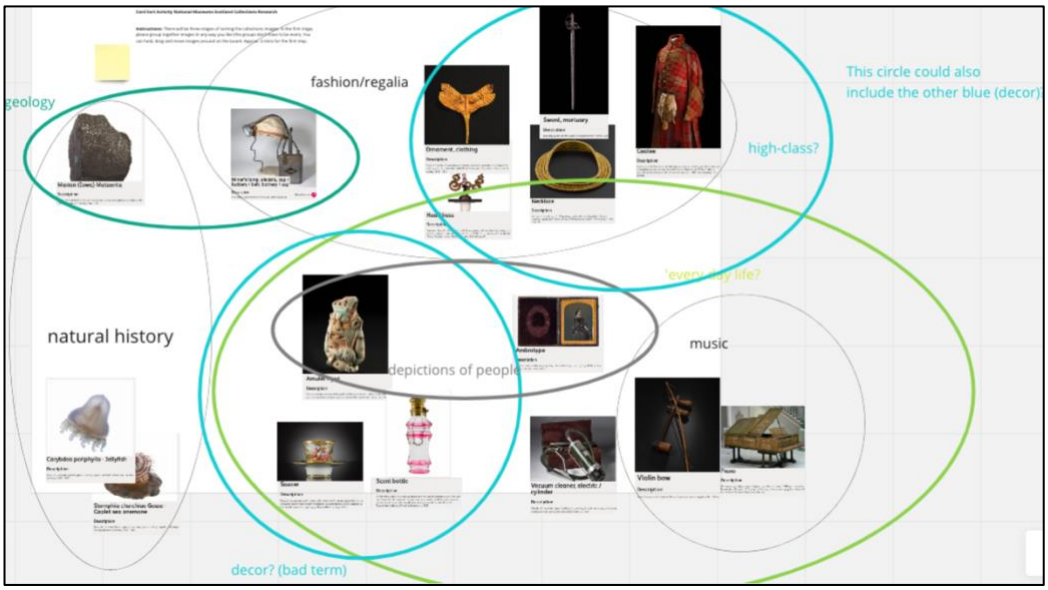


Figure 5-24 Jordan's card sort activity

Figure 5-25 Ballera Baloo's in-person card-sort activity

6 Prototyped frameworks

The following prototyped frameworks for applying keywords and captions to the collections' images are inspired by Klavans et al's (2014) matrix. However, the frameworks prioritise elements that the Museums' staff perceive to be missing from the metadata and the types of terms users commonly searched for in the Search our Collection's (SoC) queries and applied to images of the collections in our surveys and interviews. Based on these findings, the frameworks prioritise visual terms associated with the general 'who' or type of person, object, or thing visible in the image and then in decreasing importance other facets of subject description.

The prototyped frameworks for captions, due to the focus of both keywords and captions on the visuality of images, drew on the suggested best practices of other institutions such as the National Gallery (US) and the Cooper Hewitt (US) regarding alt-text.

The frameworks pictured in this section includes both a 'formulaic' and 'decision tree' framework for applying keywords and captions. These different styles have similar content for keywords (figure 6-1 & 6-2) and captioning (figure 6-3, 6-4, 6-5, & 6-6). Alternative versions of the framework were reviewed by all members of the UofG and Museums' project team (alternative version can be found in Appendix D). The four frameworks included here were further evaluated by staff from across the different Museums' departments at a workshop (24 May 2023) and subsequently tweaked.

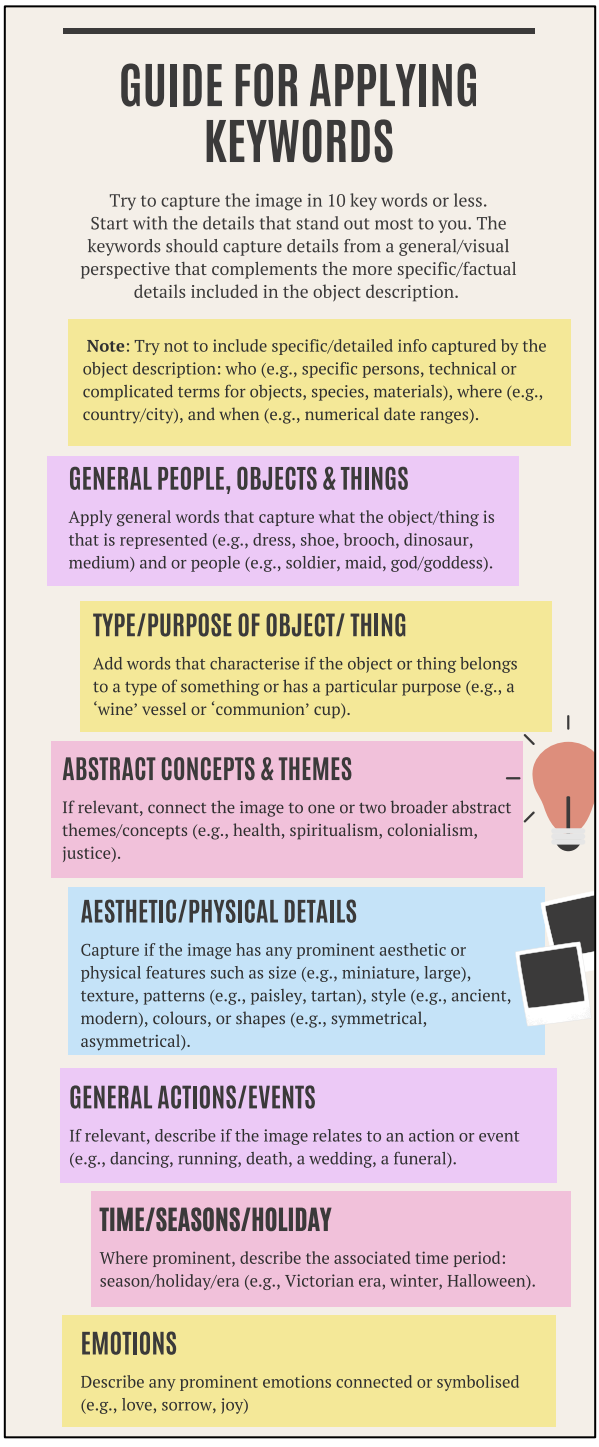


Figure 6-1 Formulaic framework for applying keywords

While there was agreement across workshop participants that another layer of metadata or changes to current collections' metadata are needed to improve the findability and searchability of the online collections, there was critical discussion on the relation between keywords, captions, and current object descriptions for serving this purpose. This mainly pertained to **how different** the object description, caption, and keywords need to or should

be in order to increase the collections' findability. Particularly, there was suggestions that the keywords and object descriptions might be sufficient, without creating too much duplication across the collections metadata that might occur with the added layer of an image caption. As such, an introductory sentence was added to the frameworks to clarify its relation to the object description. Some of the questions staff raised were to do with the entanglement between the search interface, the way it is and may be reshaped in the near future, and how this connects to the collections metadata.

Generally, it was agreed by workshop participants that applying keywords may be the best first step with potentially the greatest impact on the collections' findability for the required staff time and investment. Therefore, questions pertaining to the purpose and the value of the caption and its relation to accessibility, findability, and feasibility are left open for future institutional discussions. With some further user testing and consultation, rather than focusing on findability, the captions could be aimed at improving accessibility, adhering more clearly to the requirements of alt-text.

