The Mapping Museums Knowledge Base: technical report

Valeri Katerinchuk, Nick Larsson, and Alexandra Poulovassilis

Birkbeck, University of London, Malet Street, London WC1E 7HX

Abstract. We have developed an RDF/S knowledge base that encompasses the first comprehensive dataset of the UK's museums. This knowledge base is making possible for the first time evidence-based research into the emergence, development and current status of the UK's museum sector. The knowledge base is accessed via a web application that allows researchers to browse, search and visualise data relating to over 4000 museums according to attributes such as subject matter, location, size, governance status, accreditation status, and year of opening/closing. This report describes the knowledge base and its usage to date.

1 Introduction

Numerous reports over recent decades have highlighted the lack of a single integrated dataset covering all of the UK's museums, e.g. most recently the Mendoza Review [13]. Without such a dataset it is hard to know what museums exist, where they are, what is their subject focus, when they opened, and what levels of visitors they have. This problem is particularly pronounced for small independent museums, a large number of which have opened in the UK since the 1970s. Without comprehensive information about the whole of the UK museum sector, it has been impossible to research the history, evolution and impact of the sector, with implications also for funders and policy makers seeking to make evidence-based decisions relating to local, regional and national development of the sector [2].

These problems motivated the work of the "Mapping Museums" project[1], as part of which we have developed a comprehensive knowledge base (KB) encompassing information relating to over 4000 UK museums that have been existence since 1960 — this is double the number of museums covered in any previous survey. We have also developed a web application that allows users to easily browse, search and visualise the knowledge base. The KB and web application are both freely accessible from the project website at www.mappingmuseums.org. This is a convenience URL which redirects to a web service maintained by our institution – Birkbeck, University of London – at http://museweb.dcs.bbk.ac.uk/home.

Mapping Museums (MM) is an interdisciplinary project involving researchers from museum studies, historical geography, oral history political history, computer science, and geographic information science[2]. Development of the KB has been a lengthy and challenging process, requiring the identification, cleansing and integration of data from numerous disparate data sources, and the development of new conceptualisations of what constitutes a "museum" and what are the key attributes of museums that need to be captured in order to support research into the evolution of the sector.

This report describes the knowledge base and its usage to date. The knowledge base can be accessed through the Browse, Search and Visualise facilities of the Mapping Museums Web Application — see https://museweb.dcs.bbk.ac.uk/aboutapp. The data can be downloaded from http://museweb.dcs.bbk.ac.uk/data in the form of an RDF/XML or CSV file. The source code of the Web Application can be downloaded from http://museweb.dcs.bbk.ac.uk/software (the source code has been deposited on GitHub at https://github.com/MappingMuseums/MappingMuseums-MMWebApp including technical documentation).

2   Related Work

The most similar dataset to the MM knowledge base is the Institute of Museum and Library Services (IMLS) dataset which includes data on approximately 30,000 museums in the USA and aims to be "useful to researchers, journalists, the public, local practitioners, and policymakers at the federal, state, and local levels for planning, evaluation, and policy making"[3]. It adopts a single-level classification of museums' thematic focus comprising just 9 terms in contrast to our fine-grained Subject Matter taxonomy comprising 21 top-level categories and 108 sub-categories (see Section 3). It does not include governance, accreditation, or museum opening/closing dates. It also adopts the International Council of Museums (ICOM) definition of what constitutes a museum [8] whereas much broader criteria were applied to decide which venues should be included in our knowledge base (see Section 3). However, IMLS does include financial information about museums, which our knowledge base does not because this was not within the research scope of the MM project. There are no plans to keep the IMLS dataset updated ("IMLS does not currently plan to maintain or update the files beyond this 2018 release") whereas we have developed facilities for the public to be able to add data about new museums into our knowledge base, and to edit data about existing museums (see Section 5).

The Finding GLAMs wiki project[4] is aiming to "build a truly worldwide database of cultural heritage institutions and their collections". For the UK, it provides a list of members of the Museum Association and access to data on accredited museums sourced from Arts Council England. These are just two of the primary data sources used to compile our knowledge base (see Section

---

[2] See the Research Team page at www.mappingmuseums.org for details.

[3] https://www.imls.gov/research-evaluation/data-collection/museum-data-files

[4] https://meta.wikimedia.org/wiki/FindingGLAMs

2) and so, thus far, Finding GLAMs is limited only to England and further does not consider unaccredited museums which form the majority of the small independent museums of particular interest to the MM project. The GLAMs project also does not provide guarantees of veracity or up-to-dateness of the listed and linked resources whereas the provenance and authority of the data in the MM knowledge base are of paramount importance in order for the data to be credible for domain experts and museum professional audiences.

Much other work over the past decade has digitised and published detailed information about cultural artefacts as Linked Open data [7, 6, 4, 12, 1, 9], typically mapping from museums' internal data representations to the CIDOC CRM [3] standard or to some other fixed schema. In contrast to these, our knowledge base aims to support research into a whole museum sector. The CIDOC CRM offers rich modelling of the cultural heritage domain and cultural heritage artefacts but it does not address museums' governance, accreditation, thematic focus, opening/closing, and geodemographic context, which need to be captured in order to carry out research on the development and status of the UK museum sector as a whole. Because of the heterogeneity of the datasets that were identified and integrated to compile the MM knowledge base, it was not possible to adopt from the outset a single fixed schema and this emerged incrementally as the domain experts' conceptualisations and requirements for key attributes that needed to be captured evolved.

