Towards Computational Art Curation: Re-imagining the city of Helsinki in occasion of its Biennial

Dario Negueruela, Ludovica Schaerf, J0sé Ballesteros, Iacopo Neri, Valentine Bernasconi

Figure 1: Example of a 360° art panorama. A real location from Helsinki is transformed according to the artistic style of its corresponding artwork.

ABSTRACT

Art curatorial practice is characterized by the presentation of an art collection in a knowledgeable way. Machine processes are characterized by their capacity to manage and analyze large amounts of data. This paper envisages machine curation and audience interaction to explore the implications of contemporary machine learning models for the curatorial world. This project was developed for the occasion of the 2023 Helsinki Art Biennial, entitled New Directions May Emerge. We use the Helsinki Art Museum (HAM) collection to re-imagine the city of Helsinki through the lens of machine perception. We use visual-textual models to place indoor artworks in public spaces, assigning fictional coordinates based on similarity scores. We transform the space that each artwork inhabits in the city by generating synthetic 360° art panoramas. We guide the generation estimating depth values from 360° panoramas at each artwork location, and machine-generated prompts of the artworks. The result of this project will be virtually presented as a web-based installation, where users can navigate an alternative version of the city while exploring and interacting with its cultural heritage at scale.

CCS CONCEPTS

• **Applied computing** → **Arts and humanities**; **Media arts**.

KEYWORDS

Helsinki Biennial, AI curation, machine perception

ACM Reference Format:

Negueruela del Castillo, D. Schaerf, L. Ballesteros, P. Neri, I. Bernasconi. V. 2023. Towards Computational Art Curation: Re-imagining the city of Helsinki in occasion of its Biennial. In *Proceedings of ACM International Conference on Multimedia (Conference ACM MM '23).* ACM, New York, NY, USA, 7 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

In this paper, we present the first steps of a curatorial work that will be exhibited on the occasion of the 2023 Helsinki Biennial of Art. The project uses Artificial Intelligence (AI) as a new means for curatorial practice, exploring the possibilities and difficulties that such new methods introduce. The proposed work in this paper

Conference ACM MM '23, October 29 – November 02, 2023, Ottawa, Canada

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00

https://doi.org/XXXXXXX.XXXXXXX

is part of a more significant project entitled *New Directions May Emerge*[1], which aims to curate a museum's art collection through the perception of the machine. The immediate product of this curation consists of an interactive website based on a modified Voronoi algorithm[2] that transforms the shape of Helsinki using the virtual dots of the collection locations. The website will dynamically update throughout the Biennial, with new locations (and therefore artworks) appearing and modifying the shape of the city every 30 minutes. The physical space of this virtual environment is undetermined, hence this paper discusses some of the possibilities for projection and the engagement they entail.

The work is based on the art collection from the Helsinki Art Museum (HAM), consisting of public artworks, such as sculptures and art installations, as well as an indoor collection. Using the city of Helsinki as a context, the goal is to present a different experience of the works of art through the navigation of a new projection of the artworks in the city. First, we situate the public artworks of the HAM collection in their original place in the city and assign the artworks from the indoor museum collection to a fictional place in Helsinki, adopting deep learning and machine learning tools. We then extract the 360° Google Maps views of all the real and fictional locations, which are later used as guiding depth maps to insert the artworks in their fictional or real surrounding space. This is achieved using Stable Diffusion [24] on the 360° views that they inhabit. The city is thereafter populated by the new machinic world, where the user can navigate a geography that blurs the world of extant reality and that of machinic fiction. Following the narrative thread proposed in [14], the project focuses on the following questions: How can we curate a collection we have never seen? How does the machine perceive art? Can the machine offer a fruitful re-contextualization of the artistic data? Can the geography of the city offer fruitful ground for this re-contextualization?