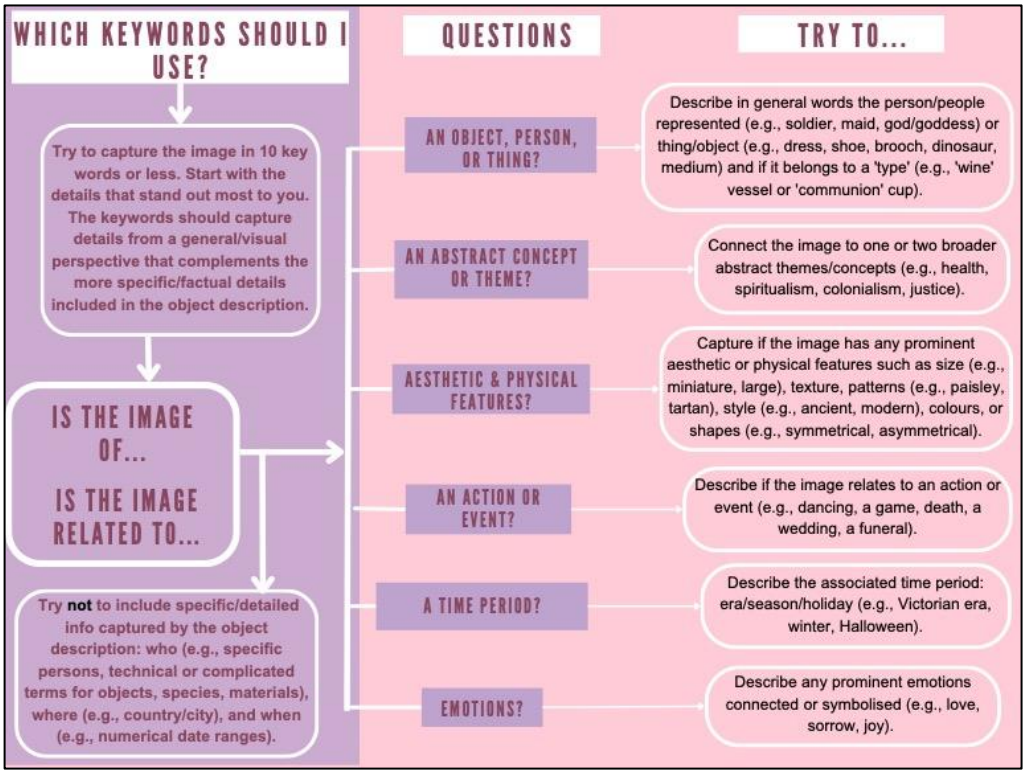


Figure 6-2 Decision tree framework for applying keywords

Staff suggested very few tweaks to the frameworks for keywording but referenced the need for a thesaurus that can capture synonyms and suggested adding a few examples to the frameworks that encompass object shape such as symmetry. There was also critical reflection that 'materials' could be excluded from keywording as it is typically included in object metadata and can be hard to guess from an image. As an additional aid, a compilation of potential keywords based on mainly the Search our Collections queries and partly the surveys and interviews are included in Appendix C. Workshop participants also suggested that additional checks for applying keywords would be needed, such as having the object description available for comparison/clarification, a buddy-system (two staff applying

keywords to cross-check) and/or for a specialist to review the keywords before publication on the web. Rather than concern for accuracy or having the 'right answer', there was a concern for having 'wrong answers' or applying inappropriate or insensitive terms.
The process of applying keywords, and who participates - whether staff, volunteers or the public is a pivotal question that requires further institutional discussion, experimentation, and evaluation.

Figure 6-3 Formulaic framework for captioning images pg. 1

Figure 6-4 Formulaic framework for captioning pg. 2

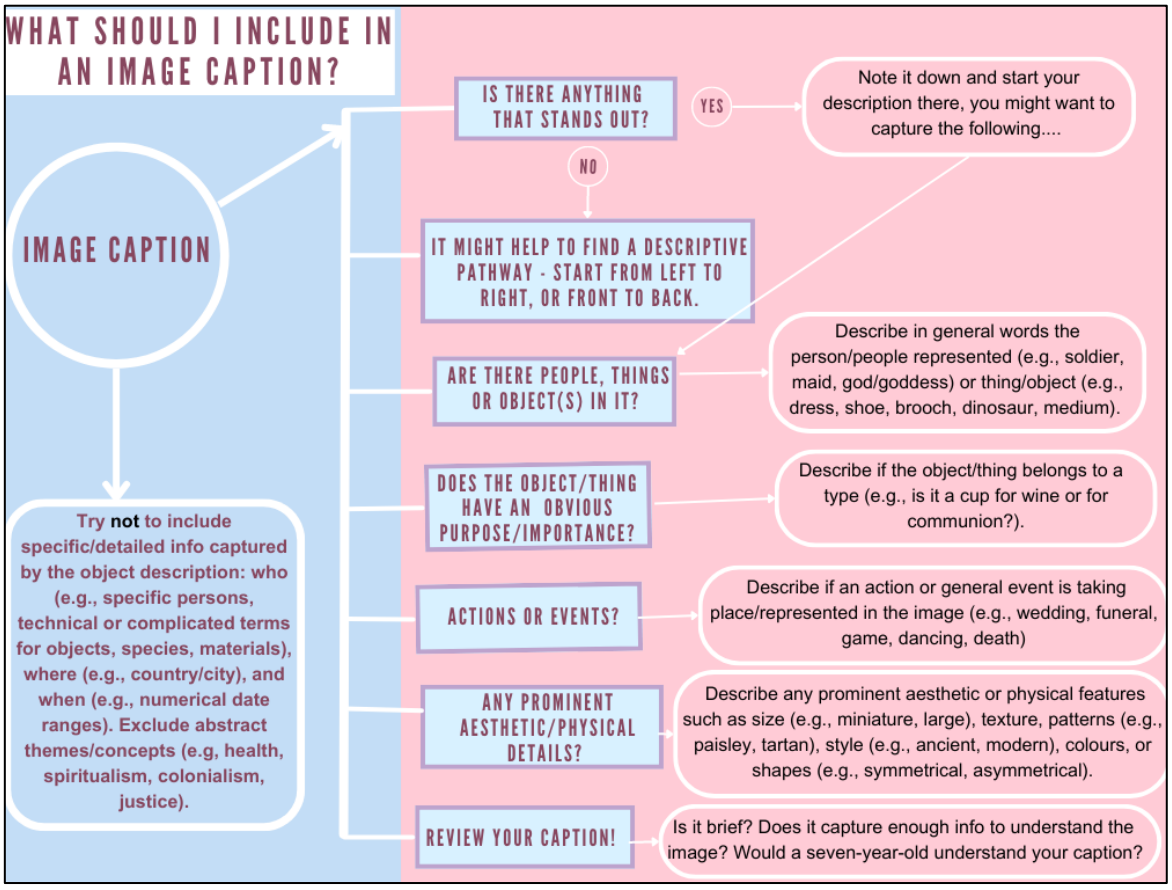


Figure 6-5 Decision tree framework for applying captions pg. 1

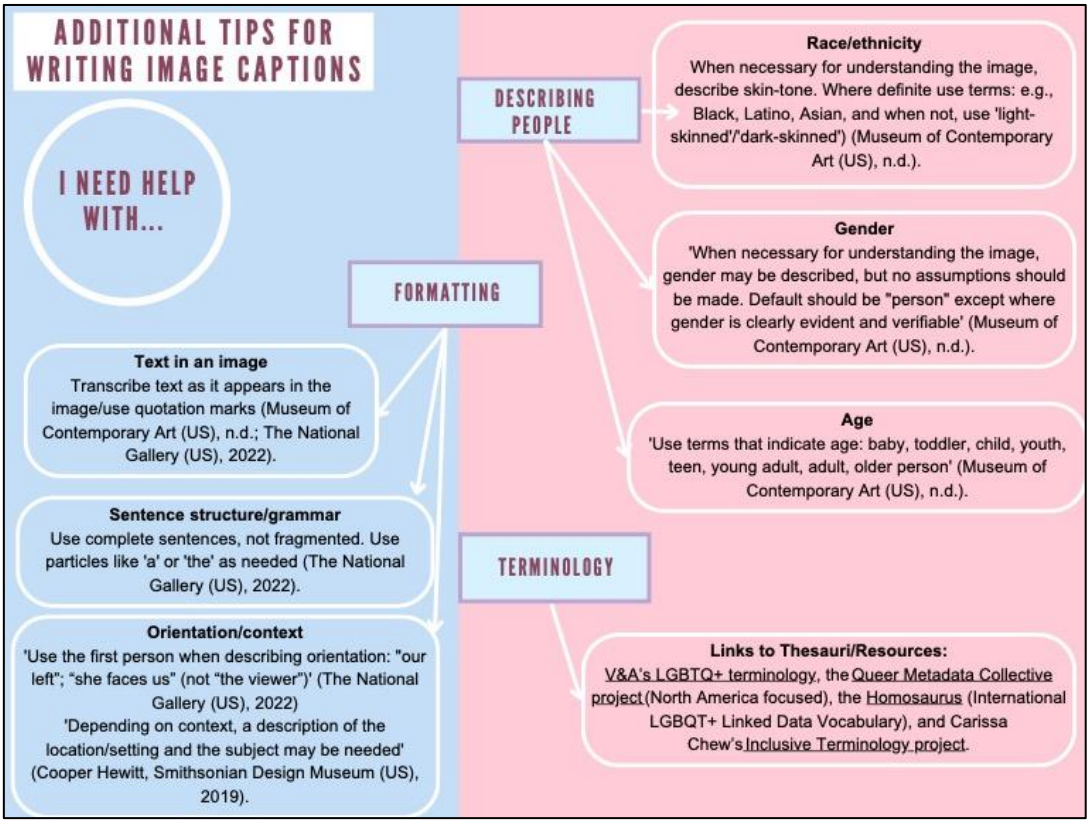


Figure 6-6 Decision tree framework for applying captions pg. 2

7 Implications for practice and recommendations

In this section we highlight recommendations directly from survey participants for the Search our Collections portal. Subsequently, based on the research findings, we make recommendation for the Museums' practices.