The CALMA digital library [17] also adopts an iterative Linked Data approach to building a layered digital library. However, that work focuses on the iterative extraction of information from a large, complex but existing online resource. In contrast, the MM knowledge base was developed concurrently with the data collection and our methodology needed to be able to handle the increasing complexity of evolving data, data conceptualisations, and users' requirements.

3   Creating the Knowledge Base

Due to the disparate and incomplete data collection within the UK's museum sector, there was no single dataset upon which to base the design of the MM knowledge base. Specifically, the MM project's museum studies experts collected and integrated data from numerous primary sources:

- The Museums and Galleries Commission [14, 25]. During 1994-1999, its Digest of Museum Statistics (**DOMUS**) survey collected information on some 1700 accredited museums.
- The Standing Committee on Museums and Galleries [22, 24].
- The Museums, Libraries and Archives Council, England [5, 13].
- The Northern Ireland Museums Council [16].
- The Scottish Museums Council (now Museums and Galleries Scotland) [23, 11].
- The Welsh Museums, Archives and Libraries Division [15].
- Arts Council England (**ACE**), which keeps a list of all the accredited museums in the UK.

- The 1983 survey of independent museums by the Association of Independent Museums (**AIM**).
- The 1987 Museums.UK database project of the Museums Association [21].
- The Museums Association's Yearbook recording details of all its members since 1955; also its Find-A-Museum service[5].

Each of these surveys had different aims and none of them cover the whole of the UK museum sector. They also do not share any common data collection standard. Other challenges faced included lost, missing and poorly archived data, and varying definitions of what constitutes a "museum"[6]. The MM museum studies experts opted to begin the data collection with the Digest of Museum Statistics (DOMUS), 1994-1999, because this was the most comprehensive of the above datasets. It comprises several hundred spreadsheets that needed to be manually cleansed and integrated to create a core data spreadsheet comprising one row per museum. This was extended with data on additional museums from the other sources listed above (there was no practically feasible method to automate the integration process). Each museum in our knowledge base has a unique identifier including as one of its components the primary source of the data about this museum, e.g. the identifier mm.domus.WM014 relates to museum WM014 originally sourced from DOMUS. The primary provenance of each museum's data is also explicitly encoded.

Data about many of the museums was incomplete, for example with missing addresses, governance status, opening and closing dates, and visitor numbers. To improve coverage, the MM museum studies experts consulted historic guidebooks, gazetteers, regional guides, specialist museums sources, and online resources (museums' websites, Wikipedia, TripAdvisor, the BBC 1986 Domesday project). They also consulted museums staff, tourist boards, local history societies, museum consultants and cultural heritage organisations. Face-to-face data validation sessions were held with the nine regional branches of the UK Museum Development Network (MDN) during which the data was examined line-by-line[7].

By the end of the main data collection phrase there was full coverage of most attributes on over 4,000 museums apart from visitor numbers, governance, opening date, and closing date: visitor numbers were known for 67% of museums, governance for 92%, year of opening for 88%, and year of closing for 68% of closed museums. For the missing governance data, we created a sub-category of Unknown governance. For missing opening dates, we recorded an interval of the form [earliest possible year of opening, latest possible year of opening] and

---

[5] https://www.museumsassociation.org/find-a-museum

[6] Detailed discussions of these issues can be found in the MM Research posts on "Problems with the Data" and "Defining museums" at http://blogs.bbk.ac.uk/mapping-museums/

[7] Details of the data compilation, cleansing and validation can be found in the Research posts on "Getting Started: Compiling the data" and "Picking the brains of the Museum Development Network", and the Events post "AIM: I'm going to map forever", at http://blogs.bbk.ac.uk/mapping-museums/

likewise for the missing closing dates: we refer readers to [20] for details of the modal logic and interval-based temporal operators that we developed within the Web Application to reason with such date intervals. For visitor numbers, the main requirement was to use these to infer the *size* of a museum, and for this the MM team decided upon the categories Huge (1M+ visitors per year), Large (50,000-1M), Medium (10,000-50,000), Small (0-10,000) and Unknown. Machine learning was used to reduce the category of Unknown from 33% of museums initially to just 1.6%, by inferring the size of museums for which there was no visitor number data from their other attributes[8].

The criteria for including venues in the MM knowledge base needed to be broader than official definitions of what is a museum (e.g. those of ICOM or the UK Museums Association) so as to ensure that small independent museums were included, which would not have been possible otherwise[9]. Generally, a venue needed to have a collection of objects on display, be open to the public, occupy a demarcated display space, and be engaged in the preservation of the display objects.