1.1 AI as a Curator

Traditionally, the work of a museum curator consists of the enrichment of a collection and cultural heritage preservation. With the creation of an exhibition, the curator performs a selection of the collection according to a narrative thread that has to be passed to the public [10]. The goal is to generate new insights into the original works of art, and elevate their physical dimension through the design of their display [4]. In recent years, with the increased availability of digital collections and tools, the notion of digital curation has become an important aspect [21], especially facing the large amount of digital data generated and their online publication to reach a wider audience. Computational art curation aims at classifying and indexing data for efficient retrieval [21], as well as creating new experiences of the artworks through new technologies [9] (e.g., virtual and augmented reality).

Unsurprisingly, the use of AI systems for artistic curation has found fertile ground, spanning from projects by Google Arts & Culture to on-site curations of museums and biennials. Google Arts & Culture's[3] experiments are pioneering applications of computational methods to the online curation of artistic datasets. For example, their 't-SNE Map' and 'Curator Table' experiments are visualization tools to see how objects, styles and artists evolve over time. Moreover, the project 'X degrees of separation' was a source of inspiration for the AI-based curatorial project *Dust and Data: The Art of Curating in the Age of Artificial Intelligence* [7].

In fact, *Dust and Data* (DAD) explores the possibilities of AI curation as an assistant to both curators and audiences; it uses semantic embeddings to recreate a curatorially specific version of the Google Arts experiment, which proposes a chain of artworks that connect one work to another, in this sense, filling the curatorial gaps in art collections [6, 7].

Another remarkable example of machine curation is the previous edition of a Biennial, the Liverpool Biennial 2021, titled *The Next Biennial Should be Curated by a Machine*[4]. The curation allows navigating the Liverpool collection through a set of alien images. For each artwork, GAN-generated images were created from the titles of existing artworks in the collections. Moreover, CLIP [22] was used to extract keywords from the artworks that were used as the link to navigate the collection.

AI curation aims to offer new insights into digital cultural artifacts. It is possible to propose personalized journeys of the collections, as well as to foster creative takes on its presentation [9]. Finally, contemporary AI curation strives to disentangle the undercovered behaviors of large models by switching from the practical data and tasks they were trained on to curatorial and artistic purposes.

A crucial element relevant to ethical considerations of AI curation is the possibility of offering a gaze on any art collection that would be free from a specific cultural framing. Indeed, as pointed out by Jones [12], there has been over the past four decades an increasing criticism of the way cultural objects that do not belong to our Western culture are treated. These objects are too often perceived as primitive artifacts, which derive from a colonial projection on non-western societies [1, 12]. The approach of the museum curator towards the objects and the narrative presented in an exhibition impacts and influences the perception of the public, and solutions have to be found to overcome this biased gaze based on the origin of artworks, and to open to "*alternative voices, histories, and representations*" [12]. Nevertheless, considering the material used in the training of most deep learning models and its strong Western anchoring, AI cannot be considered in itself as the solution but as another possibility for experiments toward cultural diversity and postcolonial views.

1.2 The HAM Dataset

The Helsinki Art Museum (HAM)'s collection consists of the core material used for the project. Defining itself as "*a city-wide art museum*" the HAM holds about 10'000 artworks, of which around 2'500 can be found in the outdoor and indoor public spaces of the city. These artworks are very diverse, such as sculptures, paintings, and drawings. The idea behind the project is to take advantage of that urban perspective and experiment with the original locations of the public works. To this end, we get access to the geographical information and the corresponding photographs of the 488 outdoor public artworks. The information consists of a set of longitude and

[1] See https://helsinkibiennaali.fi/en/story/helsinki-biennial-2023-brings-together-29-artists-and-collectives/.
[2] See https://en.wikipedia.org/wiki/Voronoi_diagram
[3] See https://experiments.withgoogle.com/collection/arts-culture.
[4] See https://ai.biennial.com/.

latitude coordinates. Additionally, 1'744 items from their indoor collection are harvested from their online platform[5]. For each item, a corresponding image representing the artwork is retrieved, as well as the title, date of creation, name of the artist, keywords in English, Finnish, and Swedish describing the piece, and the object ID in the official collection.