7.1 Recommendations from survey participants

- Make more images of the collections accessible online and enable/support image use and reuse
- Interlink collections images to other resources
- Enable users to search in different ways (location/date with clearer categories and dropdowns/suggestions)
- More detailed object descriptions for non-Western collections (work with community groups to develop these)

For instance, the following users advocated for:

'More images to the online collection. This is singularly the most important thing. It's still useful to know these items are in the collection but significantly more helpful if there are photos, including multiple angles.'

'More detailed search, keywords, choose words to search/clear categories, objects with images vs without, objects with more or higher quality descriptions vs without, sections.'

7.2 Our recommendations based on the research findings

7.2.1 Terms for keywords/captions

1. Keywords and captions should prioritise general terms for 'who/what' type of objects, figures, people, or things are visible in or associated with the collections' images (Section 5.2 & 5.3).
2. Secondary priorities may include applying terms related to actions (mourning, knitting) and abstract concepts visible in the image in the 'who/what' category – people, objects, things, ideals/values (Section 5.2, 5.3 & 5.4).
3. Additionally, applying visual terms (colour, shape, texture) could enable differing entry points into the collections (Section 5.2 & 5.3).
4. As clothing appears to have the most user interest from the SoC queries, related objects could be a good starting point for applying keywords and captions, albeit this popularity could be due to staff using the SoC portal to prepare for the upcoming 'Beyond the Little Black Dress' exhibition (Section 5.2.2).

5. When implementing keywording – create a 'check' system, e.g., one that includes two individuals keywording an image or a review process by a specialist (Section 6).

6. Incorporate sustainable channels or methods for gathering Museums staff feedback, checking keywords, and providing insight for continuous tweaks to the Search our Collections portal (Section 5.1 & 6).

7. Teachers/care workers were not well represented in the survey, but it could be useful to organise/tag collections by school related themes/more abstract concepts based on interviews with the Museums' Learning Enablers.

7.2.2 System affordances

8. Design a more 'helpful' search interface that offers suggestions or drop-downs that can evidence different ways users can search or explore the collections (Section 5.1 & 5.3).

9. Ensure high-quality images are accessible and that visitors are aware of how they can be used, and that these images are linked with other material and media both internally and externally to enable user browsing (Section 5.1 & 5.3).

7.2.3 Underrepresented histories/themes

10. Involve groups to develop virtual trails, maps, or canned searches connected to underrepresented histories to improve user experiences of online collections while diversifying knowledge production and access to the online collections (Section 5.4).

11. Some of the 'curious individuals' interviewed were enthusiastic about women's histories/mental health, climate change and rewilding, and Indigenous collections – and have connections through which the museum could seek further participation/participants.

Implementing some of the suggestions above could be a powerful step in moving the online collections away from silos and divisions grounded in the museums' collections departments and towards challenging the long-lasting legacies of the collections organisation based on the catalogue card system. In turn, a variety of keywords and themes, could enable novel ways for users to explore, search, and engage with the online collections. In particular, exploring themes connected to underrepresented histories and in participatory ways that diversify knowledge production, could be a promising method of contributing to institutional concerns and goals regarding the decolonisation of collections practices. The ability of users to explore, search, and browse the online collections is intertwined with not only the enhancement of image metadata but as evidenced through the findings, the affordances of the search portal interface. Specifically, the Search our Collections portal must be helpful and able to clarify how users can search or browse, which are factors that should be considered during website re-development.

While these suggestions are aimed at enabling access for users, attracting different user groups will be a core challenge. This challenge could be addressed by reaching out to different groups that are less evident in the research data (grandparents, teachers, creatives), commissioning art based on the collections, and facilitating participatory projects that increase both awareness and engagement across networks and platforms with the online collections. The institution should invest labour and time on particular collections – in this case, clothing appeared to be the most popular category of objects search and could be a good starting point for investing resources into further images and media, and for trialling keywords and captions.

References

Armitage, L. H., & Enser, P. G. B. (1997). Analysis of user need in image archives. *Journal of Information Science*, *23*(4), 287–299. https://doi.org/10.1177/016555159702300403

Bailey-Ross, C. (2021). *Online User Research Literature Review: UK Gallery, Library, Archive and Museum (GLAM) Digital Collection*. Zenodo. https://doi.org/10.5281/zenodo.5779826

Benedetti, B. (2002). Virtuality and Reality in Enterprise's Museum, Art Museum, Archeological Sites. *Quaderni Della Fondazione Piaggio, Nuova Serie: Cultura Europea e Musei*, *6*(1).

Brooke, G. (2022, June 23). Cultural Content: Explore the Collections [Substack newsletter]. *Cultural Content*. https://culturalcontent.substack.com/p/cultural-content-explore-the-collections

Choi, Y., & Rasmussen, E. M. (2003). Searching for images: The analysis of users' queries for image retrieval in American history. *Journal of the American Society for Information Science and Technology*, *54*(6), 498–511. https://doi.org/10.1002/asi.10237

Chowdhury, G. G., & Chowdhury, S. (2011). *Information users and usability in the digital age*. Facet Publishing. https://doi.org/10.29085/9781856049757

Chowdhury, G., Gibson, R., Chowdhury, S., Ross, J., St Clair Inglis, C., Cawston, R., & Ganley, C. (2022). *Digital footprints and search pathways: Working with National Collections in Scotland during Covid-19 lockdown to design future*. Zenodo. https://doi.org/10.5281/zenodo.6602365

Ciecko, B. (2020). AI Sees What? The Good, the Bad, and the Ugly of Machine Vision for Museum Collections. *MW20:MW2020*. Museums and the Web 2020. https://mw20.museweb.net/paper/ai-sees-what-the-good-the-bad-and-the-ugly-of-machine-vision-for-museum-collections/index.html

Clough, P., Hill, T., Paramita, M. L., & Goodale, P. (2017). Europeana: What Users Search for and Why. In *Research and Advanced Technology for Digital Libraries* (Vol. 1–Book, Section, pp. 207–219). Springer International Publishing. https://doi.org/10.1007/978-3-319-67008-9_17

Collections Trust. (2022a). *Terminology guidance*. Collections Trust. https://collectionstrust.org.uk/terminology/terminology-guidance/

Collections Trust. (2022b). *What's new in Spectrum 5.1? – Collections Trust*. Collections Trust. https://collectionstrust.org.uk/spectrum/spectrum-5/summary-of-changes/

Dahlgren, A., & Hansson, K. (2020). The Diversity Paradox: Conflicting Demands on Metadata Production in Cultural Heritage Collections. *Digital Culture & Society*, *6*(2), 239–256. https://doi.org/10.14361/dcs-2020-0212

Dahlgren, A. N. (2022). Image metadata. From information management to interpretative practice. *Museum Management and Curatorship*, 1–21. https://doi.org/10.1080/09647775.2022.2073562

Dobreva, M., O'Dwyer, A., & Konstantelos, L. (2011). User needs in digitization. In L. Hughes (Ed.), *Evaluating and Measuring the Value, Use and Impact of Digital Collections* (pp. 73–84). Facet. https://doi.org/10.29085/9781856049085.007

Falk, J. H., & Dierking, L. D. (2016). *The museum experience revisited*. Routledge. https://doi.org/10.4324/9781315417851

Fidel, R. (1997). The image retrieval task: Implications for the design and evaluation of image databases. *New Review of Hypermedia and Multimedia*, *3*(1), 181–199. https://doi.org/10.1080/13614569708914689

Giaccardi, E. (2006). Collective Storytelling and Social Creativity in the Virtual Museum: A Case Study. *Design Issues*, *22*(3), 29–41. https://doi.org/10.1162/desi.2006.22.3.29

Hale, M. (2019). *Searching for Art Records: A Log Analysis of the Ackland Art Museum's Collection Search System* [University of Carolina]. https://doi.org/10.17615/hh31-4g44

Hall, M. M. (2018). Digital Museum Map. In E. Méndez, F. Crestani, C. Ribeiro, G. David, & J. C. Lopes (Eds.), *Digital Libraries for Open Knowledge* (pp. 304–307). Springer International Publishing.