The incremental data collection was accompanied in parallel by the incremental design of the conceptual model of our KB. As part of this effort, a new Subject Matter taxonomy needed to be designed so as to cover the subject areas targeted by newer independent museums. The most recent taxonomy for classifying the overall subject of a museum (as opposed to the artefacts or collections within it) was devised for the DOMUS project in 1994. However, this did not fully encompass the popular "non-academic" subject areas that are the focus of some newer independent museums. More recent taxonomies have been designed to categorize museum collections, but these had specific emphases not relevant to the needs of the MM project. While taking careful note of historic and existing taxonomies, the MM museum studies experts developed a new classification for museums' subject matter. Museums were grouped into recognizable major categories such as 'arts' and 'transport'. If a number of similar museums did not easily fit into the existing classes new classes were devised, and sub-categories were introduced when a group was large and unwieldy. The number of sub-categories varies by subject type. Large categories such as 'transport' have several sub-categories, while the relatively small category 'food and drink' has none. More inclusive terminology was used than was previously the case (e.g. 'military' was replaced by 'war and conflict').

This new Subject Matter taxonomy can be viewed using the Browse facility of the MM Web Application by clicking on the Subject Matter attribute. Like the data, this taxonomy was reviewed and evolved in collaboration with several ex-

---

[8] Details of the variability of the visitor numbers data and the derivation of the Size attribute are discussed in the Research posts on "Missing, massaged, and just wrong: Problems with visitor numbers" and "How big is that museum" at http://blogs.bbk.ac.uk/mappingmuseums/

[9] Discussions of the MM project's evolving conceptualisation of what constitutes a museum can be found in the Research posts on "Defining museums" and "Surveying Museums: What's in and What's out?" at http://blogs.bbk.ac.uk/mapping-museums/

ternal museum consultants and with members of all nine Museum Development Network groups.

The incremental accrual of both the data and the domain experts' conceptualisations mandated the use of semantic web technologies: representing both data and schema information in the form of (subject, predicate, object) triples allowed the evolving relationships between entities to be easily extended as new knowledge and data accrued. The KB underwent 10 formal releases within the first 24 months of the project; between these, numerous additional iterations took place to validate new data conceptualisations and new data, improve data upload facilities, and correct data errors. This pace of change would have been impossible to handle within the time and resourcing constraints of the MM project using structured data modelling/management approaches such as relational databases.

Using semantic web technologies also made it straightforward to extend the evolving schema of the KB with other taxonomies as the researchers' need for these became apparent, specifically with (i) the Administrative Area hierarchy of the Office for National Statistics[10], and (ii) the Deprivation Index[11] and Area Classification[12] relating to the administrative area of each museum, in order allow detailed geographical analysis of the museums data. We chose to model all of these external taxonomies using RDFS superclass/subclass relationships in order to retain their original modelling and therefore promote reusability by others who are familiar with those taxonomies.

To generate the KB a pre-processing step is first applied to extend the core spreadsheet with additional data for each museum, specifically: its inferred size, its geographical coordinates — derived using the ONS Postcode Lookup Dataset[13], the Administrative Area in which it is located, and the associated Deprivation Index and Area Classification information. An RDFS template is generated from the spreadsheet header and combined with the spreadsheet data to create a set of RDF/S triples in N3 format. This is converted to RDF/XML using RDF2RDF[14] and finally loaded into Virtuoso[15] (we refer readers to [20] for full details of the KB generation process).

Because of the timing and funding constraints of the MM project, the development of the KB and the Web Application (Web App) needed to start well before the end of the data collection. Also, the domain experts' research to be

---

[10] https://data.gov.uk/dataset/7709b64e-369f-41f4-96ce-1f05efde9834/national-statistics-postcode-lookup-august-2017

[11] https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015, https://gov.wales/statistics-and-research/welsh-index-multiple-deprivation/?lang=en, http://simd.scot/2016/, https://www.nisra.gov.uk/statistics/deprivation/northern-ireland-multiple-deprivation-measure-2017-nimdm2017

[12] https://www.ons.gov.uk/methodology/geography/geographicalproducts/areaclassifications/2011areaclassifications/datasets

[13] https://data.gov.uk/dataset/7709b64e-369f-41f4-96ce-1f05efde9834/national-statistics-postcode-lookup-august-2017

[14] http://www.l3s.de/~minack/rdf2rdf

[15] https://virtuoso.openlinksw.com/rdf/

supported by the KB and Web App needed to start before these had been finalised. Thus, we adopted an iterative, participatory methodology to develop the KB and Web App. There were 10 major iterations of the KB over the first 24 months of the project, each created using the above process, and overlapping with the project's main data collection phase (months 1-20), elucidation of the project's research questions (months 1-15), and the domain experts' research into the development of the UK museum sector (month 16 to date). Early versions of the Web App were developed alongside the KB, enabling schema and data validation to easily be carried out from the outset by the domain experts using the Web App itself.

The Web App comprises three main facilities:

Browse: Supports user-friendly navigation through the KB schema, allowing drill-down into the museums data through properties such as Accreditation, Governance, Size, Subject Matter, Location, Year opened, Year closed. The browsing results (i.e. subsets of museums) can be viewed on a zoomable map of the UK (see Figures 1, 2), as a list (Figure 3, right-hand pane), or museum-by-museum showing full details for each museum (Figure 5).

Search: Allows users to undertake finer-grained search over the museums data by specifying one or more filter conditions (see Figure 3, left-hand pane). Again, results can be viewed on a map, as a list, or museum-by-museum.

