Digital Museum Map

Mark M Hall[1][0000−0003−0081−4277]

Martin-Luther-Universität Halle-Wittenberg, Von-Seckendorff-Platz 1, 06120 Halle, Germany mark.hall@informatik.uni-halle.de

Abstract. The digitisation of cultural heritage has created large digital collections that have the potential to open up our cultural heritage. However, the search box, which for non-expert users presents a significant obstacle, remains the primary interface for accessing these. This demo presents a fully automated, data-driven system for generating a generous interface for exploring digital cultural heritage (DCH) collections.

Keywords: Digital Cultural Heritage · Generous Interfaces · Browsing.

1 Introduction

The ongoing digitisation of cultural heritage artefacts has made large swathes of our cultural heritage, which previously were hidden in storage, available to the general public [7]. However, available does not mean accessible, particularly not for non-expert users, as the primary interaction method remains the search box. For non-expert users this often represents a significant hurdle, as they struggle to construct appropriate search queries [11,3,4].

Generous interfaces aim to address this by allowing the non-expert to explore the collection without requiring them to immediately enter query terms. Using a variety of approaches, they are more generous in offering up what the user might want to and can look at. What current generous interfaces generally lack is the ability to give the user an initial overview over what is available, before letting them freely explore that overview. The Digital Museum Map attempts to address this by providing an automatically generated physical-museum-style overview map, which the users can then use to explore the collection.

2 Generous Interfaces

"Generous interfaces" [10] aim to address the limitations of the search box by providing a generous overview or sample of the kind of things the user can find in the collection, which can then be explored through browsing. This primacy of browsing has the potential to enable rich interactions with the collection [2,8,1], particularly for non-expert users who have a strong preference for browsing [6,9]. Whitelaw overviews generous interfaces [10] and the Tyne & Wear Collection Dive [5] is also often highlighted as a good generous interface. These examples highlight one of the main limitations of current generous interfaces; they either do not provide a complete overview, or, where there is an overview, it is very shallow and does not allow the user to zoom into the data.

3 Digital Museum Map

The Digital Museum Map[1] (DMM) system addresses these limitations by providing an initial overview visualisation that provides the user with an overview over what is available in the collection and which the user can then interactively zoom into to explore the collection.

Data Loading: The DMM is a fully automated and data-agnostic system, thus to apply it to any DCH collection, all that is needed is a custom loading component. In this demo the source data is a collection of 1962 items from the National Museums Liverpool's World Museum Egyptology collection. These cover a wide range of objects from personal jewellery to architectural elements drawn from all periods of ancient Egypt. The items in this collection are annotated with the following meta-data attributes: category, culture, materials, measurements, date made, place made, measurements, collector, and date collected. Not all items are annotated with values for all attributes and the map generation takes this into account when generating the overview.

Structure Generation: To generate the structure that underlies the DMM the system first generates a fully data-driven hierarchical organisational structure and then annotates each node in the structure with context information.

The basic principle of the hierarchical structure is to recursively split the data-set into individual "rooms" of around 50-70 items that belong together conceptually (as defined through their meta-data). The target room size is driven by the number of items that can easily be displayed in the interface without requiring too much scrolling. The splitting algorithm uses a recursive divide-and-conquer approach together with a series of heuristics to automatically determine which attribute to use to split the data and how to group attribute values together to create cohesive nodes higher in the structure.

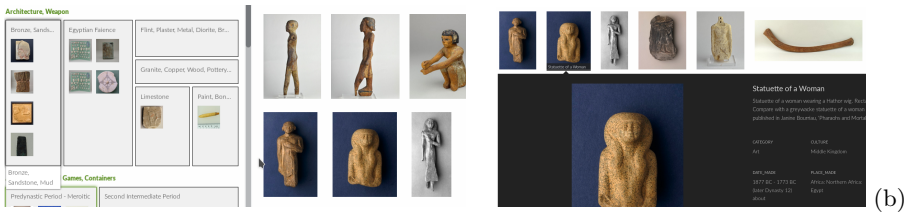
At each recursive step the algorithm analyses the distribution of attribute values for the current data sub-set to split. For each attribute it calculates how much of the data sub-set the attribute values cover and how many unique values that attribute has in the sub-set. It then chooses the attribute with the highest coverage and if there is a tie the one with the smallest number of attribute values as the attribute to use for partitioning.

The initial partitioning generally creates more than 10 nodes, many of which have few items, which would create a poor structure. Thus in the next step partition values are aggregated, either to create "rooms" with 50-70 items or at most 10 higher-level organisational nodes. Where the attribute has an intrinsic ordering (numeric, temporal, spatial) the values are first sorted using that ordering. Then consecutive values are grouped to achieve either the target room size or target number of nodes. Where there is no intrinsic ordering the values are sorted by the number of items assigned to each value. Then a round-robin approach is used to evenly distribute the values into the target number of nodes.