Hall, M. M., & Walsh, D. (2021). Exploring digital cultural heritage through browsing. In *Information and Knowledge Organisation in Digital Humanities*. Routledge.

Han, H., & Wolfram, D. (2016). An exploration of search session patterns in an image-based digital library. *Journal of Information Science*, *42*(4), 477–491. https://doi.org/10.1177/0165551515598952

Hollink, L., Schreiber, A. T., Wielinga, B. J., & Worring, M. (2004). Classification of user image descriptions. *International Journal of Human-Computer Studies*, *61*(5), 601–626. https://doi.org/10.1016/j.ijhcs.2004.03.002

Hopes, D. (2014). Digital CoPs and Robbers: Communities of practice and the use of digital artefacts. *Museum Management and Curatorship*, *29*(5), 498–518. https://doi.org/10.1080/09647775.2014.959703

Hourihane, C. (2002). It Begins with the Cataloguer: Subject Access to Images and the Cataloguer's Perspective. In M. Baca (Ed.), *Introduction to Art Image Access: Issues, Tools, Standards, strategies* (pp. 40–59). https://www.getty.edu/publications/resources/virtuallibrary/0892366664.pdf

Jörgensen, C. (1998). Attributes of images in describing tasks. *Information Processing & Management*, *34*(2), 161–174. https://doi.org/10.1016/S0306-4573(97)00077-0

Klavans, J. L., LaPlante, R., & Golbeck, J. (2014). Subject matter categorization of tags applied to digital images from art museums. *Journal of the Association for Information Science and Technology*, *65*(1), 3–12. https://doi.org/10.1002/asi.22950

Martin, J. M. (2021). Records, Responsibility, and Power: An Overview of Cataloging Ethics. *Cataloging & Classification Quarterly*, *59*(2–3), 281–304. https://doi.org/10.1080/01639374.2020.1871458

Mihelj, S., Leguina, A., & Downey, J. (2019). Culture is digital: Cultural participation, diversity and the digital divide. *New Media & Society*, *21*(7), 1465–1485. https://doi.org/10.1177/1461444818822816

Museums Association. (2023, April 17). *Public engagement & public benefit*. Museums Association. https://www.museumsassociation.org/campaigns/ethics/code-of-ethics/1-public-engagement-public-benefit/

Navarrete, T., & Mackenzie Owen, J. (2016). The Museum as Information Space: Metadata and Documentation. In K. J. Borowiecki, N. Forbes, & A. Fresa (Eds.), *Cultural Heritage in a Changing World* (pp. 111–123). Springer International Publishing. https://doi.org/10.1007/978-3-319-29544-2_7

Newbury, D., & Brennan, D. (2021, November 9). *Digital Collections as Product*. Museum Computer Network. https://mcn.edu/resources/digital-collections-as-product/

One Further. (2022, October 20). *The Cultural Content Report 2022*. One Further. https://onefurther.com/blog/cultural-content-survey-2022

Panofsky, E. (1972). *Studies in iconology: Humanistic themes in the art of the Renaissance*. Westview Press.

Priestner, A. (2018, September 13). *UX In libraries: Cognitive Maps*. CILIP: The Library and Information Association. https://www.cilip.org.uk/page/CognitiveMapsMember

Shatford, S. (1986). Analyzing the Subject of a Picture: A Theoretical Approach. *Cataloging & Classification Quarterly*, *6*(3), 39–62. https://doi.org/10.1300/J104v06n03_04

Shatford, S. (2002). Subject Access to Art Images Sara Shatford Layne. In M. Baca (Ed.), *Introduction to Art Image Access: Issues, Tools, Standards, Strategies* (pp. 1–20). https://www.getty.edu/publications/resources/virtuallibrary/0892366664.pdf

Skov, M., & Ingwersen, P. (2014). Museum Web search behavior of special interest visitors. *Library & Information Science Research*, *36*(2), 91–98. https://doi.org/10.1016/j.lisr.2013.11.004

Slawsky, D. (2007). Building a keyword library for description of visual assets: Thesaurus basics. *Journal of Digital Asset Management*, *3*(3), 130–138. https://doi.org/10.1057/palgrave.dam.3650074

The Audience Agency. (2022, October 1). *Cultural Participation Monitor Autumn 2022*. The Audience Agency. https://www.theaudienceagency.org/evidence/covid-19-cultural-participation-monitor/recent-key-insights/autumn-2022-

Turner, H. (2016). Critical Histories of Museum Catalogues [Special Issue]. *Museum Anthropology*, *39*(2), 99–191. https://doi.org/10.1111/muan.12118

Turner, H. (2020). *Cataloguing culture: Legacies of colonialism in museum documentation*. UBC Press. https://go.exlibris.link/D1vKQPhr

University of Glasgow. (n.d.). *UX Framework—Cognitive mapping*. University of Glasgow - MyGlasgow. Retrieved 17 May 2023, from https://www.gla.ac.uk/myglasgow/ux/cognitivemapping/

Vaughan, C. (2018). The Language of Cataloguing: Deconstructing and Decolonizing Systems of Organization in Libraries. *Dalhousie Journal of Interdisciplinary Management*, *14*(0), Article 0. https://ojs.library.dal.ca/djim/article/view/7853

Villaespesa, E. (2019). Museum Collections and Online Users: Development of a Segmentation Model for the Metropolitan Museum of Art. *Visitor Studies*, *22*(2), 233–252. https://doi.org/10.1080/10645578.2019.1668679

Villaespesa, E., & Crider, S. (2021). A critical comparison analysis between human and machine-generated tags for the Metropolitan Museum of Art's collection. *Journal of Documentation*, *77*(4), 946–964. https://doi.org/10.1108/JD-04-2020-0060

Villaespesa, E., & French, A. (2019). *AI, Visitor Experience, and Museum Operations: A Closer Look at the Possible* (pp. 101–113).

Walsh, D., Hall, M. M., Clough, P., & Foster, J. (2020). Characterising online museum users: A study of the National Museums Liverpool museum website. *International Journal on Digital Libraries*, *21*(1), 75–87. https://doi.org/10.1007/s00799-018-0248-8

Warwick, C., Terras, M., Huntington, P., & Pappa, N. (2008). If You Build It Will They Come? The LAIRAH Study: Quantifying the Use of Online Resources in the Arts and Humanities through Statistical Analysis of User Log Data. *Literary and Linguistic Computing*, *23*(1), 85–102. https://doi.org/10.1093/llc/fqm045

Appendices

A. Appendix A: a copy of main survey questions

[SURVEY INTRO]

Have your say!

Help us improve how you can search National Museums Scotland's collections.

We are inviting you to participate in a research project about surfacing the national collections through Images. This study is being carried out by the University of Glasgow and National Museums Scotland. This project investigates who is using and not using National Museums Scotland's (the Museums') Collections webpages, and how we can describe and apply keywords to images of the collections in ways that improve access. You can find more information about the use of data and this study in the Participant Information sheet.

So, you want to take part?
You are welcome to participate:

- Whether you have used the Museums' Collections webpages or not
- If you are 18 and over
- Can read and write English

Your participation is entirely **voluntary**, your answers can be **anonymous**, and you can **skip** any question. You will be asked about your: demographic profile (e.g., age, gender); how you explore the web; to take part in an activity (describe an image through keywords); and whether you would be interested in being involved in a follow-up interview or focus group.

What happens to your data?
If you provide an email (optional) you can withdraw at any time up until the data is anonymised by the researchers (**June 1st, 2023**). All data is kept confidential and in secure storage, and anonymised data will be used in future public outputs without reference back to you. For further information about this project please see the information sheet, and/or contact the researcher Dr Cassandra Kist: Cassandra.kist@glasgow.ac.uk.

If you have any concerns regarding the conduct of the research project, you can contact the College of Arts Ethics Officer (email: arts-ethics@glasgow.ac.uk).

By submitting a response, you consent to take part in the survey and agree to the terms for data processing as outlined in the Participant information sheet.

[START]

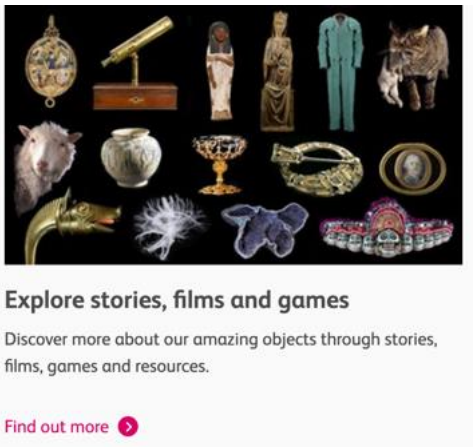
Survey to be shared over social media/with students/community groups.