We thus collect a total of 2'232 items divided into two distinctive sets referred to as the public art, corresponding to the outdoor public artworks, and the indoor art, referring to the indoor collection of the HAM.

2 METHODS

In this project, we strive to present and recreate the HAM collection as a new entity inhabiting and embodying the city of Helsinki. To this end, our curatorial pipeline begins with the geolocation of all the artworks of the collection, including the indoor artworks that do not have a physical location. We proceed in two steps: we employ an image-to-text model to extract a compressed representation of both the public and indoor art, and we successively utilize this representation to assign fictional coordinates to the indoor collection based on their similarity to the public artworks. Given the new coordinates, we proceed to induce the artworks to embody their space at that coordinate: we extract the panoramic 360° view of each artwork from its corresponding location and use diffusion-based models [24] to turn the 360° panoramas into an immersive space representing the artwork. The output image is generated using depth images of extracted panoramas and machine-generated prompts as input guidance for the model (Figure 6).

2.1 Image to CLIP representations

"A painting of a person standing in front of a body of water, by Cornelia MacIntyre Foley, persian folklore illustration, river and trees and hills, cd cover artwork, protagonist in foreground, wanderers traveling from afar, in a desert oasis lake, watercolour, auction catalogue photo, inspired by Janet Fish"

Figure 2: Example artwork with CLIP Interrogator extracted text. Artwork: Lepistö, P. "Kolme vesilintua metsälammella". Courtesy of the HAM collection.

As a first step, we extract the visual and textual features from all images in the collection using the CLIP-based model CLIP-Interrogator [22] (Figure 2). In fact, we take advantage of the zero-shot performance of the CLIP model released by OpenAI, which produces Stable Diffusion 1.5 compatible prompts. Using CLIP-Interrogator, we store two outputs: the prompts $\mathbf{t}$ and the embeddings $\mathbf{z}^*$. For each image $\mathbf{x} \in \mathcal{X}$, the interrogator maps x to the image embedding $\mathbf{z}_I \in \mathcal{Z}_I$ using the `ViT-L-14` model. It leverages contrastive-based learned features to map image embeddings $\mathbf{z}_I$ to text embeddings $\mathbf{z}_T \in \mathcal{Z}_T$, where $\mathbf{z}_I, \mathbf{z}_T \in \mathbb{R}^m$, $m = 768$. Each text embedding $\mathbf{z}_T$ is decoded into a text prompt $\mathbf{t} \in \mathcal{T}$. The prompts will be used in Section 2.4 as the inputs for the Stable Diffusion generation. We wish to consider both linguistic and visual information to assign the fictional coordinates in Section 2.2. Therefore, we represent each artwork (both indoor and public) as the concatenation $\mathbf{z}^* \in \mathcal{Z}^*$ of $\mathbf{z}_I$ and $\mathbf{z}_T$, where $\mathbf{z}^* \in \mathbb{R}^{m+m}$.

2.2 CLIP to Fictional Coordinates

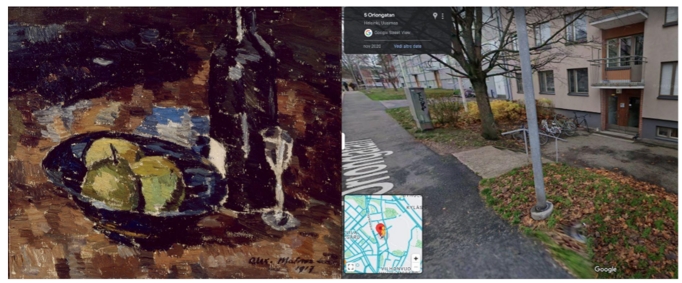

Figure 3: Example indoor artwork (left) with predicted location (right). Artwork: Matson, A. "Asetelma". Courtesy of the HAM collection.