Visualise: Provides bar charts, line graphs and heatmaps to support research into the opening and closing patterns of museums over time, allowing drilling down into dimensions such as Accreditation, Governance, Size, Subject Matter, Location (see Figure 4).
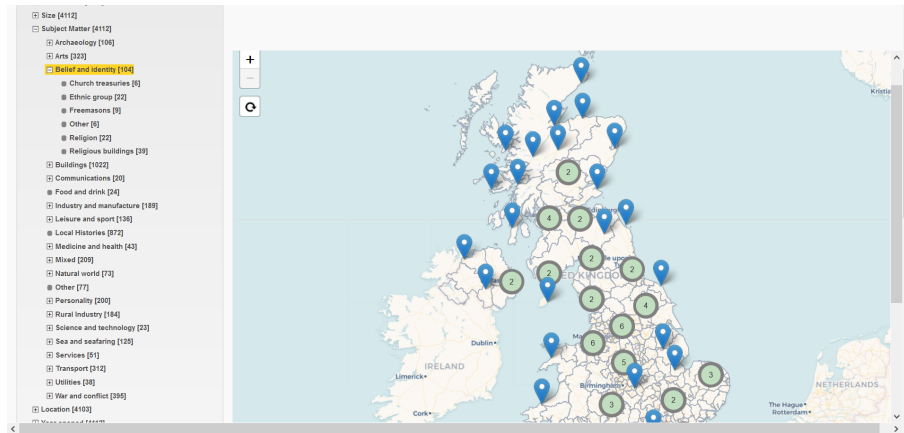
Fig. 1. Browse - all museums with subject focus Belief and identity. The Subject Matter taxonomy is partially opened up in the left-hand pane
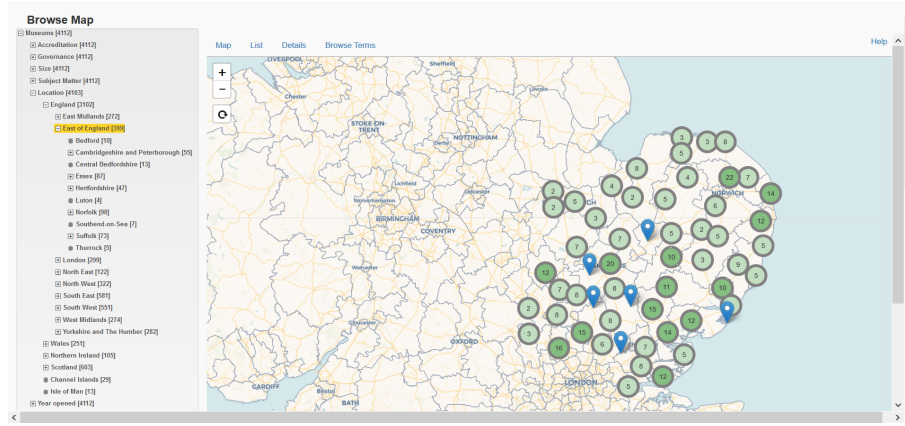
Fig. 2. Browse - all museums located in the East of England. The Location taxonomy is partially opened up in left-hand pane

The Web App has a three-tier architecture comprising a Web Browser-based client served by a Web Server connecting to a Database Server. The KB supports a SPARQL end point for communicating with the web server. We refer readers to [20] for details of the Web App's functionalities and the technologies used for its implementation. The source code has been deposited on GitHub at https://github.com/MappingMuseums/MappingMuseums-MMWebApp including technical documentation.

Our methodology to develop the KB has commonalities with DILIGENT [18] and HCOME [10], which focus on the evolving nature of ontology specification. However, in contrast to these and other ontology development methodologies, our KB had to be developed concurrently with the data collection and with the development of the Web App. Similarly to HCOME, the MM project's museum studies experts were central to the KB development life-cycle and tools were provided for them to be able to fulfil their role, notably the Browse facility of the Web App into which we immediately embedded the evolving versions of the ontology.

Our ontology uses the class, subclass, object property, data property, domain and range axioms of RDFS. There are 788 classes (of which 100+ relate to the MM Subject Matter taxonomy and 500+ relate to the external taxonomies from the ONS), 52 properties and 748 subclass axioms.

The full set of properties associated with each museum can be viewed by selecting to view museum-by-museum details using the Details view of Browse, illustrated in Figure 5. The data relating to a museum can also be viewed using a resolvable URL of the form http://museweb.dcs.bbk.ac.uk/Museum/*id* where *id* is the museum's unique identifier within our KB, e.g. mm.domus.WM014 for the Bilston Craft Gallery museum.

Fig. 3. Search - a user query comprising four filters; results shown in List view

The web page on Data, at http://museweb.dcs.bbk.ac.uk/data, provides a download of the KB in the form of a machine-readable RDF/XML file so that it can be used as a stand-alone resource by other applications and ontology browsing tools[16]. Also downloadable from the same page is a document containing metadata about the RDF/XML file, and a set of schema diagrams of the KB schema, partitioned into its major components for easier viewing. That page also lists the usage rights of the data: free to use under the terms of the Creative Commons (BY) license; and the canonical citation for using any downloaded data: "Data downloaded from the Mapping Museums website at www.mappingmuseums.org, Accessed on (date)."