1. Have you used National Museums Scotland's (the Museums) Collections webpages?

Yes
No (skip to Survey 2 question 2)

1b) I have used the following Collections webpages (select all that apply) (Proceed to Survey 1 question 2)

A) Search the Collections Database [Add screenshot- will be larger]



B) Explore Stories, Films, and Games [Add screenshot – will be larger]

Survey one

Section 1: [Understanding] your motivations

2. Which best describes the main purpose for your last visit to Museums' online collection? (please select **only one**)

- I am planning my visit to National Museums Scotland
- I followed a link on a site/social media/email to find out more about what was posted

- I am visiting for work (please specify)
- I am doing homework/coursework for school/college/university
- General interest (e.g. about art, science, technology, natural history, etc)
- I am looking for inspiration for my own creative work
- I want to share images on social media or post it on my blog/website
- Other (please specify)

3. What information or content were you looking for? (e.g., images, information on a type of object, specific collections, type of activity).

(Open-ended)

Section 2: [Understanding your] familiarity with cultural heritage and/or museum collections

4. I last visited the Museums' Collections webpages as a/an… (select all that apply):

Curious individual
Artist/producer/creative
Cultural heritage/art enthusiast
Nature/natural history enthusiast
Science/technology enthusiast
Academic researcher/scholar/scientist
Student
Museum/heritage professional
Personal/family history researcher
Teacher
Parent/grandparent
Tourist/travel enthusiast
Other:_____

5. How would you describe your knowledge of cultural heritage and/or museum collections?

Little or no knowledge
General
Specialist

Section 3: [Understanding] how you search or browse

6. When visiting the Museums' Collections webpages on your last visit, did you do any of the following (select all that apply):

Browsed collection stories, films, and/or games
Searched the collections for a specific object or object type - Terms searched: ___
Searched/browsed the collections by a general topic/subject - Terms searched: __
Browsed Featured Collections categories (e.g., Bagpipes, Ancient Egypt)

Searched the collections through Advanced Search specifying material, object name, or collections. Please describe: _____
Other: _____

7. I am most interested in exploring images of the Museums' collections from the following themes (select all that apply):

Life events (e.g., marriage, death, birth, illness)
Actions represented (e.g., giving, fighting, playing)
Actions abstracted or symbolised (e.g., courage, justice, generosity)
Figures or things represented (e.g., person, child, dinosaur, cat)
Mythical figures represented (e.g., unicorn, goddess, fairies)
General location represented (e.g., garden, gallery, jungle)
Location or place symbolised (e.g., paradise, Mount Olympus, hell)
Emotions associated or represented (e.g., love, anger, sadness)
Type of photographs (e.g., angle, style, portrait, quality)
Shapes, colours, and forms visible in the images
Under-represented histories
Politics and society (past/contemporary social movements)
Seasonal holidays/celebrations
Other topics/themes: _____

Section 4: Labelling activity

8. Help us make the following images searchable. Please attribute the keywords you may expect to search the following images by.

Choose some key words that describe what the image is of and about:

(open-ended)

Choose some key words that describe what the image is of and about:

(open-ended)

Choose some key words that describe what the image is of and about:

(open-ended)

Choose some key words that describe what the image is of and about:

(open-ended)

9. Do you have any further suggestions for improving your experiences of using the Museums' Collections webpages? (optional)

(Open-ended)

Section 5: Demographics

To understand the use of the Museums' Collections webpages and reach of this survey we are asking optional demographic questions.

10. Are you aged?
    18–24
    25–34
    35–44
    45–54
    55–64
    65–74
    75 and over
    Prefer not to say

11. I identify my gender as a… [Note: these options are based on the best practices of the UK National Office for Statistics]:
    Woman
    Man
    Non-binary/genderqueer/agender/gender fluid
    Prefer not to say
    Other

12. Where are you currently based?
    Scotland Urban
    Scotland Rural
    The rest of the UK
    European Union
    Rest of the world
    Prefer not to say

13. Please check your ethnic group/identity [Note: these options are based on the Scottish Government's best practice for census data]:

A. White [Drop down option list when clicking A,B,C..]
    Scottish
    Other British
    Irish
    Polish
    Gypsy/Traveller
    Roma
    Showman/Showwoman
    Other, please describe

B. Mixed or Multiple ethnic groups
    Any Mixed or Multiple ethnic groups, please describe

C. Asian, Asian Scottish or Asian British
    Pakistani, Pakistani Scottish or Pakistani British
    Indian, Indian Scottish or Indian British

Bangladeshi, Bangladeshi Scottish or Bangladeshi British
Chinese, Chinese Scottish or Chinese British
Other, please describe

D. African

African, African Scottish or African British
Other, please describe

E. Caribbean or Black

Caribbean, Caribbean Scottish or Caribbean British
Black, Black Scottish or Black British
Other, please describe

F. Other ethnic group

Arab, Arab Scottish or Arab British
Other, please describe

Prefer not to say

Want to help us further? We are looking for individuals to participate in an interview or focus group either online or in-person over the weeks: April 17th-29th for around 45 mins. We can cater to your availability between 9am-7pm BST. This will involve exploring topics relevant to you and keywording a wider breadth of collection images.

☐ I would be interested in taking part in an interview/focus group about how the National Museums Scotland describes images and can be contacted with further information.

Email: _____

All emails will be held in secure storage until July 2023 and then deleted. Please see the information sheet for further details about data use and storage.

[CLOSING]

Thank you for participating in this survey! If you have any further questions, concerns or would like to be updated about the project please contact Dr Cassandra Kist at Cassandra.kist@glasgow.ac.uk.

[END]

B. Appendix B: Images of the Museums' collections used in card sort activity

Coatee

Description

Coatee jacket in Ross tartan with twenty-one brass ball buttons, part of a tartan suit originally made and worn for George IV's visit to Scotland in 1822 and taken to Australia by Donald Munro Ross when he emigrated in 1864, Scotland, early 19th century

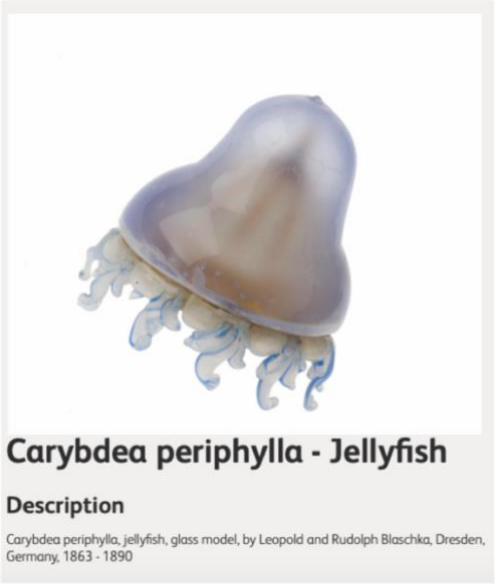
Figure B-1 Coatee image

Ambrotype

Description

1/4 plate tinted ambrotype, depicting a woman holding a chair, by Negretti & Zambra, Hatton Garden, London, 1850s

Figure B-2 Ambrotype image

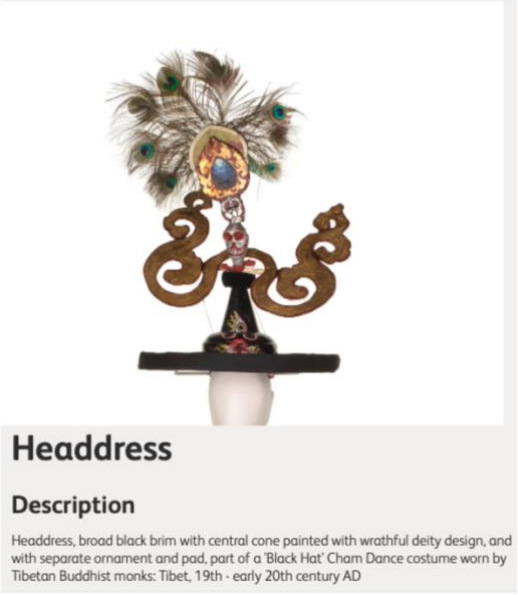


Carybdea periphylla - Jellyfish

Description

Carybdea periphylla, jellyfish, glass model, by Leopold and Rudolph Blaschka, Dresden, Germany, 1863 - 1890

Figure B-3 Jellyfish image



Headdress

Description

Headdress, broad black brim with central cone painted with wrathful deity design, and with separate ornament and pad, part of a 'Black Hat' Cham Dance costume worn by Tibetan Buddhist monks: Tibet, 19th - early 20th century AD