Successively, we determine fictional coordinates (Figure 3) for the 1'744 images of the indoor collection using information from the known geolocations (latitudes and longitudes) of the public artworks $\mathbf{y}_{\mathbf{z}^*_{public}} \in \mathcal{Y}$ and the feature vectors $\mathbf{z}^*$ obtained in the previous step. We experiment with several predictive and similarity-based algorithms. We split the public collection data $\mathbf{z}^*_{public} \in \mathcal{Z}^*$, $\mathbf{y}_{\mathbf{z}^*_{public}} \in \mathcal{Y}$ into 70% training and 30% validation, and predict on $\mathbf{z}^*_{indoor} \in \mathcal{Z}^*$ to obtain the fictional locations of the indoor artworks $\hat{\mathbf{y}}_{\mathbf{z}^*_{indoor}} \in \mathcal{Y}$.

As a first effort, we train a selection of canonical machine learning regression models, Linear Regression [8], decision tree-based methods such as Random Forest [11], XGBoost [2], and AdaBoost [26], a Support Vector Machine Regression (SVR)[6] [3], and a simple 5-layer Feed Forward Neural Network to predict $\mathbf{y}$ using $\mathbf{z}^*$ and test different preprocessing techniques for dimensionality reduction and normalization. We explore UMAP [18] and PCA [20] methods for dimensionality reduction, to 200, 50, and 2 dimensions each and we normalize the data using Standard Normal or Min Max scalers.

Alternatively, we also experiment with an unsupervised GPS-inspired similarity method to compute the coordinates of the indoor artworks $\hat{\mathbf{y}}_{\mathbf{z}^*_{indoor}}$. We use the feature vectors $\mathbf{z}^*_{indoor}$ to find the three most similar public artworks. Practically, we construct a Ball Tree [5] using $\mathbf{z}^*_{public}$ and we query the tree with each $\mathbf{z}^*_{i\,indoor}$. We retrieve, for each indoor artwork, the three items of the public

[5]See https://ham.finna.fi/?lng=en-gb for the full collection.

[6]All the models used are available from sklearn, we use the default hyperparameters and do not perform any hyperparameter tuning.

artworks $\mathbf{z}^*_{j\,public}$ where $j = 1, 2, 3$ with the smallest Euclidean distance d as:

$$\mathbf{z}^*_{1\,public} = \underset{i \in indoor, k \in public}{\arg\min} \; d(\mathbf{z}^*_k, \mathbf{z}^*_i)$$

$$\mathbf{z}^*_{2\,public} = \underset{i \in indoor, k \in public}{\arg\min} \; d(\mathbf{z}^*_k | \mathbf{z}_1, \mathbf{z}^*_i)$$

$$\mathbf{z}^*_{3\,public} = \underset{i \in indoor, k \in public}{\arg\min} \; d(\mathbf{z}^*_k | (\mathbf{z}_1, \mathbf{z}_2), \mathbf{z}^*_i)$$

Finally, we use the coordinates $\mathbf{y}_{\mathbf{z}^*_{j\,public}}$ of the three most similar public artworks $\mathbf{z}^*_{j\,public}$, $j = 1, 2, 3$ to triangulate the fictional coordinate of the indoor artwork as the centroid of the triangle simply as:

$$\hat{\mathbf{y}}_{\mathbf{z}^*_{i\,indoor}} = \frac{\mathbf{y}_{\mathbf{z}^*_{1\,public}} + \mathbf{y}_{\mathbf{z}^*_{2\,public}} + \mathbf{y}_{\mathbf{z}^*_{3\,public}}}{3}$$

2.3 Fictional and Real Coordinates to Panoramas

Once the fictional location of the indoor artworks $\hat{\mathbf{y}}_{\mathbf{z}^*_{\text{indoor}}}$ is calculated, we begin with the inspection of the local conditions of each data point, the 360° panorama street view. We employ the Google Street View API to gather the panorama street views $\mathbf{v}$ at each latitude and longitude tuple (both $\mathbf{y}_{\mathbf{z}^*_{\text{public}}}$ and $\hat{\mathbf{y}}_{\mathbf{z}^*_{\text{indoor}}}$).