The web page on Software, at http://museweb.dcs.bbk.ac.uk/software, allows download of the source code of the Web App from GitHub[17] where technical documentation relating to the Web App is also available. That page lists the usage rights of the software: free to use and extend under the GNU General Public License.

The data and software usage rights are also summarised on the Copyright & Access page of the website, at http://museweb.dcs.bbk.ac.uk/copyright.

Accessible from the project website is a Glossary of terms used in the KB as a whole. The Web App includes Help sections for each of its Browse, Search and Visualise facilities, as well as a glossary for the subset of terms associated

---

[16] This file contains the schema, followed by the data. It does not contain quite the entire MM KB because historical data about some museum attributes has not been published as it was deemed too incomplete by the project's museum studies experts. Due to the specialist and disparate nature of our primary data, it has not been possible within the resource constraints of the project to enrich the KB with OWL:sameAs links in a way that would be systematic and comprehensive; but that could be an area of further research in a future context of additional funding.
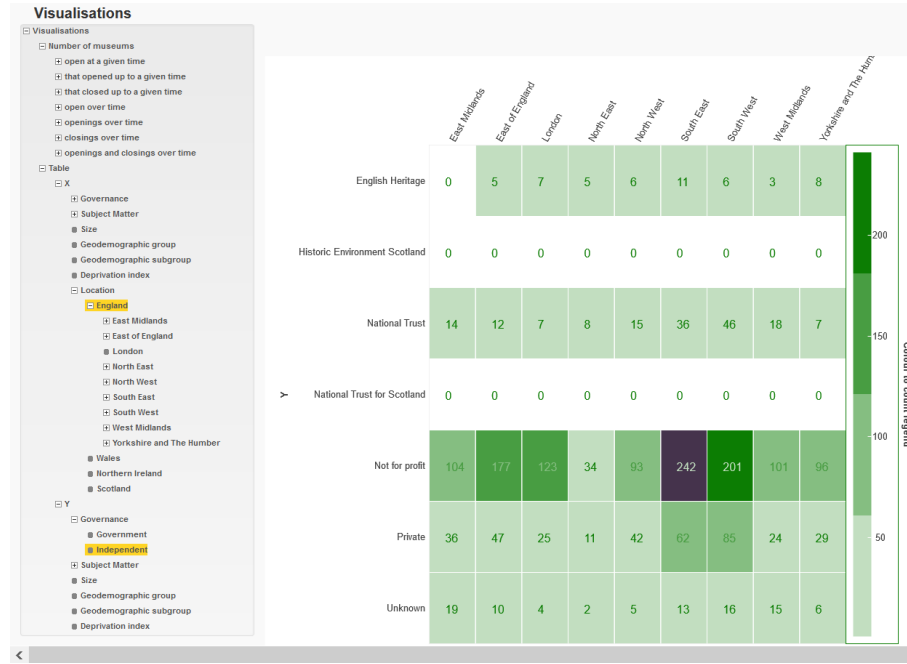
[17] https://github.com/MappingMuseums/MappingMuseums-MMWebApp

Fig. 4. Visualise - heatmap showing the number of independent museums of each type in each region of England

with each of these facilities. An Online Guide and a Video are available showing users how to access the KB through Browse, Search and Visualise.

4   Usage

As discussed above, the process of creating the KB enabled, in parallel, the formulation of detailed research questions to be investigated by the MM project's humanities scholars. These included:

- identification of broad patterns of museum openings and closings;
- analysis of detailed patterns of sector development, focusing on the relationships between high rates of museum openings/closings and attributes such as accreditation, governance, location, size and subject matter;
- using the Area Classification and Deprivation Index to analyse in detail the geodemographic context of museums' openings/closings, and cross-correlating these aspects with other museum attributes;
- establishing criteria for selecting museums for follow-on interview-based research and for undertaking archival research into the broader context of areas of museum growth.

Fig. 5. Details of the Ullapool Museum And Visitor Centre (scrolling down the list of attributes, left-to-right, top-to-bottom)

Our Knowledge Base has made possible for the first time this kind of research, and the Publications page at http://museweb.dcs.bbk.ac.uk/publications lists articles reporting on the outcomes of the research to date. A major publication, published on 17 March 2020 at the same time that the resource itself was first made publicly available, was Fiona Candlin *et al.* "Mapping Museums 1960-2020: A report on the data". This report provides a summary of the findings to date from the data analysis, including the identification of periods of growth and contraction of the sector, and variations in museums' opening and closing depending on location, size, governance, accreditation and subject matter.