Figure B-4 Headdress image

Figure B-6 Sea anemone image

Figure B-5 Vacuum image

Figure B-7 Miner's lamp image

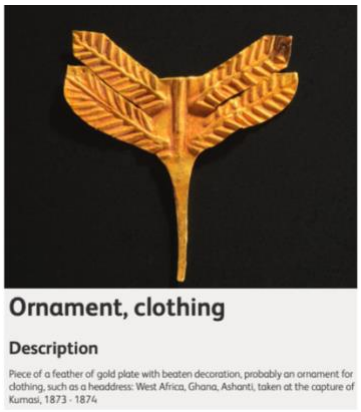


Figure B-8 Ornament image

Figure B-10 Piano image

Figure B-9 Amulet image

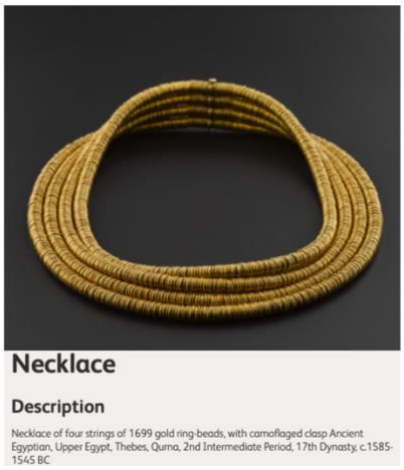


Figure B-11 Necklace image

Figure B-12 Saucer image

Figure B-14 Meteorite image

Figure B-13 Mortuary sword image

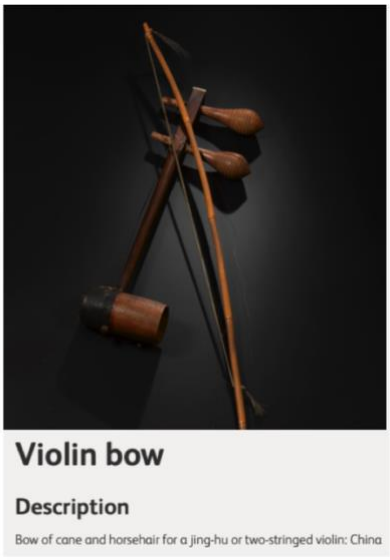


Figure B-15 Violin bow image

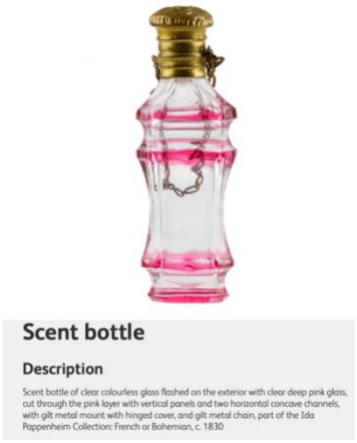


Figure B-16 Scent bottle image

C. Appendix C: popular ‘general’ and ‘abstract’ terms from fieldwork

- Most of the following terms are from the Search our Collections user queries sample from 2020-2021
- Highlighted terms in yellow are additional suggestions based on Klavans et al.’s (2014) matrix for image description (section 4.3) and suggested thesauri (section 4.4)
- Terms highlighted in green are based on user keywords from the surveys and from the mind-map and card sort interview activities (section 5.4)

Table C-1 Terms for general objects/things

Clothing	Accessories	Dishware/ food and drink	Instruments
Dress	Brooch	Bowl	instrument
Armour	Pendant	Plate	Bagpipe
Fashion plate	Necklace	Teapot	Highland bagpipe
Jacket	Earring'	Dish	Harp
Skirt	Ring	Goblet	Bell
Coat	Finger ring	Jug	Small pipes, bagpipe
Shirt	Ornament	Spoon	Rattle
Bodice	Mask	Cup	Bellows bagpipe
Glove	Hair ornament	Chalice	Trumpet
Coat, man's	Bead	Pottery/sherd	Lute
Trousers	Bag	Saucer	Drum
Mitten	Hat	Tea Bowl	Union bagpipe
Belt	Purse	Bottle	Lowlands bagpipe
Cloak	Bracelet	Teapot cover	Violin
Shoe	Chain	Pot	Zither
Robe	Armlet	Tea set	Violin bow
Bustle	Locket	Sugar bowl	Whistle
Apron	Pouch	Coffee pot	Guitar
Thong	Bangle	Form	Loom backstrap
Hood	Pin	Glass	Talking drum
Collar	Neckpiece	Tea caddy	Horn
Dress, woman's	Fibula	Cup stand	Clapper
Boot	Cap	Plate, soup	Drumstick
bonnet	Arm ring	Glass, wine	Touting-horn
sandal	Ear ornament	Vessel stone/ cup	Guitar
Blouse	Armband	Bowl, Toddy	Loom backstrap
kilt	Necklet	Tankard	Talking drum
Petticoat	String of beads	Cup / cover cup	Horn
Helmet	Nose ornament	Decanter	Clapper

Stocking	Badge	Teacup	Drumstick
Legging	Anklet	Pottery	Touting-horn
Button	Watch	Tumbler	Bagpipe
Jersey	Wig	Pan	harp
Shawl, beadwork	Belt / piece	Hot water jug	Metal, Bell
Headdress	Belt buckle	Cup / cover cup	bell
Evening dress	Medallion	Cup cover	drum
Girdle	Arm ornament	Coffee pot cover	
Waistcoat, man's	Accessory	Chopstick	
Snowshoe	Hairpin	Beaker	
Helmet mask		Plate	
Trousers, man's		bowl	
Visor			
Helmet, parade			
Plate armour / piece			
Tartan / portion			
Bootee, baby's			
Poncho			
Sandal			
Garter			
Coat, uniform			
Shoulder covering			
Dress, chief's			
mask			
Bodice, dress			
Flying helmet			
Apron, double			

Table C-2 Terms for general objects/things continued

Materials (generic)	Tools/tech	Animals/reptiles/insects
Wood	Tape recorder	Sea urchin
Metal	Clock	jawless fish
Horn	Weight	Mole
Cloth	Powder horn	beaked whale
Bone	Badge Mould	Sheep
Textile	Microscope	Bottlenose dolphin
Pottery	Rotor	insect
Crystal	Tray	river dolphin
Hair	Propeller	Harbour porpoise
Fibre	Hammer	White-beaked dolphin
Bead	Robot	Fly

Fabric	Drone	Narwhal
Paint	Printing press	Sea anemone
Paste	Microscope, compound	Great scallop
Shell	Key	Termite
Skin	Loom	Tortoise crab
Wire	Telephone	African elephant
Pigment	Foot plough	Beetle
String	Ristle plough	Scarab
Thread	Flywhisk	Ant
Rock crystal	Microscope, pocket	sea cow
Tartan	Rocket	Red deer
Feather	Boiler	Minke whale
Gut	Electrotype	Fruit bat
Pitch	Thermometer	Centipede
Bark	Television receiver	Blue whale
Sand	Thread winder	Polar bear
Nit	Tower clock	Seal
Gemstone	Spindle whorl	Marsupial Mole
Thread	Apparatus	Komodo dragon
Ash	Wind turbine	Lizard
Strap	Earphone	Nightjar
Synthetic material	Electric motor	Iberian lynx
Fur	Thermometer	Lion-tailed macaque
Greenstone	Water dropper	Pearly nautilus
Mud	Bellows	Thresher shark
Clay	Harness	Spider
Gas	Microscope, simple	Canadian lynx
Grass	Slide rule	hippopotamus
Root	Shepperd’s horn	Mangrove crab
Cord	Thermometer, Six's	Pallas' cat
Rope	Poker	sheep
Poster paint	Pitchometer	Deer
Leaf	Loom, hand weaving	monkey
Hard board	Fly whisk	bird
	Lamp, electric	hedgehog
	Electrophone stand	bear
		Brown rat
		Kitten
		owl
		moth
		insect

		birds
		cow
		mouse
		reptile
		Lion

Table C-3 Terms for person/group[12]

Role	Gender	Age	General
Cottier	Woman	Infant	Person
Enslaved	Man	Baby	People
Maiden	Transgender	Child	
Coigrich	Non-binary	Adult	
Shepperd	Two-Spirited	Older adult	
King			
Chief			
Maid			
Soldier			

Table C-4 Terms for actions/events

Action	Event
Curling	Wedding
Travelling	War
Forging	Plague
knitting	Death
Jumps	Festival
Football	Celebration
Communicating	
Dancing	
Weaving	
Mourning	

Table C-5 Terms for general time

General era	Season	Holiday
Ancient Egypt	Snow	Halloween
Medieval	Winter	Day of the Dead
Victorian	Spring	Valentine's
Tudor	Summer	Christmas
Renaissance	Autumn	

[12] For terms about people, particularly pertaining to gender or ethnicity, the following thesauri may be useful: the Homosaurus (International LGBTQ+ Linked Data Vocabulary) and Carissa Chew's Inclusive Terminology project.