Helsinki offers a very varied landscape, spanning from coastal settings and gulfs to urban areas and parks. On the one hand, this variation provides fertile ground for the following image generation phase. On the other, the pipeline is challenged by several locations where the 360° street view is unavailable. To overcome this limitation, an iterative process queries the Google Street View API with increasing radii to ensure proximity to the predicted position, while permitting local adjustments. In the case of locations with no available street view panorama within a radius of 250 meters, such as in the middle of the sea, a 360° panorama view with an aspect ratio of 19:6 is generated using Midjourney[7].

2.4 Panoramas and CLIP prompts to Art Panoramas

Finally, using the panorama views $\mathbf{v}$ of each location as depth maps, and prompts $\mathbf{t}$, we generate landscape artworks that semantically depict the original art piece but use the real context as the canvas. To this end, ControlNet[8] [28] plays a key role in guiding the generation with an input depth map, computed via MiDaS from $\mathbf{v}$ [23]. Through their combination - assisted with asymmetric tiling[9] - we influence the Stable Diffusion[10] generation towards pertaining visual consistency between the real and the imagined landscapes. Finally, the resolution of the artwork is increased by 2x using ESRGAN [27], leading to the resulting art panoramas $\mathbf{a}$ (Figure 6).

[7]See midjourney.com.
[8]We use the code from the official release on Github, v1.0.
[9]See https://github.com/tjm35/asymmetric-tiling-sd-webui/.
[10]We use the code from the huggingface release v1.5, using 30 inference steps and Euler sampling.

3 RESULTS

In this section, we describe our quantitative and qualitative results. Quantitative results are those related to the computation of fictional coordinates and panoramas. Qualitative results relate to image similarity retrieval and the generation of art panoramas.

3.1 Fictional coordinates

As we wish to understand the amount of variation captured by the model, we use the R^2 on the test set in order to evaluate the performance of the models in Section 2.2. We find the best-performing model to be the Random Forest model with no dimensionality reduction and Min Max scaler, which obtained an R^2 of 0.091. This is very close to a regression to the mean, which has an R^2 of 0, indicating that the models are not able to capture the complexity of this task only relying on machine visual and textual descriptions of artworks to extrapolate coordinates (Figure 4).

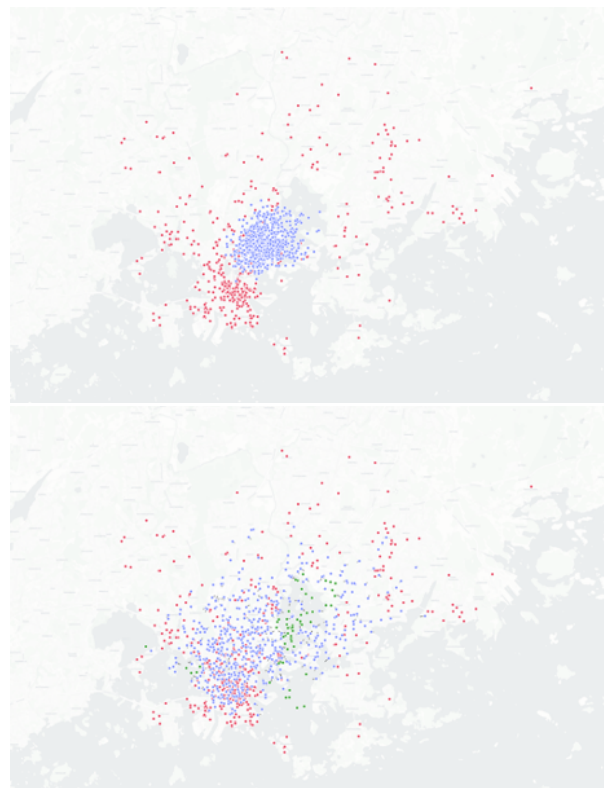

Figure 4: Top: Map of Helsinki showing the public artworks $\mathbf{y}_{\mathbf{z}^*_{public}}$ in red, and the predicted locations $\hat{\mathbf{y}}_{\mathbf{z}^*_{indoor}}$ in blue. Clustered blue points illustrate how the Random Forest model does not capture variation in the data. Bottom: Updated map with the GPS-inspired similarity method. Green points indicate sea locations that did not have a 360° panoramic image.

