Computational filling of curatorial gaps in a fine arts exhibition

Arthur Flexer

Institute of Computational Perception
Johannes Kepler University Linz, Austria
arthur.flexer@jku.at

Abstract

We present an interactive computational approach to fill curatorial gaps between artworks in a fine arts exhibition. We describe the algorithmic details of our semantic approach based on word embedding of keywords and how we include additional curatorial constraints. We present an installation at a museum exhibition and discuss lessons learned during our arts based research.

Introduction

This work is part of the arts-based research project "Dust and Data: the Art of Curating in the Age of Artificial Intelligence"[1], where we explore how machine learning can help both curators and audiences discover and navigate large museum collections. In a previous workshop contribution (Flexer 2020) we presented an approach to compute smooth semantic pathways between works of art. In this paper we report about an installation at our project exhibition "Dust and Data - Artificial Intelligence im Museum" at the Austrian Museum of Folk Life and Folk Art in Vienna, Austria, adapting and extending this approach. Our installation asks the question about curatorial gaps between artworks shown in an exhibition. What works of art exist in the holdings of the museum that fit the curatorial narrative but did not succeed in becoming part of the exhibition?

Our approach is inspired by the project "X Degrees of Separation"[2] by "Google Arts and Culture", which explores the "hidden paths through culture" by analyzing visual features of artworks to find pathways between any two artifacts through a chain of artworks. While we find these pathways aesthetically pleasing, we are, from a curatorial perspective, more interested in finding pathways of the semantic meaning of works of art. We chose this semantic driven approach because vital information about a piece of art cannot be found in the artwork itself. Think e.g. of subjecting the "Mona Lisa" to an automatic visual analysis. Computational results will tell you that it is a picture of a young woman, in front of a landscape, and (if your algorithm is really good) is sort of smiling. This information of course totally misses the significance of the painting for (Western) art history, its immense relevance and the many connotations it has. All of this rather

is a societal construct and result of centuries of discourse and reception history. Another problem with analysing the visual content of paintings is that state-of-the-art image analysis is usually trained on photographs and generalization to paintings is not trivial, especially for more abstract artforms (Kim et al. 2019).

As a consequence, we chose to use word embedding (Mikolov et al. 2013) to embed keywords of a museum collection and obtain pathways through the resulting semantic space as well as to add curatorial semantic constraints. Word embedding encodes semantic similarities between words by modelling the context to their neighboring words in a large training text corpus.

Data

We obtained 4365 artworks and their keywords from Belvedere's (Vienna, Austria) online collection[3], representing art from the middle ages to the present time. We exclude all sculptures and three-dimensional art, keeping 3421 artworks which are mainly paintings and drawings.

To demonstrate our curatorial approach of filling gaps, we chose part of a room in Belvedere's permanent exhibition. It is a room about "Viennese portraiture in the Biedermeier period", assembling the "greatest portrait painters" from this period. The three paintings marked (a), (d) and (g) in figure 1 are part of the original exhibition. Artworks in between ((b), (c), (e) and (f)) are proposed by our algorithm described in the next section.

Methods

Semantic embedding: The online collection is indexed with 3585 different keywords which are often very specialized with a fourth of them being assigned only once. Since this keyword index is therefore rather sparse and of only limited help for organization of the collection, we use natural language processing to compute similarities between keywords thereby obtaining a semantic embedding space. Specifically we use the tool `spacy`[4] to remove stop words and extract only nouns, this e.g. changes the original keyword 'scepter, ruling staff (as symbol of highest force)' into the four keywords 'scepter', 'ruling staff', 'symbol' and 'force'. Since

some of the artworks are not indexed by any keywords, we gain additional keywords by applying the same procedure to the titles of artworks also. This results in 6216 different keywords for which word embeddings actually exist. Please note that we translate all keywords from German to English for this paper. We use the German `fasttext`[5] word embedding, which has been trained on about 3 million words from the *Wikipedia-* and 19 million words from the *Common Crawl*-corpus (Mikolov et al. 2018). This gives us a vector representation of size 100 for every keyword, with the cosine between vectors indicating semantic similarity. A cosine of 1 signifies perfect semantic similarity and 0 no similarity at all. To obtain a similarity $c\bar{o}s(a,b)$ between any two artworks A_a and A_b with k_a and k_b keywords, we simply average all possible crosswise cosine distances between keyword lists.

