EuroRV3: EuroVis Workshop on Reproducibility, Verification, and Validation in Visualization (2018)
K. Lawonn, N. Smit, L. Linsen, and R. Kosara (Editors)

Visualizing Uncertainty in Cultural Heritage Collections

F. Windhager[1], V. A. Filipov[2], S. Salisu[1], and E. Mayr[1]

[1]Department for Knowledge and Communication Management, Danube University Krems, Austria
[2]Centre for Visual Analytics Science and Technology, Vienna University of Technology, Austria

Abstract

Visualizations of cultural heritage collections play an increasing role in supporting the sensemaking processes of visitors and researchers. While many visualization systems provide overview and exploration options for non-expert audiences or casual users, others support the in-depth analysis of collection experts like curators or art historians. These visualizations often rely on object metadata, which have been generated from historical catalogs or textual sources, each with their own standards of precision, certainty, and data quality. To make these levels of data quality transparent, the visualization of uncertainty poses a vital challenge for interfaces across the board. We introduce the PolyCube system as a visualization framework which provides four different perspectives on the geo-temporal origins of cultural collection data, including coordinated multiple views, animation, color coding, and a space-time cube representation. With specific regard to this multi-perspective framework we develop options to visualize spatio-temporal uncertainty and discuss ways and means of their coherent implementation.

Categories and Subject Descriptors (according to ACM CCS): [Human-centered computing]: Visualization toolkits—Visual analytics—Information visualization

1. Introduction

Due to the increasing digitization of contents from cultural heritage collections (CHC), massive stocks of pictures, texts or other cultural objects have become accessible on the web [WFS*18]. Visualizations of these collections can provide overview and orientation for visitors, they can foster immersion and contemplation of details on demand, or encourage open-ended browsing and exploration [DCW11]. While selected strategies of visualizing data quality and uncertainty have already been explored or mentioned in cultural heritage [KBD13,JFCS17], the field misses a more systematic discussion of options to represent uncertainty. This paper discusses such a design space with specific focus on time-oriented representations, which play a crucial role in CHC [DPC17]. These visualizations often combine certain data aspects (such as geographic origin of objects), with temporal information (e.g., on dates of origin) - with both data dimensions possibly affected by various levels of uncertainty. With this concept paper we discuss options to represent spatio-temporal data uncertainty with regard to different representation techniques, namely coordinated multiple views, animation, color coding, and space-time cube representations. These four techniques have been implemented in the PolyCube CHC visualization system, which is lacking options for uncertainty representation up to now [WMS*16, WSSM18].

To address the corresponding non-trivial design challenge on a general level, we summarize key aspects of the CHC domain with regard to users, tasks and data (sec. 2). We explore related works and methods of uncertainty visualization (sec. 3) and conceptually apply them to the PolyCube visualization framework (sec. 4). As for an outlook, we discuss future challenges for the design of advanced visualization systems in the CHC data domain.

2. Digital Cultural Heritage Collections - Data, Users & Tasks

Utilizing a triangular knowledge elicitation approach to the design of visual analytics tools [MA14], we outline major characteristics of our research area with regard to CHC data, users, and tasks.

2.1. Users

Many approaches to the visualization of CHC data aim to develop accessible representations for public audiences, which consist of casual, non-expert users [MFM*16]. Walsh and Hall [WH15] describe a casual user as someone "who has just stumbled across [the digital] collection in the same way that they would wander into the CH institution's physical space" (p. 1). In contrast to museum experts, casual users of digital collections do not have much knowledge on the structure and metadata of the CHC and how to navigate or search within. Also on the visualization side, casual users do not necessarily have expertise in visual-analytical thinking or in decoding uncertainty cues within visualizations.

2.2. Tasks

Casual use of CHC takes place in a free exploration setting: No external tasks exist, rather users are driven by some general in-

DOI: 10.2312/eurorv3.20181142

terest, to figure out what a CHC is about and find something interesting. To support them in this activity, users should firstly be supported in *gaining an overview*, e.g., by providing conceptual orientation or different perspectives on the CHC. Studies on the casual use of digital CHC showed that visitors needed an orientation phase before they started browsing or searching the exhibition [VCHR13, DCW11]. In museums, conceptual orientation, that is information on the museums' content and structure is often provided at the entrance. Overview visualizations can serve as a similar "entrance hall" for digital CHC by giving a overview and direction.