This result is not surprising, as the reasons behind the allocation of public artworks throughout a city may involve a wide range of cultural, societal, and urban factors that are not considered in the study.

As an alternative method to overcome such low prediction variation, we test a GPS-inspired similarity method. As there are no ground truth locations for the indoor artworks, we are not interested in correctly predicting locations, but rather, in a meaningful way of representing the locations through the eyes of the machine. In fact, we aim to recreate the space of the city, an alternative Helsinki through the lens of machine perception. The unsupervised method is justified exactly by this premise, that the space should follow a semantic continuity rather than a strictly geographical one. Furthermore, the triangulation is inspired by GPS localization, reconnecting symbolically to the bases of geolocation.

This approach results in a roughly even and well-distributed set of locations, which are easy to navigate and distinctive from each other (Figure 4). Therefore, we select GPS-inspired similarity as the most suitable method to generate fictional locations. We obtain a dataset of 1'744 predicted locations of which we are able to retrieve 1'681 equirectangular street view images within a range of 250 meters. The remaining 3.61% are assigned the set of generated HDR (High Dynamic Range) images of forest (14%) and sea (86%) landscapes.

3.2 Art Panoramas

Figure 5: Example of the top three most similar public artworks (right) to an indoor painting sample (left). Artworks (from left to right top to bottom): Sallinen, T. "Hihhulit"; Juva, K. "Arkkienkeli Mikael"; Sörensen-Ringi, H. "Jäähyväiset"; Kaasinen, T. "Ihmisiä". Courtesy of the HAM collection.

Due to the nature of the project, we examine the rest of the results qualitatively. First, after inspecting the three most similar images from a randomly selected set of indoor artworks, we draw some general impressions on the nature of such similarity. The retrieved images seem to have both conceptual and visual similarities to the original image (Figure 5), where the concept of religion is used in the retrieval of the first image, even in the absence of any physical connection. In some cases, similarities are found across modalities, with instances of connections between paintings and statues. For example, formalistic properties are shared between query and retrieved images (e.g., a drawing of a square triggers cubic-shaped sculptures in retrieval). On the other hand, some public artworks are retrieved repeatedly, sometimes without any clear connection, indicating that the search space is not equally likely for all images. The limited conceptual connections are unsurprising as CLIP prompts function as textual descriptions of the visual, and not as an art historical explanation of the artwork. Moreover, we note that CLIP-Interrogator only adopts a limited vocabulary in the extraction of the textual description, which is not suited for art historical descriptions.

Figure 6: Images involved in the immersive panorama generation of Figure 2 (left). Panorama of predicted location (top right), Depth map (middle right), art panorama using depth map and CLIP prompt (bottom right). Artwork: Lepistö, P. "Kolme vesilintua metsälammella". Courtesy of the HAM collection.

With the generation of 360° art panoramas using depth maps as a guide, we aim to maintain the spatial geometry of the image close to that of the original street views. While in most cases that geometry is kept and the perceived resulting space still captures the 3-dimensional quality of the original urban setting (Figure 6), the results show that the built physical features can become highly transformed depending on the graphical and pictorial style of the source artwork. For instance, very atmospheric pictorial styles, where contours are very blurred and diffuse, tend to generate resulting 360° art panoramas with less recognizable geometries of specific physical features of the original urban elements, while keeping the same perception of depth and overall composition. We see that most of the panoramas broadly reflect high coherence with respect to the semantics but significant shifts in color palette and brushstroke. Concretely, we highlight how the textual information used to generate art panoramas are often insufficient for an appropriate match of color palette and style properties of artworks. The perceived space changes with the artistic style of the original artwork, rendering a new dreamed Helsinki as seen through the collection by the proposed pipeline. The immersion of the resulting 360° art panoramas is generally satisfactory as tested through a generic online 360° panorama viewer VR[11]. It is to be noted that those panoramic views are experienced through a fixed standpoint and do not each result in a synthetic 3D navigable space. Nevertheless, using 360° panoramas allows us to capture a complete spherical view of the image surroundings, hence leading users towards a credible immersive experience.