A formative evaluation of the first complete version of the KB was undertaken in summer/autumn 2018. For this, 15 participants were recruited via the MM project's network of museum professionals, comprising consultants, digital archivists, museum accreditation assessors, officers from the Museum Development Network (MDN), and staff from AIM and ACE. These participants had not been involved in earlier stages of the research. Details of the activities we designed for this formative evaluation study can be found in [20] as well as analysis of the participants' performance in the activities and their responses to a questionnaire seeking their views on the usability and usefulness of the resource. The participants' responses (see [20]) provided an overall endorsement of the resource and also articulated its usefulness beyond the research objectives of the MM project. Specifically, the sixth and seventh questions in the questionnaire sought participants' views of the usefulness of the resource for themselves in their own role (sixth question) and for other people or institutions (seventh question). Participants' replies included:

"It would be useful to understand all museums in one area and look at accreditation status. Additionally helpful to understand dates of closure" [Cultural Heritage expert]

"Provides great statistics + trends to show development, particularly of independent museums sector - gives data for setting strategy." [Museum Development Officer]

"understanding where to target cluster locations for training etc. E.g. if high concentration of unaccredited museums could target 'recruitment' drive" [Museum Accreditation Manager]

"Baseline for understanding what museums are open in research of historic loans, merged collections and accreditation status ... useful to know who are considered National and Accredited for potential loans/collaborations/disposals ... could hugely support future collections management" [Cultural Heritage expert]

"Museums would like to find other museums like them and see patterns of their development over time" [Regional Museum Development Officer]

Additional comments on the usefulness of the resource made during the end-of-session group discussion include:

"Would be useful for accreditation team because we could identify clusters of unaccredited museums and target them - and hold sessions in those areas" [Museum accreditation manager]

"Would be useful when you're talking to an individual museum, because you could identify like organisations and put them in touch" [Museum Development Officer].

The formative evaluation study was also very useful in identifying areas of enhancement to the Web App. The MM museum studies experts continued to engage with professionals from across the UK museum sector for the next twelve months to further understand their detailed requirements from such a resource (views where obtained from approximately 45 people working in policy, consultancy, museum development, specialist museums Trusts, and museum associations). The usage scenarios that emerged (over 60 scenarios) are listed in [19], and they informed the finalisation of the KB and Web App into an alpha version released in autumn 2019 as well as the design and development of the overall project website through which these and related resources are accessed. A summative evaluation study was undertaken at that time with a new set of ten independent domain experts (details are reported in [19]), the outcomes of which informed the beta release of the KB, Web App and website in March 2020.

The beta release of the resource went live on 17 March 2020, although unfortunately the planned launch event on that date had to be cancelled due to the ongoing global Covid-19 pandemic. Over 120 professionals from the UK museum sector had booked to attend the event, indicating the sector's strong interest in the work. Despite the cancellation, the project team has received much positive feedback since the beta release.

The final release of the WebApp software was on 21 December 2021. The KB schema has not changed since the beta release and is available/viewable as described earlier. Usage statistics of the resource between the time of the beta release and the final software release are shown in the Appendix. The KB data will continue to be updated into the foreseeable future using the Add and Edit facilities described below.

5   Adding and Editing museum data going forwards

We released as part of the beta release two new web services that allow update of data relating to existing museums in the KB and addition of data about new museums. Two forms accessible via the Contact Us tab on the website allow the public to upload such data, thus supporting a form of "crowdsourcing" going forwards. The submitted data are first validated and if necessary corrected by the MM domain experts before being uploaded to the KB, after which they are immediately visible within the Browse, Search and Visualise facilities.

This Add/Edit process consists of three stages. The first stage provides a form-based web interface allowing external users to submit data about new museums or suggest edits to data about existing museums. In the second stage, a member of the MM project team inspects the proposed Adds and Edits via a more extensive password-protected form that allows confirmation or correction of the user's input and expansion of their input to include data about additional attributes (see below). In the third stage of the process the Curator confirms the Added or Edited data, leading to the execution of code that carries out the updates on the KB.

In the first stage, the user-facing Add and Edit web forms include the user's Name and Email as required fields, to enable the Curator to contact them for more information or clarification if necessary. Fields are provided for entering the museum's name, address, and postcode as well as the year opened/closed, subject matter, governance, and visitor numbers as optional fields. Additional ad hoc information may be provided in an Extra Information field. The year opened/closed fields allow the input of a single year, or a range of years in the form of XXXX:YYYY if a precise year of opening/closing is not known. Subject matter and Governance are selectable from a dropdown list of possible values, to avoid ambiguity. Visitor numbers are selectable from a dropdown list of choices that map into values of the museum Size category (1-10,000 $\rightarrow$ Small, 10,001-50,000 $\rightarrow$ Medium, 50,001-1M $\rightarrow$ Large, 1M+ $\rightarrow$ Huge).

The Edit form provides an autocomplete facility for the museum name field to allow the user to easily find a museum whose data they wish to amend. If the name itself needs to be amended, this change can be entered into the Extra Information field

In the second stage the additional fields available to the Curator to add/edit are: Alternate Museum Name, Year Museum First Documented, Year Museum First Documented Source, Year Opened Source, Year Closed Source, Founder, Founder source, Governance source. For a new museum, Subject matter, Size, Governance, Accreditation are all mandatory. A Notes field is provided to allow the Curator to add any additional comments (this field can be searched over through the Search facility of the Web App).

As part of the third stage, the Latitude, Longitude, and Administrative Area attributes of the new/edited museum are automatically (re)generated from its new/updated postcode. The method for doing so uses the NSPL 2017 postcode dataset from the Office of National Statistics – details are given in [20]. Simi-

larly, the Deprivation index and Geodemographic attributes of the museum are automatically (re)generated from its new/updated postcode – see [20].