Curatorial semantic constraint: Looking at the keywords of the three paintings from the original exhibition (marked (a), (d) and (g) in figure 1), one can see that most of them are purely descriptive, e.g. 'headgear', 'necklace', 'bonnet', 'eye contact', probably not doing the semantic content of the artworks full justice. We also believe that one underlying semantic topic of the Biedermeier room is 'gender', with all but one painting depicting females. We therefore add an additional algorithmic constraint by requiring all suggested artworks to respect both the requirement of being part of a pathway and having a 'gender' related keyword. Since 'gender' is not a keyword in the Belvedere taxonomy, we use word embedding to obtain Belvedere keywords with high similarity to the topic of 'gender'. This translation step yields the following top ranking keywords with cosine similarity between 0.60 and 0.45: *Islam, religion, headscarf, education, blinder, equal opportunities, religions, hacking, asylum, robe, female labor, femaleness, fan, skirt, medicine, ornament, force, avowal, psychiatry, delusion, blindness, doctrine*. There are a number of keywords in the same high similarity range which we excluded from this list for being too general: *non-, context, science, instrument, conversation, points of view, attribute, natural sciences*. There are 133 artworks with at least one of these keywords in the database of 3421 mostly paintings and drawings.

Choosing artworks: To compute artworks to fill the gap between a start artwork A_s and an end artwork A_e, we do the following for a database of n artworks A_i:

1. for all $i = 1, ..., n$ artworks compute cosine similarities to start artwork $c\bar{o}s(i,s)$ and end artwork $c\bar{o}s(i,e)$

2. find nn_s artworks with largest similarity $c\bar{o}s(i,s)$ to A_s; find nn_e artworks with largest similarity $c\bar{o}s(i,e)$ to A_e; keep only $m = |nn_s \cup nn_e|$ artworks for further processing

3. for all $i = 1, ..., m$ artworks compute a similarity ratio:

$$R(i) = c\bar{o}s(i,s)/c\bar{o}s(i,e) \qquad (1)$$

4. order all $i = 1, ..., m$ artworks according to their similarity ratio $R(i)$

Artworks which are closer to start artwork A_s than to end artwork A_e will have a similarity ratio $R(i) > 1$, while those

closer to A_e than to A_s will have $R(i) < 1$. Artworks which have equal similarity to both A_s and A_e will have a similarity ratio $R(i)$ around 1. Since this is also true for artworks with small but equal similarity, they have to be excluded from the pathway as outliers, keeping only the nn_s (nn_e) artworks closest to either A_s or A_e in step 2. For our revision of the Biedermeier room we chose $nn_s = nn_e = 30$. Values greater than 30 yielded artworks with too little similarity.

Results

In figure 1, the three paintings marked (a), (d) and (g) are part of the original exhibition, with the other four paintings being one specific revision solution obtained with the algorithm described above. These artworks at positions (b), (c), (e) and (f) fill the curatorial gaps by forming a transition between (a) and (d), and (d) and (g) respectively. At the same time each of the obtained artworks obeys the curatorial constraint of having at least one of the keywords from the 'gender' list. This constraint results in only 133 artworks from 3421 being eligible for the solution. In step 2 of our algorithm we keep the $|nn_s| = |nn_e| = 30$ closest artworks in consideration for our solution between artworks (a) and (d), and (d) and (g) respectively. Removing duplicates in these lists via $m = |nn_s \cup nn_e|$ leaves us with 18 artworks to fill the gap between (a) and (g), and 19 artworks between (d) and (g). From these lists of 18 or 19 artworks we can now randomly select two artworks each to fill the curatorial gaps. Since in step 4 of our algorithm these artworks are ordered according to their similarity ratios, every selection needs to respect this ordering to obey the requirement of a smooth transition. For $f = 18$ artworks, there are $(f(f-1))/2 = 153$ possibilities to choose two artworks respecting the ordering. For $f = 19$ artworks there are 171 possibilities.