Table C-6 Terms for general where - type of place

Place/building
Settle
Space
croft
church
factory
Shrine

Table C-7 Terms for abstract objects, beings, and ideas

Abstract concepts/emotions	Spiritual objects	Mythical/religious/ spiritual beings	Abstract themes
Soul	Shabti	Mermaid	Health/Wellbeing
Tradition	Dao	Unicorn	Belief: Religion, spiritualism, mythology
propaganda	Charmstone	God	Colonialism
penance	Paten	Goddess	Politics and society
peace	Caster	Fairy	Identity: gender, socioeconomic status, religion
Love	Crucifix	Nymph	Life and Death
Joy	Crosier head	Cherub	Environment/climate change
Sorrow	Ikupasuy		Outer space
Happiness	Oracle bone		
Sadness	Magic knife		
precious	Heper ay		
intimate	Prayer wheel		
heritage	Shesh nag		
history	Islamic		
Culture	Amulet		
memento mori	Charm		
	Pons ikayop		
	charm		
	Amulet / wadjet eye		
	Heart scarab		
	Lunula		
	Reliquary		
	Charm box		
	Reliquary		
	Charm box / half		

	Charm box, labourer's		
	Charm box / invocation		
	Communion flagon		
	Communion cup		
	Beaker, communion		
	Plate, communion		
	Vessel metal / chrismatory		
	Sculpture / figure / avatar / Navanita Krishna		
	Ibeji figure		
	Figure / water nymph		
	Votive bronze / uraeus		
	Figure / Surya / chariot / horse / charioteer / wife		
	Katsina doll		
	Figure / luohan		
	Guardian figure		
	Votive statuette / god		
	Votive statuette		
	totem pole		
	Figure, haniwa / horse		
	Votive statuette / bull		
	Thangka		
	Votive statuette / god / falcon		

Table C-8 General terms for visual, aesthetic & physical attributes

Colours	Lines	Shapes	Patterns	Physical attributes
Blue	Straight	Cross	Paisley	Fragment
Red	Sharp	Circle	Tartan	Miniature
Black	Curve	Curve	Marbled	Piece
Green	Layered		Gradient	Hoard

Orange	Cut			Knitted
Yellow				Decorative
Colourful				Embroidered
Pink				Reproduction
Gold				Smooth
				Ancient
				Classical
				Modern
				Beaded/beadwork
				Ornate