The four forms (the User's and Curator's Add and Edit forms) are implemented as webpages using HTML, CSS and Javascript. The server-side validation, data storage, and autogeneration facilities are implemented in Python. The insertion and update of data are implemented using Python-generated SPARQL statements submitted to the underlying Virtuoso repository which hosts the KB.

Through provision of this data insertion/update pipeline, the KB will be able to be extended into the foreseeable future through community efforts, supported by our institution's museum studies experts and IT personnel.

6   Conclusions

In this report we have described the Mapping Museums knowledge base, which constitutes the first comprehensive dataset of the UK's museums, covering longitudinal data from 1960 to the present day. We have also described the usage to date of this resource, which is making possible for the first time research into the emergence, development and status of the whole of the UK's museum sector. The disparate nature of the input data, incremental data collection and evolving knowledge accrual mandated the use of semantic web technologies to create the resource, which could not have been achieved within the time and resource constraints of the project otherwise.

Poulovassilis *et al.* [20] detail the methodology and methods used to develop the first prototype of the Knowledge Base and Web Application, including the design and outcomes of a formative evaluation study undertaken with an independent group of domain experts. Poulovassilis *et al.* [19] detail the methodology used to complete the development of the Knowledge Base, Web Application and project website, including the design and outcomes of a summative evaluation study undertaken with another group of independent experts. These evaluation studies provided positive endorsement of the resource and also identified applications beyond the original research objectives of the MM project in areas such as:

- fostering networking between similar museums,
- setting regional museum development strategy,
- targeting training for museum professionals to improve accreditation rates,
- understanding museums' collections management practices,
- assisting the work of organisations such as AIM, ACE and MDN, and
- informing government policy.

The Mapping Museums Knowledge Base was developed with input and feedback from numerous stakeholder organisations including: Arts Council England (ACE); Department of Digital, Culture, Media & Sport (DCMS); Museums, Archives, Libraries Division (MALD) Wales; Museums Galleries Scotland (MGS); Northern Ireland Museums Council (NIMC); the Museums Development Network (MDN), comprising nine regional organisations; National Trust;

Historic Houses Association; Association of Independent Museums (AIM); Museums Association UK.

During their interactions with the MM project team, staff from these organisations have articulated the uniqueness of the Knowledge Base and the ways that it will be of use to them going forwards. It was staff from these organisations who recommended to the MM team that a report be produced to present in summary form key findings from the data and who advised on the analyses that would be most useful for the sector (resulting in the March 2020 report mentioned in Section 4). Staff from these organisations also recommended that we produce the Online Guide and Training Video to support user stakeholders in accessing the KB through the Web App and conducting their own analyses. Some of their recommendations, such as regional workshops on using the KB, were put on hold due to the Covid-19 pandemic but the project team hopes deliver these in face-to-face or virtual mode as soon as is feasible.

At present, sector stakeholders each gather different data about the UK's museums, for differing purposes, at different timescales and with different degrees of sector coverage. These data collection efforts are not coordinated and there is no facility for data synchronisation and integration between organisations. Hence there is potential for the approach that we have adopted in the Mapping Museums project, grounded in the use of semantic web technologies, to inform the data collection and management practices of the whole sector going forwards.

Our adoption of semantic web technologies means that the Knowledge Base is readily extensible with additional schema and data in a future context of additional funding, for example to add information on visitors' demographics, museums' accreditation history, and museums' financial information, thereby supporting new use cases that would also be of relevance to the above stakeholders.

Our data contains substantial new information, particularly on small museums, and a copy of the RDF/XML data and metadata files was deposited at the end of September 2021 (the formal end of the MM project) in our institution's research data repository[18]. A hard copy was also deposited in the Micromuseums Archive at the Bishopsgate Institute[19] together with sound recordings of interviews with staff from over 40 museums and other printed materials collected by the MM project's researchers. The Mapping Museums Knowledge Base, Web App and website will continue to be maintained for free public use by our institution for the foreseeable future at http://museweb.dcs.bbk.ac.uk/home.