In line with the primary need of users to develop conceptual orientation, this paper focuses on multiple types of overview visualizations, which offer insights into the spatio-temporal origins of CHC. More specifically, we investigate options to represent spatial and temporal uncertainty, which often originates from CHC data.

How will the visualization of uncertainty influence users during their orientation and exploration activities? Hullman et al. [HAS11, p. 2220] assume that visual representations of uncertainty trigger more thorough processing of information and makes users double-check their inferences. Goodchild and Janelle [GJ10] assume that the processes of data analysis and modeling even *require* critical thinking about a variety of profound issues, such as accuracy, uncertainty, representation, scale, or (data) ethics [p. 8]. In the CHC context we assume that an adequate representation of widely existing data uncertainty can bring transparency, awareness and trust into the users' sensemaking process [SSK*16].

2.3. Data

To file and catalog their collections, cultural institutions (such as galleries, libraries, archives, or museums) document objects according to multiple (meta-)data dimensions, including place and time of origin (see Fig.1, left). In the following we discuss possible uncertainties in spatial and temporal data, which can arise from the documentation of the approximately point-like events of object creation.

2.3.1. Spatial Data Uncertainty (SDU)

Curators of CHC encounter at least three levels of SDU. Due to the relatively simple data structure in the CHC domain, we merge various dimensions of SDU, which have already been analytically distinguished (such as accuracy, precision, completeness, or credibility; cf. [MRH*05]) into a "general factor" of SDU. Yet this g-factor can be dissolved for more specific visual encodings later on.

- *Certain location data* define an objects' geographic origin in sufficient detail and precision. The delineation of "sufficiency" hereby depends on specific CHC domains and their information standards, which vary greatly from ancient artifacts to digital photo collections.
- *Uncertain location data* commonly arise either from vague place names (referring to regions, countries or continents) - or from vague signifiers connected to a precise localization (like "in the region of"). Accordingly, we distinguish *point-based* SDU and *polygon-based* SDU, which can appear on various *magnitude levels*, like "SDU1" (minor spatial uncertainty) and "SDU2" (major spatial uncertainty) (cf. Fig. 1).
- *Unknown location data* constitute maximum spatial uncertainty of curatorial knowledge, which is documented by missing specifications or other non-signifiers.

2.3.2. Temporal Data Uncertainty (TDU)

Also for the temporal data dimension, three levels of data quality or uncertainty can be analytically distinguished.

- *Certain dates of origin* are the best case, where an objects' time stamp marks a point in time with sufficient precision, with "sufficiency" again being tied to local CHC standards.
- *Uncertain dates of origin* are commonly marked by signifiers such as "around", "about", "circa", or question marks. In this context, imprecise time stamps can be either *point-based* (e.g., "around 1850"), or *interval-based* (e.g., "19th century", or even whole epochs like "Old Egyptian"). Also for TDU, magnitude levels obviously can vary (e.g., TDU1 and TDU2, cf. Fig. 1).
- *Unknown dates of origin* are the curatorial worst case, where no information on temporal origins is available, indicated by empty fields, question marks, or other non-distinctive signifiers.

2.3.3. Spatio-Temporal Data Uncertainty (STDU)

Spatio-temporal data uncertainty results from the combination of the above-mentioned dimensions and thus inherits all definitions and distinctions from temporal and spatial data definitions.

3. Related Work

In the CHC domain *SDU* is omnipresent; therefore, it requires ways and means to visually encode these data and to clearly communicate them. As many uncertainty visualization (UV) techniques have been proposed in other domains, we summarize prominent options and focus on a selection further down.

A study by MacEachren et al. [MRO*12] compared different techniques of SDU visualization with regard to intuitiveness and efficiency. Options of UV for point-like markers on maps included color saturation, color hue, color value, fuzziness, size, but also shape and texture. Kardos et al. [KMB] go beyond static UV options (like blur, fog, or pixel mixture) to also explore dynamic techniques like blinking pixel and animation [KMB]. For further elaboration, we will focus on the SDU visualization options of color saturation and blinking pixel.

Also *TDU* is omnipresent in CHC data, as curators frequently lack precise information on the temporal origins of cultural objects or items. To avoid simplifications or communication of false certainty, Kräutli et al. [KBD13] investigated different representations for historical TDU, including opacity gradient, variable line dashing, and variable wavelength.

A user study by Gschwandtner et al. [GBFM16] on variations of TDU representations included gradient plots, violin plots, accumulated probability, (centered) error bars, and ambiguation. The results of the evaluation suggest that the use of ambiguation is very effective to represent TDU - by using a lighter color value to depict the uncertain periods.