4 DISCUSSION

Previous sections have presented the first steps of curatorial work for the 2023 Helsinki Biennial. The process of creating a new spatial projection of artworks from the HAM collection entails ethical and scholarly issues. These are part of a larger narrative that continuously unfolds and is focused on the ways creative AI applications (en)force global cultures.

4.1 Towards a curatorial machine

The stack of models presented works as a curatorial agent, not simply following another sophisticated search engine paradigm, but as a mediating actor. In modern society, we are already witnessing machine curatorial processes that guide cultural aesthetic preferences [17] (e.g., recommender systems), but this example showcases and highlights new emergent practices that point to a shift in the modalities involved in cultural curation and artistic production. Particularly, agency comes into question in this work, a modular stacking of several algorithms from a variety of tasks and applications. The complexity of such products makes it unsurprising that both scholars and the broader public are ready to project authorship to these models as we have recently seen in debates about ChatGPT and GPT3-4 [15, 25, 29]. In our case, the capacities of a curatorial agent which relate to the possibility of connecting images and texts to a common embedding space, are inherited from CLIP. Furthermore, the capacity to establish translations between the visual, textual, and spatial, situates the imagined collection in a new geography that is neither unreal nor illusory as it is shared and can be experienced, much like curation. Curation is no longer reserved for humans, just as this curatorial practice transgresses the domain-specific knowledge of the art world. As machine curation develops and normalizes, understanding the inner functioning of machine learning models becomes an increasingly crucial literacy, relevant to a general set of skills in contemporary artistic curation.

4.2 Avenues for public engagement

As this curation is not a physical exhibition, it can inhabit endless physical spaces, each creating a diverse public engagement. The first possibility of physical interaction is inside the walls of the HAM. This can take the form, among others, of a 2D projection screen, a 3D immersive curved screen, or a Virtual Reality headset. Due to the nature of the panoramas, we believe the best interaction would be achieved through a cylindrical screen or a dome, which would immerse the public in a 360 degrees setting of the surrounded mutated city, and, unlike the VR headset, this projection is suited for the fixed central standpoint of the produced panoramas. The geographical nature of the project opens a second avenue for physical interaction outside the museum space: the street. Here, we believe an interesting interaction can take place as a mobile AR experience.

[11] See renderstuff.com.

The public would engage with the machinic Helsinki at the location of the real Helsinki, superimposing the two worlds simultaneously.

These avenues lead us to ask what the effects of this newly imagined urban landscape on the "real" Helsinki are, and what interactions it will unleash once deployed. The impact of digital versions of a given environment, especially urban, is being discussed within the digital twin and smart city scene [16], but the effects of such an imagined digital version on the behaviors, attitudes, urban development, and other aspects of the social and built urban environment are yet to be assessed.

In addition, the potential effects of such an experience in fostering enhanced levels of spatial agency are also to be considered and assessed. Aesthetic experiences can be conducive to emotional reactions, which in turn alter not only the way we perceive the space around us but also can modulate our perception of affordances and therefore, our spatial capacities [19]. Our project seeks to revisit the urban spaces of Helsinki through the imagined panoramas, inviting the public to engage differently with their urban imaginary. An AR implementation that locally reacts to the location of the visitors, either through geolocation estimation via GPS triangulation or through on-site QR code scanning can add a crucial component, resulting in a more convincing situated experience and therefore, a more impactful emotional and affective engagement.