---

[18] https://researchdata.bbk.ac.uk/
[19] https://www.bishopsgate.org.uk/

References

1. Brownlow, R., et al.: An ontological approach to creating an Andean Weaving knowledge base. Journal on Computing and Cultural Heritage **8**(2), 11 (2015)
2. Candlin, F., Larkin, J., Ballatore, A., Poulovassilis, A.: The Missing Museums: Accreditation, surveys, and an alternative account of the UK sector. Cultural Trends **29**(1), 50–67 (2020)
3. CIDOC CRM Special Interest Group: Definition of the CIDOC Conceptual Reference Model, Version 5.0.4 (November 2011)
4. De Boer, V., et al.: Supporting linked data production for cultural heritage institutes: the amsterdam museum case study. In: Extended Semantic Web Conference. pp. 733–747 (2012)
5. Greenwood, H., Maynard, S.: Digest of Statistics for Museums, Libraries and Archives. London: Museums Libraries Archives (2006)
6. Haslhofer, B., Isaac, A.: data.europeana.eu - The Europeana Linked Open Data Pilot. In: International Conference on Dublin Core and Metadata Applications (2011)
7. Hyvönen, E., et al.: MuseumFinland - Finnish museums on the semantic web. Web Semantics: Science, Services and Agents on the World Wide Web **3**(2-3), 224–241 (2005)
8. International Council on Museums: Development of the Museum Definition according to ICOM Statutes (2007-1946). Available at: http://archives.icom.museum/hist_def_eng.html [Accessed 26 March 2020] (2007)
9. Knoblock, C.A., et al.: Lessons learned in building linked data for the American art collaborative. In: International Semantic Web Conference. pp. 263–279 (2017)
10. Kotis, K., Vouros, G.A.: Human-centered ontology engineering: The HCOME methodology. Knowledge and Information Systems **10**(1), 109–131 (2006)
11. Martinolli, M.: Visitor Attraction Monitor Report. Glasgow: Museums Galleries Scotland (2014)
12. Matsumura, F., et al.: Producing and consuming linked open data on art with a local community. In: COLD Workshop, at the 3rd Int. Conf. on Consuming Linked Data. pp. 51–62. CEUR-WS.org (2012)
13. Mendoza, N.: The Mendoza Review: an independent review of museums in England. London: Department for Digital, Culture, Media and Sport (2017)
14. Morris, B.: Report 1987-88: Specially featuring independent museums. London: Museums and Galleries Commission (1988)
15. Newman, T.: Spotlight on Museums. Museums, Archives and Libraries Division Welsh Government (2015)
16. NIMC: Northern Ireland Museums Council: Mapping Trends in Local Museums Survey. https://www.nimc.co.uk/research-and-publications/ (2016)
17. Page, K.R., Bechhofer, S., Fazekas, G., Weigl, D.M., Wilmering, T.: Realising a layered digital library: exploration and analysis of the live music archive through linked data. In: 2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL). pp. 1–10 (2017)
18. Pinto, H.S., Staab, S., Tempich, C.: DILIGENT: Towards a fine-grained methodology for Distributed, Loosely-controlled and evolvInG. In: 16th European Conference on Artificial Intelligence. vol. 110, pp. 393–397 (2004)
19. Poulovassilis, A., Katerinchuk, N., Candlin, F.: Designing a system to chart the development of the UK museum sector: an iterative approach. In Preparation (2022)

20. Poulovassilis, A., Larsson, N., Candlin, F., Larkin, J., Ballatore, A.: Creating a Knowledge Base to research the history of UK Museums through Rapid Application Development. ACM Journal on Computing and Cultural Heritage **12**(4), 1–27 (2020)
21. Prince, D., Higgins-McLoughlin, B.: Museums UK: the findings of the Museums Data-Base Project. London: Museums Association (1987)
22. SCMG: Standing Commission on Museums and Galleries: Survey of Provincial Museums and Galleries (Rosse report) (1963)
23. SMC: A collective insight: Scotland's National Audit; full findings report. Edinburgh: Scottish Museums Council (2002)
24. Wright, C.W.: Provincial Museums & Galleries, a report. London: Department of Education and Science (1973)
25. Wright, M., et al.: UK Museums Retrospective Statistics Projects. Library & Information Statistics Unit, Loughborough University (2001)

A   Usage Statistics

No. of views/ accesses	Mar 2020	Apr 2020	May 2020	June 2020	July 2020	Aug 2020	Sep 2020	Oct 2020	Nov 2020	Dec 2020
2020 Report	166	139	98	67	47	45	77	45	60	48
Online Guide	19	22	24	21	13	17	14	6	7	11
Video	41	21	26	5	4	5	6	5	16	1
Browse	351	344	215	205	556	300	168	217	193	160
Search	305	278	301	185	255	249	146	128	179	240
Visualise	336	147	164	212	108	301	176	109	216	329
Interview transcript files	6	17	63	17	4	3	5	1	2	3
RDF/XML data file	not available	36	16	11	21	16	14	15	15	22
All other accesses to the MM website	20752	18613	13960	11503	14165	15172	10349	11013	10152	9958

Fig. 6. Mapping Museums Usage Statistics 17th March to 31st December 2020

No. of views/ accesses	Jan 2021	Feb 2021	Mar 2021	Apr 2021	May 2021	June 2021	July 2021	Aug 2021	Sept 2021	Oct 2021	Nov 2021	Dec 2021
2020 Report	53	58	60	63	59	50	33	27	65	70	61	63
Online Guide	7	8	5	11	6	6	5	7	8	2	6	6
Video	3	3	4	1	1	1	1	1	1	1	1	1
Browse	233	163	341	241	214	139	116	129	185	150	167	116
Search	295	257	314	200	197	147	98	103	255	164	188	117
Visualise	185	204	384	162	181	154	100	50	161	221	144	101
Interview transcript files	44	52	9	6	6	1	1	1	15	0	14	14
RDF/XML data file	10	11	38	27	19	19	17	23	8	14	11	10
All other accesses to the MM website	25549	23725	40749	30879	29792	14750	25949	18452	26572	38551	30477	35119

Fig. 7. Mapping Museums Usage Statistics January to December 2021