Returning to figure 1, we would now like to discuss this exemplary solution out of the many possible ones. Painting (b) is suggested because its keyword 'femaleness' (in bold face in figure 1) is a gender keyword and its keyword 'necklace' makes it similar to the keywords of painting (a) ('earrings', 'pearl necklace') and to painting (d) ('brooch', 'bracelet'). Keyword 'portrait of a girl' is also related to 'portrait' of painting (a). All these similarities together explain why painting (b) fits into the transition between (a) and (d) while at the same time being related to the concept of 'gender'.

Similar arguments can be given for the other filling artworks. Painting (c) has the gender keyword 'fan' which also appears for (a). It also has the keyword 'inner room' (and 'church interior') which also appears for (d). Painting (e) has the gender keyword 'force', but also 'empire' and 'emperor' which all relate to keyword 'princess' of (d). Keyword 'eye contact' appears also for (d) and (g). Painting (f) has the gender keyword 'head scarf' which connects to the keyword 'feather hat' of (d). Keyword 'eye contact' appears for (d), (f) and (g). The keyword 'spouse' also connects to 'woman' and 'lady' of painting (g).

Leaving the discussion about the one specific solution in figure 1, it is interesting to see which gender keywords appear for all 18 plus 19 possible transition artworks. Keyword 'head scarf' appears 18 times, 'fan' 8 times, 'force' 5 times, 'skirt' 3 times, 'femaleness', 'religion', 'ornament' and 'avowal'

(a) headgear, chair, portrait, gloves, figure, earrings, person, pearl, armchair, fan, pearl necklace, mitten, lady

(b) **femaleness**, necklace, portrait of a girl

(c) veil, church, figure, **fan**, inner room, church interior

(d) Marie, feather hat, eye contact, clasp, curtain, bracelet, brooch, countess, bangle, figure, primary motif, draperies, princess, ring, inner room, buckle, Chorinsky, Esterházy

(e) **force**, empire, eye contact, gown, scepter, Franz, holy, emperor, symbol, cuirass, Lothringen, breastplate, ruling staff, coat, Stephan, garb, fleece

(f) eye contact, **head scarf**, Schrotzber, figure, Eleonore, artist's, spouse, Stohl

(g) eye contact, half profile, portrait, person, bonnet, lace hood, woman, inner room, ribbons, lady

Figure 1: Depicted is one exemplary solution of a revision of the Biedermeier room. The three paintings marked (a), (d) and (g) are part of the original exhibition. Artworks in between ((b), (c), (e) and (f)) are proposed by our algorithm. Captions show all respective keywords, with gender related keywords printed in bold. All images by Belvedere, Vienna, Austria (CC BY-SA 4.0, https://creativecommons.org/licenses/by-sa/4.0/).

once, all others never. It is noteworthy that the majority of keywords ('head scarf', 'fan', 'skirt') describe apparel or accessories. These keywords are of course closer to the many keywords of the original Biedermeier paintings (a), (d) and (g) also describing apparel or accessories. It would be interesting to keep the curatorial constraint of requiring a 'gender' keyword for every filling artwork, but excluding these keywords from the computation of similarity between artworks. This could result in gender related artworks more detached from keywords describing apparel.

Somewhat problematic are keywords describing pictorial organization (Bove, Heusinger, and Kailus 2016) like 'multiple layer room', since the word embedding is not able to grasp such subtle semantics. Names of historical persons as keywords (e.g. 'Chorinsky' or 'Esterházy') are often not adequately embedded in the semantic space and very general terms like 'men' or 'figure' can also lead the algorithm astray.

Returning to the full list of gender related keywords which we obtained via word embedding, it is also striking that many keywords point to a stereotypical discourse of gender, quickly derailing towards topics of 'religion' and 'Islam' and a compulsion to wear 'headscarfs', or a discussion of 'femaleness' and 'force', probably pointing to women still being subjected to violence in today's society. This is also why we like to term the use of word embedding in this context world embedding: it confronts the very rigid taxonomy of the Belvedere keywords (based on Iconclass[6], a multilingual classification system for cultural content) with everyday language as represented in the textual training data of the word embedding. It thereby re-contextualizes or even re-socializes taxonomic

art histories via natural language processing since it uncovers biases and prejudice in our use of language and (re-?) introduces them to the world of fine arts.