To visualize *STDU*, spatial and temporal UV options have to be

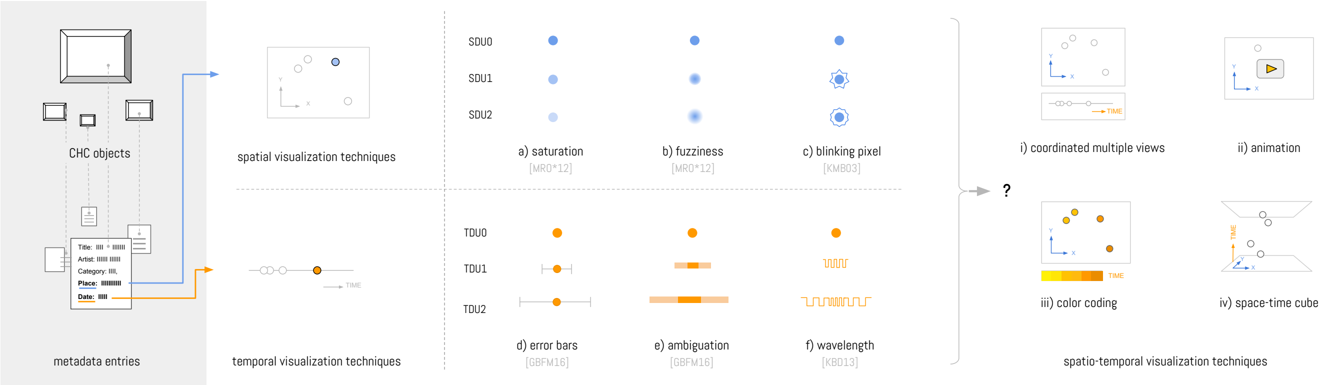

Figure 1: *The research question of this paper arises from the challenge to transfer options to visualize spatial (top) and temporal (bottom) object origins with various uncertainty levels (center) into the design space of spatio-temporal visualization techniques (right).*

combined. This poses a non-trivial design challenge, that has not been addressed in the CHC domain up to now. To analyze story accounts, Shrestha et al. [SZM14] proposed to plot latitude and longitude along two vertical axes, while a horizontal time-axis unfolds a flat (2D) coordinate system in which the spatio-temporal events of stories are plotted. While this technique discloses interesting patterns for analysts, it also distorts existing mental maps of territories due to its unconventional encoding choices. A work on the visualization of dengue fever epidemics by Delmelle et al. utilized a more promising approach for STDU visualization in a 3D environment [DDC*14]. This approach builds on the scaffold of a space-time cube representation, which will be elaborated further down.

4. Uncertainty Visualization in the PolyCube Framework

The PolyCube framework aims to provide highly effective spatio-temporal overviews for CHCs [WMS*16, WSSM18]. It aims for the optimization of its visualizations' user experience (to attract also non-expert users) and of their analytical and synthetical efficiency, to convey an overview on CHC in a minimum amount of time for casual users. As multiple visualization techniques can convey spatio-temporal overviews - and every single technique has its strengths and limitations [KPS14] - the framework offers multiple spatio-temporal representations. Thereby the system combines their analytical strengths in a synoptic visualization system and users can choose the right perspective for different tasks. To this end, the PolyCube interface offers (1) coordinated multiple views (CMV), combining a map and a timeline, (2) animation (ANI), mapping time to time, (3) a color coding (CC) perspective, mapping time to a color scale, and (4) a space-time cube representation (STC), mapping time to an additional spatial dimension (see Fig.1, right). To interconnect these perspectives visually, their selection and construction is mediated by seamless transitions, utilizing the transformational power of STC operations as a cognitive scaffold [BDA*16, WSSM18]. To further preserve the users' conceptual orientation, the visual encoding and design choices are kept consistent across different views [QH17, BWK00].

While the first implementation and evaluation cycles built on close to ideal (i.e., spatio-temporally precise) data [WSSM18], an exploration of CHC data from a large art history museum demonstrated the need to take massive amounts and significant degrees of data uncertainty into account. In the following, we discuss options to visually represent STDU in multiple spatial and temporal combinations and for all four spatio-temporal visualization techniques: Figure 2 provides an overview of the design space of the PolyCube framework, showing the development area of required uncertainty visualizations in case of i) SDU (top row), ii) TDU (middle row), and iii) STDU (bottom row). In the following we draw together options for these areas - with a specific focus on finding a consistent design solution across all four views. For exemplification purposes - and due to a design rationale of minimizing visual clutter - we decided on a color saturation technique to model SDU, and error bars to represent TDU. The concrete UV techniques that will be implemented in PolyCube have not been fixed, yet.