D. Appendix D: Alternative formats for keywords and image caption frameworks

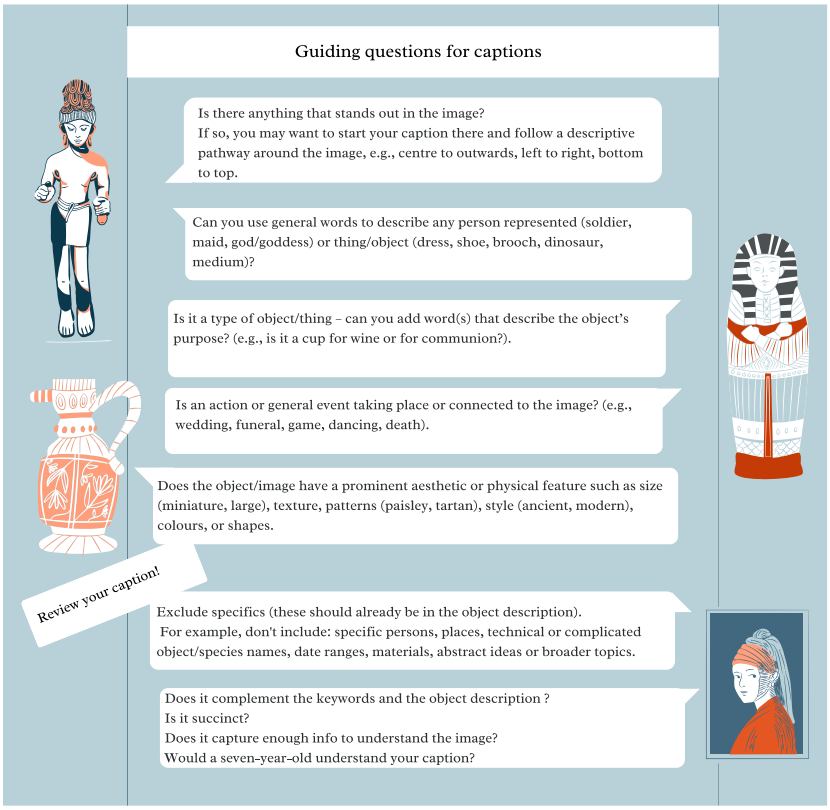


Figure D-1 Guiding questions framework for captions

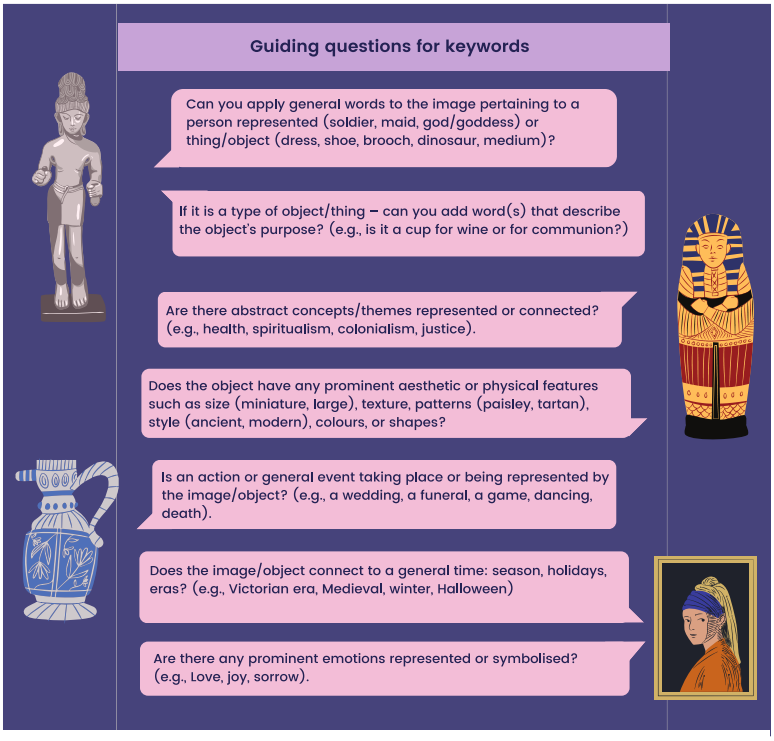


Figure D-2 Guiding questions framework for keywords

E. Appendix E User/non-user surveys - additional data charts

E.1 Search activities, users of cultural heritage images, and level of knowledge

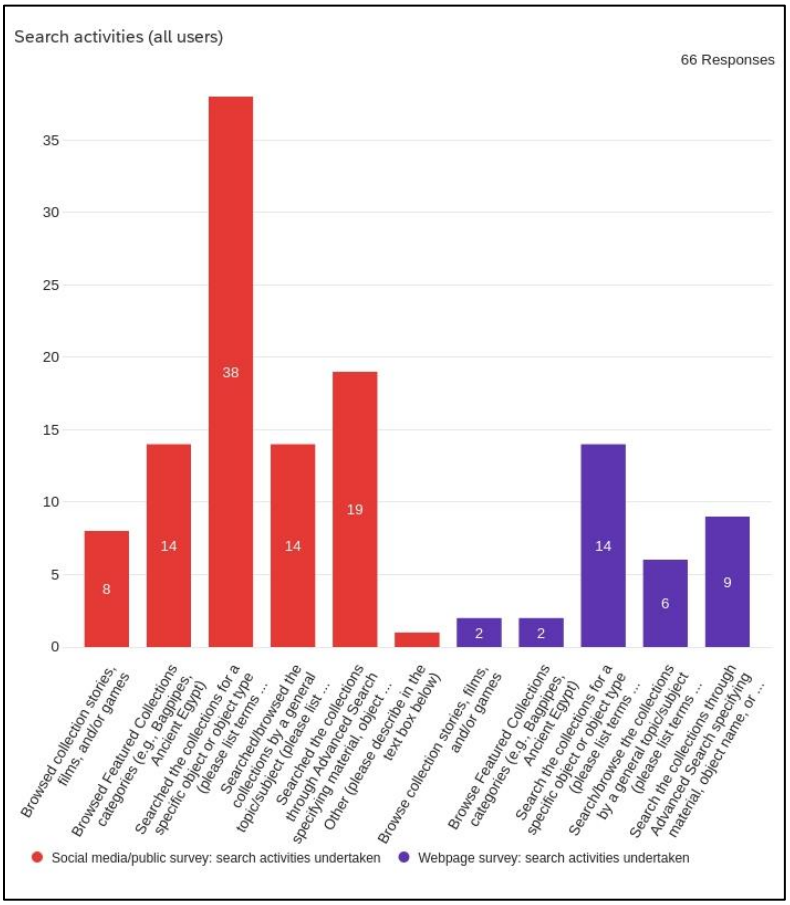


Figure E-1 Search activities

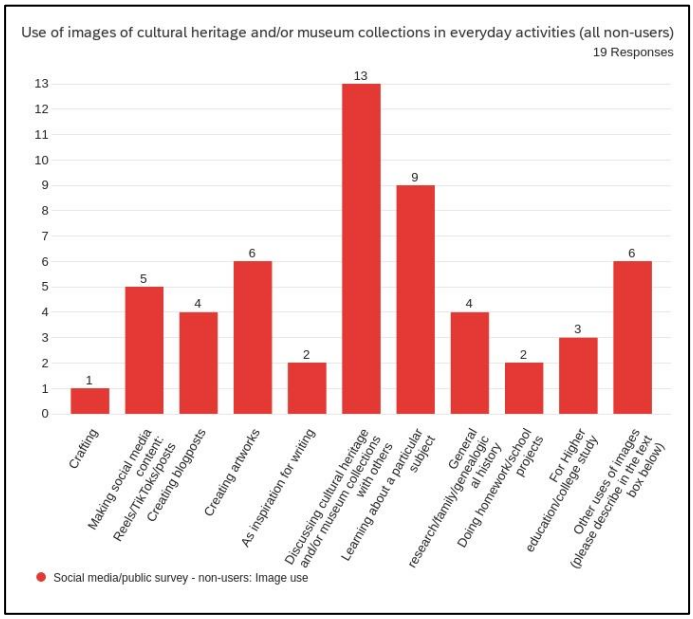


Figure E-2 Use of cultural heritage images

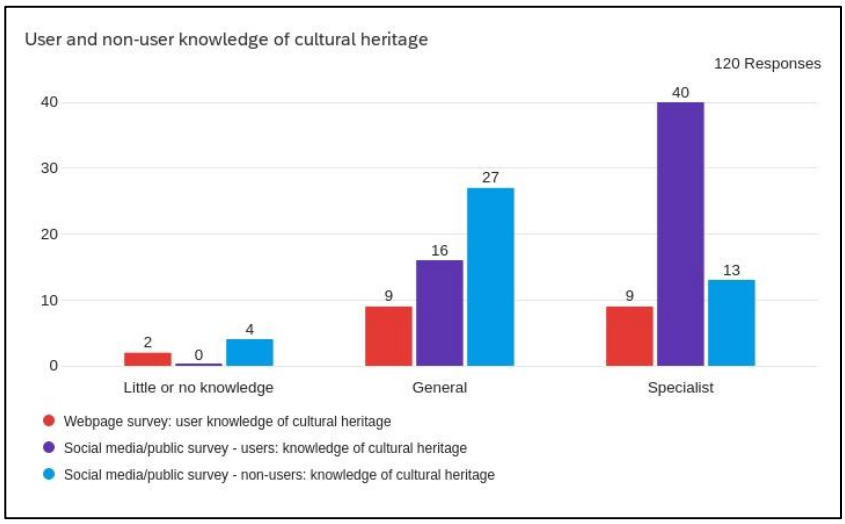


Figure E-3 Level of knowledge

E.2 Types of terms applied by all users in survey labelling tasks 3 and 4 (pie charts)

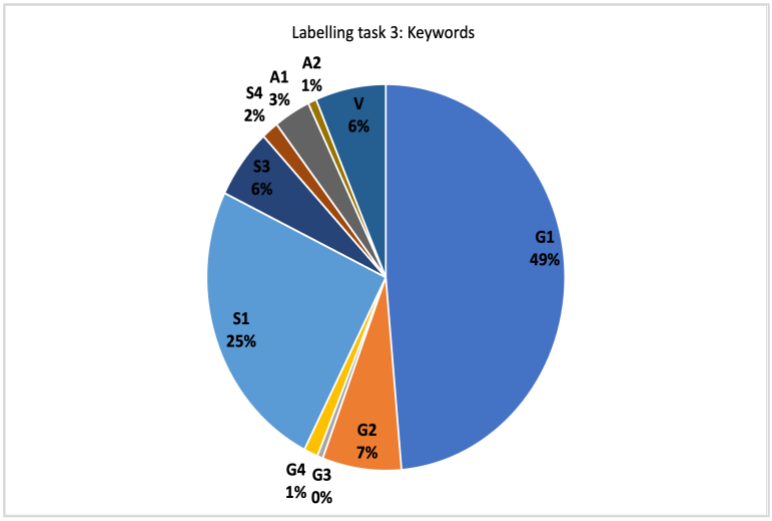


Figure E-4 Types of terms applied to 'brooches'

Figure E-5 Labelling task 3 'brooches'

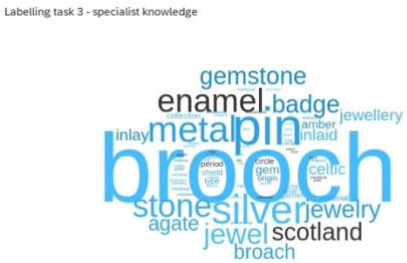


Figure E-7 Labelling task 3 – terms applied by specialist knowledge users

Figure E-6 Labelling task 3 – terms applied by general knowledge users

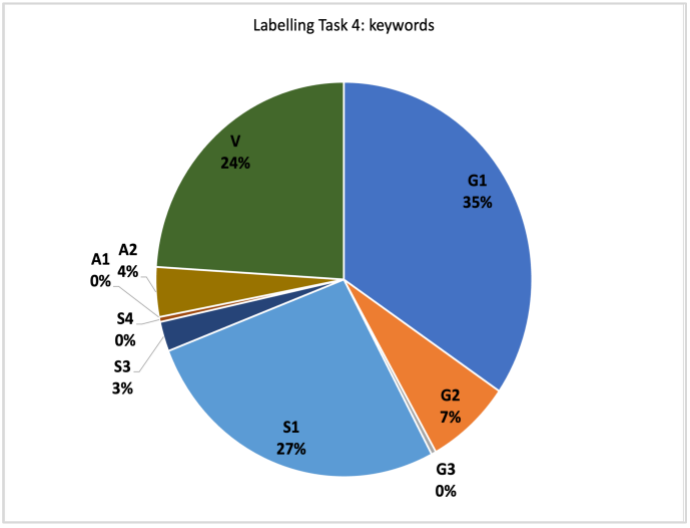


Figure E-9 Types of terms applied to the 'agate'

Figure E-8 Labelling task 4 'agate'

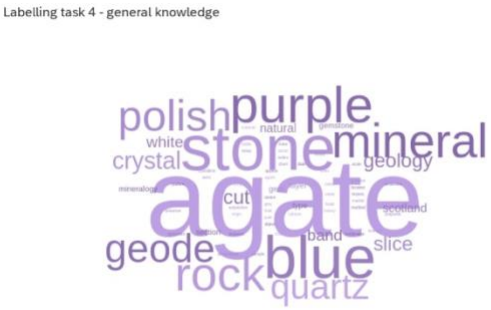


Figure E-10 Labelling task 4 – terms applied by general knowledge users

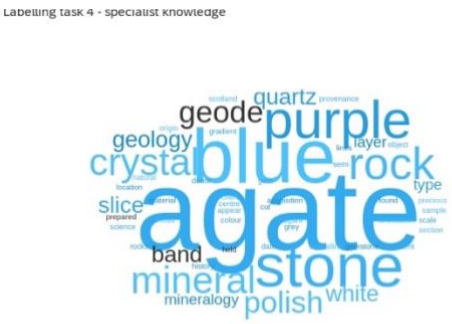


Figure E-11 Labelling task 4 – terms applied by specialist knowledge users