The actual installation of the Biedermeier room revision was part of our project exhibition "Dust and Data - Artificial Intelligence im Museum" at the Austrian Museum of Folk Life and Folk Art in Vienna, Austria. The installation is a half-scale copy of part of Belvedere's Biedermeier room and can be seen in figure 2. The three original paintings (positions (a), (d) and (g) in figure 1) are shown as reproductions printed on linen to set them apart from the artworks at the curatorial gap positions. Artworks selected by our algorithm are projected at their respective positions including keyword information. Please note that to keep keyword information readable, only single artworks are projected between (a) and (d), and (d) and (g) respectively. The artworks are shown in a repeated random order, realizing a sort of flickering representing the many artworks that fit the curatorial gaps but have not been shown by the original curation.

Conclusion

We have presented an approach to compute pathways between works of art that also follow an overarching curatorial constraint, enabling audiences to discover transitions based on semantic content instead of visual information. There are three lessons we have learned while building the art installation based on this technology described in our paper:

(i) Generally speaking, **semantic** approaches should be more helpful for **building a curatorial narrative** (Wolff, Mulholland, and Collins 2012) than a purely aesthetic procedure. After all, museum curation relying on visual information only is hard to conceive. Our specific procedure allows to

---

[6]http://www.iconclass.org/

Figure 2: Installation of the revised Biedermeier room at the Austrian Museum of Folk Life and Folk Art in Vienna, Austria. Photo by Christoph Panzer.

answer the question about curatorial gaps between artworks shown in an already existing exhibition.

(ii) Using a machine learning tool like word embedding, **curating becomes a joint endeavor of man and machine**, where curatorial decisions have to be formulated as input and constraints to the algorithm. As such we tried "to design programs that can enhance human creativity without necessarily being creative themselves"[7], which is one of the goals of computational creativity. A fact which is hardly ever discussed is that even a simple curatorial Google search already is an interaction of man and machine, with algorithms to a certain extent (oblique to the curator) shaping their curatorial enterprise by showing specific selections of information only. All these man/machine approaches are able to uncover algorithmic biases in the methods used, as e.g. stereotypical representations of societal discourse in word embedding.

(iii) Looking towards **future extensions** of our work it can be said that of course we could analyse longer (art historic) texts about artworks with the same methodology thereby gaining much richer semantic context then by relying on simple keywords only. Another possible extension is to embed semantic and visual information simultaneously which could yield curatorial solutions that respect semantic and visual constraints at the same time (Frome et al. 2013; Kim et al. 2019).

Acknowledgments

---

[7]http://computationalcreativity.net/home/about/computational-creativity/

References

Bove, J.; Heusinger, L.; and Kailus, A. 2016. *Marburger Informations-, Dokumentations-und Administrations-System (MIDAS).: Handbuch und CD. hand, 4, cd, onl*. Universitätsbibliothek Heidelberg.

Flexer, A. 2020. Discovering x degrees of keyword separation in a fine arts collection. In *Proceedings of the 37th International Conference on Machine Learning, Machine Learning for Media Discovery Workshop, PMLR 108*.

Frome, A.; Corrado, G. S.; Shlens, J.; Bengio, S.; Dean, J.; Ranzato, M.; and Mikolov, T. 2013. Devise: A deep visual-semantic embedding model. In *Advances in neural information processing systems*, 2121–2129.

Kim, D.; Xu, J.; Elgammal, A.; and Mazzone, M. 2019. Computational analysis of content in fine art paintings. In *Proceedings of the 10th International Conference on Computational Creativity*.

Mikolov, T.; Chen, K.; Corrado, G.; and Dean, J. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Mikolov, T.; Grave, E.; Bojanowski, P.; Puhrsch, C.; and Joulin, A. 2018. Advances in pre-training distributed word representations. In *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2018)*.

Wolff, A.; Mulholland, P.; and Collins, T. 2012. Storyspace: a story-driven approach for creating museum narratives. In *Proceedings of the 23rd ACM conference on Hypertext and social media*, 89–98.