4.1. Uncertainty Visualization for Coordinated Multiple Views

From an implementation perspective, a CMV system brings along no novel challenge for the visualization of SDU, TDU, and STDU. On the plus side, every UV option already documented for 2D maps or timelines (see Fig.1, center) could also be applied for CMV (see Fig.2, first column). On the negative side, CMV systems enforce a split-attention effect on users and require increased visual work – and increased cognitive effort – for the mental integration of two separated representations, when it comes to the creation of a bigger (spatio-temporal) picture [SWSM16].

4.2. Uncertainty Visualization for Animation

Animation allows the straightforward use of color saturation for SDU, while the encoding of TDU has to be transformed from a spatially linear encoding (like error bars) to a visual encoding that is compatible with an ANI's dynamic encoding of time. For this purpose dynamic encoding (like a fading effect) or a second type of static encodings (such as fuzziness) could be utilized (see Fig.2, second column). As for the desired consistency of design choices [QH17, BDA*16], this already means a first breach of ideal conditions, or at least the need to communicate the new encoding method when switching from CMV to an ANI perspective.

Figure 2: *Design space for spatial, temporal and spatio-temporal uncertainty visualization for four different visualization techniques.*

4.3. Uncertainty Visualization for Color Coding

The color coding technique poses another challenge for the aim of UV encoding consistency. Exemplarily, the spatial UV method of color saturation easily interferes with the encoding of time via color scale. Against this background, a more discriminating encoding technique for SDU (such as blinking pixel) could be recommended. On the other hand, TDU could again be encoded via fuzziness, to keep the design consistent with ANI (see Fig.2, third column). Yet from a cognitive perspective, the introduction of a new SDU encoding technique has to be counted as another factor detrimental to the ideal of complete consistency or full visual momentum [BF12].

4.4. Uncertainty Visualization in the Space-Time Cube

The STC allows to return to the initial UV options of saturation for SDU and error bars for TDU (see Fig.2, fourth column), but also fuzziness and blinking pixel could be implemented. As demonstrated by Delmelle et al. [DDC*14] also more complex uncertainty shapes could be generated based on a STC representation. From a cognitive perspective, the occlusion effects of 3D encodings are known to require increased interactions (such as rotating or zooming), or the implementation of other measures of occlusion management [ET08].

5. Discussion & Conclusion

In this paper we outlined options to represent uncertainty of spatio-temporal event data in the CHC domain. To the best of our knowledge, hardly any applications for cultural heritage exist, with the exceptions of [KBD13, JW13] for TDU. Maybe this is due to the fact that "museums have practiced the concealment of uncertainty" [DK15, p.114]. Yet our exploration of CHC data showed that uncertainty poses a huge and omnipresent challenge in this domain: Meta-data highly vary in data quality, precision and provenance, and are frequently documented in an heterogeneous fashion. To make such uncertainty and data diversity transparent, visualization systems have to include uncertainty indicators. Otherwise they risk to omit objects by filtering their seemingly deficient entries, or to visualize them with a misleading degree of over-precision not actually present in the data.

To address this non-trivial visualization challenge in a more systematic manner, we outlined solutions for four different spatio-temporal visualization techniques, including CMV, ANI, CC, and STC representations. While exploring options of UV design for an advanced CHC interface including these four views, we also encountered a substantial challenge to the desideratum of consistent system design on a more general level [QH17]. According to well-established design guidelines, the additional complexity introduced by multiple views should be "balanced by ease of learning, which is facilitated by consistency" [BWK00, p.117]. Yet according to the best of our knowledge, no representation of SDU and TDU can be globally applied across all spatio-temporal perspectives without generating interferences of encoding choices – or less-than ideal effects on a local design level. We consider this hypothesis about necessary consistency compromises to point to a new kind of high-level design challenge, which requires further research.