[1] Available at https://museum-map.uzi.uni-halle.de

After the organisational structure is generated, a title is assigned to each node. For ordered attributes, the first and last value are used, while for all other attributes all values are joined using commas. Each nodes is also annotated with context information drawn from Wikipedia.

Map Display: The resulting organisational structure forms the basis for the browser-based digital museum, which consists of three aspects: the museum map, the main item viewer, and the contextual information display (see Figure 1). For the main item viewer the aim was to have a display that lets the user view a large number of items at the same time. Thus initially only the item images are shown. The user has to click on an image to see the item's detailed meta-data. This is driven by the experience of visiting a physical museum, where upon entering a room, the visitor can see all items from a distance, and can then choose which ones they want to look at in more detail.



(a) (b)

Fig. 1. (a) One "floor" of the museum map with the selected "room" highlighted and items shown on the right. (b) The item viewer showing details for one item.

The museum map provides an overview over the top two levels of the organisational structure, where each group in the top level becomes a "floor" of the digital museum, while groups from level two form "rooms" within the "floor". The visual size of the rooms represents the approximate number of items in that room. To place a room, the ratio between the number of items in the room to place and the number of items in the other rooms of the floor is calculated. The space in the "floor" is then split using these ratios. It is always the longer side of the free space that is split, to ensure a relatively even combination of vertical and horizontal splits. For each room, a sample of items are shown as small thumbnails to give an idea of what the room contains.

When the user selects a room on the map, the main display shows the items in that room. Additionally, in the bottom-left corner the contextual information is updated to list the Wikipedia articles that the room was annotated with.

4 Conclusion

The Digital Museum Map automatically generates a data-driven, explorable overview for any DCH collection. It uses a series of heuristics to recursively

split the data until the leaf nodes contain 50-70 items for display. This structure is used to provide the non-expert user with an overview over what the digital museum contains and allows them to then freely explore that overview.

At the same time there are a number of planned improvements to the system. Currently attribute semantics are not taken into account when grouping. Additionally, while browsing is a powerful interaction method, users will also want to search the data-set and we plan to integrate search into the current interface. Finally, while the DMM's heuristics are tuned to provide a sensible split of the data, we are also investigating how to let users specify how they would like to see the data, to provide a fully user-driven and personalised DMM interface.

Acknowledgements

Thank go to National Museums Liverpool for providing their data.

References

1. Bates, M.J.: What is browsing-really? a model drawing from behavioural science research (4) (2007)
2. Chang, S.J., Rice, R.E.: Browsing: A multidimensional framework. **28**, 231–76 (1993)
3. Geser, G.: Resource discovery - position paper: Putting the users first **6**, 7–12 (2004)
4. Hall, M.M., de Lacalle, O.L., Soroa, A., Clough, P.D., Agirre, E.: Enabling the Discovery of Digital Cultural Heritage Objects through Wikipedia. In: Proceedings of the LaTeCH workshop held at EACL 2012 (2012)
5. John, C.: I dont know what im looking for: Better understanding public usage and behaviours with tyne & wear archives & museums online collections. Published January **29**, 2016 (2016)
6. Lopatovska, I., Bierlein, I., Lember, H., Meyer, E.: Exploring requirements for online art collections. Proceedings of the Association for Information Science and Technology **50**(1), 1–4 (2013)
7. Petras, V., Hill, T., Stiller, J., Gäde, M.: Europeana – asearch engine for digitised cultural heritage material. Datenbank-Spektrum **17**(1), 41–46 (Mar 2017). https://doi.org/10.1007/s13222-016-0238-1, https://doi.org/10.1007/s13222-016-0238-1
8. Toms, E.G.: Understanding and facilitating the browsing of electronic text **52**(3), 423–452 (2000)
9. Walsh, D., Hall, M., Clough, P., Foster, J.: The ghost in the museum website: Investigating the general public's interactions with museum websites. In: Kamps, J., Tsakonas, G., Manolopoulos, Y., Iliadis, L., Karydis, I. (eds.) Research and Advanced Technology for Digital Libraries. pp. 434–445. Springer International Publishing, Cham (2017)
10. Whitelaw, M.: Generous interfaces for digital cultural collections **9**(1) (2015), http://www.digitalhumanities.org/dhq/vol/9/1/000205/000205.html
11. Wilson, M.L., Elsweiler, D.: Casual-leisure searching: the exploratory search scenarios that break our current models. In: Proceedings of HCIR. pp. 28–31 (2010)