While the implementation and comparative evaluation of the proposed UV techniques poses practical challenges for the near future, we see further challenges with specific regard to the management of visual clutter and complexity, which is inevitably raised by encoding uncertainty as an additional data dimension [RLN07]. Especially for interfaces conceptualized for casual use, smart and flexible solutions are needed to find dynamic trade-offs between the representation of uncertainty and the need to shield non-experts from an overkill of visual complexity. The aim should be to balance the quest for a more honest and critical visualization design (fostering transparency, awareness and trust [DFCC13, SSKK14]) with the strive for a sublime and inspirational (but not all too complex or confusing) experience, that oftentimes motivates the contemplation of cultural objects and artwork collections.

Acknowledgments
This work was supported by a grant from the Austrian Science Fund (FWF), No. P28363-24 and from the NFB, No. SC16-032.

References

[BDA*16] Bach B., Dragicevic P., Archambault D., Hurter C., Carpendale S.: A Descriptive Framework for Temporal Data Visualizations Based on Generalized Space-Time Cubes. In *Computer Graphics Forum* (2016), Wiley Online Library. URL: http://onlinelibrary.wiley.com/doi/10.1111/cgf.12804/full. 3

[BF12] Bennett K. B., Flach J. M.: Visual momentum redux. *International Journal of Human-Computer Studies 70*, 6 (2012), 399–414. URL: http://www.sciencedirect.com/science/article/pii/S1071581912000146. 4

[BWK00] Baldonado M. Q. W., Woodruff A., Kuchinsky A.: Guidelines for Using Multiple Views in Information Visualization. In *Proceedings of the Working Conference on Advanced Visual Interfaces* (New York, NY, USA, 2000), AVI '00, ACM, pp. 110–119. doi:10.1145/345513.345271. 3, 4

[DCW11] Dörk M., Carpendale S., Williamson C.: The information flaneur: A fresh look at information seeking. In *Proc. SIGCHI Conference on Human Factors in Computing Systems* (2011), ACM, pp. 1215–1224. 1, 2

[DDC*14] Delmelle E., Dony C., Casas I., Jia M., Tang W.: Visualizing the impact of space-time uncertainties on dengue fever patterns. *International Journal of Geographical Information Science 28*, 5 (May 2014), 1107–1127. doi:10.1080/13658816.2013.871285. 3, 4

[DFCC13] Dörk M., Feng P., Collins C., Carpendale S.: Critical InfoVis: Exploring the politics of visualization. In *CHI Extended Abstracts on Human Factors in Computing Systems (CHI EA)* (2013), ACM, pp. 2189–2198. 4

[DK15] Davis S. B., Kräutli F.: The idea and image of historical time: Interactions between design and digital humanities. *Visible Language 49*, 3 (2015), 101. 4

[DPC17] Dörk M., Pietsch C., Credico G.: One view is not enough. High-level visualizations of a large cultural collection. *Information Design Journal 23:1* (2017), 39–47. 1

[ET08] Elmqvist N., Tsigas P.: A taxonomy of 3d occlusion management for visualization. *IEEE Transactions on Visualization and Computer Graphics 14*, 5 (2008), 1095–1109. 4

[GBFM16] Gschwandtner T., Bögl M., Federico P., Miksch S.: Visual Encodings of Temporal Uncertainty: A Comparative User Study. *IEEE Transactions on Visualization and Computer Graphics Volume: 22*, Issue: 1 (Jan. 2016), 539 – 548. doi:10.1109/TVCG.2015.2467752. 2

[GJ10] Goodchild M. F., Janelle D. G.: Toward critical spatial thinking in the social sciences and humanities. *GeoJournal 75*, 1 (2010), 3–13. 2

[HAS11] Hullman J., Adar E., Shah P.: Benefitting infovis with visual difficulties. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (2011), 2213–2222. 2

[JFCS17] Jänicke S., Franzini G., Cheema M., Scheuermann G.: Visual text analysis in digital humanities. In *Computer Graphics Forum* (2017), vol. 36, Wiley Online Library, pp. 226–250. 1

[JW13] Jänicke S., Wrisley D. J.: Visualizing uncertainty: How to use the fuzzy data of 550 medieval texts. In *Proceedings of the Digital Humanities* (2013). 4

[KBD13] Kräutli F., Boyd Davis S.: Known unknowns: representing uncertainty in historical time. In *Proc. Electronic Visualisation and the Arts (EVA) 2013, London* (2013). 1, 2, 4

[KMB] Kardos J., Moore A., Benwell G. L.: The visualisation of uncertainty in spatially-referenced attribute data using trustworthy data structures. In *Proc. of the 15th Annual Colloquium of the Spatial Information Research Centre (SIRC 2003: Land, Place and Space), Dunedin, New Zealand*. 2

[KPS14] Kriglstein S., Pohl M., Smuc M.: Pep up your time machine: recommendations for the design of information visualizations of time-dependent data. In *Handbook of human centric visualization*. Springer, 2014, pp. 203–225. 3

[MA14] Miksch S., Aigner W.: A matter of time: Applying a data-users-tasks design triangle to visual analytics of time-oriented data. *Computers & Graphics 38* (2014), 286–290. URL: http://www.sciencedirect.com/science/article/pii/S0097849313001817. 1

[MFM*16] Mayr E., Federico P., Miksch S., Schreder G., Smuc M., Windhager F.: Visualization of Cultural Heritage Data for Casual Users. Collins C., (Ed.), IEEE VIS. URL: http://publik.tuwien.ac.at/files/PubDat_250950.pdf. 1

[MRH*05] MacEachren A. M., Robinson A., Hopper S., Gardner S., Murray R., Gahegan M., Hetzler E.: Visualizing geospatial information uncertainty: What we know and what we need to know. *Cartography and Geographic Information Science 32*, 3 (2005), 139–160. 2

[MRO*12] MacEachren A. M., Roth R. E., O'Brien J., Li B., Swingley D., Gahegan M.: Visual Semiotics amp;amp; Uncertainty Visualization: An Empirical Study. *IEEE Transactions on Visualization and Computer Graphics 18*, 12 (Dec. 2012), 2496–2505. doi:10.1109/TVCG.2012.279. 2

[QH17] Qu Z., Hullman J.: Keeping Multiple Views Consistent: Constraints, Validations, and Exceptions in Visualization Authoring. *IEEE transactions on visualization and computer graphics* (Aug. 2017). doi:10.1109/TVCG.2017.2744198. 3, 4

[RLN07] Rosenholtz R., Li Y., Nakano L.: Measuring visual clutter. *Journal of Vision 7*, 2 (2007), 17. 4

[SSK*16] Sacha D., Senaratne H., Kwon B. C., Ellis G., Keim D. A.: The Role of Uncertainty, Awareness, and Trust in Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 240–249. doi:10.1109/TVCG.2015.2467591. 2

[SSKK14] Sacha D., Senaratne H., Kwon B. C., Keim D. A.: Uncertainty propagation and trust building in visual analytics. In *IEEE VIS 2014* (2014). 4

[SWSM16] Schreder G., Windhager F., Smuc M., Mayr E.: A Mental Models Perspective on Designing Information Visualizations for Political Communication. *JeDEM - eJournal of eDemocracy and Open Government 8*, 3 (Dec. 2016), 80–99. URL: http://www.jedem.org/index.php/jedem/article/view/443. 3

[SZM14] Shrestha A., Zhu Y., Miller B.: Visualizing uncertainty in spatio-temporal data. *ACM SIGKDD Workshop on Interactive Data Exploration and Analytics* (2014), 117–126. 3

[VCHR13] Villa R., Clough P. D., Hall M. M., Rutter S. A.: Search or browse? casual information access to a cultural heritage collection. In *EuroHCIR* (2013), Citeseer, pp. 19–22. 2

[WFS*18] Windhager F., Federico P., Schreder G., Glinka K., Dörk M., Miksch S., Mayr E.: Visualization of Cultural Heritage Collection Data: State of the Art and Future Challenges. *Transactions of Visualization and Computer Graphics (TVCG), forthcoming* (2018). 1

[WH15] Walsh D., Hall M. M.: Just looking around: Supporting casual users initial encounters with Digital Cultural Heritage. In *Supporting Complex Search Tasks* (Vienna, 2015), vol. 1338, CEUR-WS. URL: http://ceur-ws.org/Vol-1338/paper_6.pdf. 1

[WMS*16] Windhager F., Mayr E., Schreder G., Smuc M., Federico P., Miksch S.: Reframing Cultural Heritage Collections in a Visualization Framework of Space-Time Cubes - Semantic Scholar. In *Proceedings of the 3rd HistoInformatics Workshop* (Krakow, June 2016), Düring M., (Ed.), vol. Vol-1632, CEUR-WS, pp. 20–24. URL: http://ceur-ws.org/Vol-1632/. 1, 3

[WSSM18] Windhager F., Salisu S., Schreder G., Mayr E.: Orchestrating Overviews. A Synoptic Approach to the Visualization of Cultural Collections. *Remaking Collections. Special Issue of the Open Library of the Humanities, forthcoming* (2018). 1, 3