4.3 Ethical considerations

In our globalized world, an art biennial is an event designed to showcase and locate a city on the map of current creative and influential cities worldwide. It is, therefore, about status and attracting attention, explicitly foregrounding the added value of the city's assets, innovative profile, and capacity to become a central player [13]. In the case of this project, we were commissioned to work with and feature the collections of the HAM collection, which is primarily composed of artworks from local artists, potentially for a global audience. We tackled such a conundrum with the clear objective of avoiding manipulating the original artworks in order to respect their artistic integrity and their rich and complex relationship with the many layers of memory, identity, and local cultural heritage.

The ethical stakes of such an operation need to be carefully delineated. On the one hand, there is a need to respect, as outlined above, the artwork regarding its cultural context, and, on the other hand, we need to consider the approach from a global public, necessarily agnostic of those specificities. This second approach requires some freedom to appropriate, recombine, and re-imagine cultural production in the process of international artistic influence and cross-fertilization.

In this conundrum, the machine (meaning the actant resulting from the stack of diverse models) becomes an aid. It allows a certain lecture of these images. However, the fact that we are using a CLIP-guided model, and getting the textual and visual embeddings from CLIP means that we are, in practice, reading a collection (the one of HAM) from another collection (the CLIP training dataset). What that exactly means remains a complex and multifaceted question, but in our case, it becomes a crucial aspect that problematizes the desired cultural agnostic approach as made through the machine.

It confirms that cultural framing is already embedded in our computational models and begs the question of how to conceive an AI curation that embraces diverse cultural frames and a non-colonial approach.

5 CONCLUSION

In this paper, we presented a novel digital curation of a collection of works of art carried out for the 2023 Helsinki Biennial. Because of our will to propose a culturally agnostic view on artistic material, we decided to represent the collection through the lens of deep learning models, proposing a machinic approach to art curation, and anchoring it in the city of Helsinki. Using the Helsinki Art Museum (HAM) collection, we considered the use of the CLIP-Interrogator to create textual descriptions of the works of art and assign them fictional coordinates around the city of Helsinki through a similarity-based algorithm. A new synthetic 360° panoramic view of the predicted location was then generated with the original depth map of the location and the CLIP prompt, thus proposing a new visual style or flavor of the city of Helsinki. In the future, this generated material will be made accessible with the implementation of a web application in collaboration with the designer Yehwan Song[12]. The goal will be to propose to the user navigation of these fictional projections, with the possibility to move from one space to the next, thus discovering the HAM collection through the gaze of the machine and opening the question of the structure of the navigation in this synthetic geography. As a complement to this free navigation, a future line of research will be to use other machine learning models to automatically generate narratives and shape new threads of exploration of that space.

REFERENCES

[1] Jean-Loup Amselle, Noal Mellott, and Julie van Dam. 2003. Primitivism and Postcolonialism in the Arts. *MLN* 118, 4 (2003), 974–988.

[2] Tianqi Chen and Carlos Guestrin. 2016. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (San Francisco, California, USA) *(KDD '16)*. ACM, New York, NY, USA, 785–794. https://doi.org/10.1145/2939672.2939785

[3] Corinna Cortes and Vladimir Vapnik. 1995. Support-vector networks. *Machine learning* 20, 3 (1995), 273–297.

[4] Audrey B Davis. 1980. A museum curator. *Public Hist.* 2, 4 (July 1980), 97–99.

[5] Mohamad Dolatshah, Ali Hadian, and Behrouz Minaei-Bidgoli. 2015. Ball*-tree: Efficient spatial indexing for constrained nearest-neighbor search in metric spaces. (Nov. 2015). arXiv:1511.00628 [cs.DB]

[6] A Flexer. [n. d.]. *Discovering X Degrees of Keyword Separation in a Fine Arts Collection.*

[7] A Flexer. 2021. Computational Filling of Curatorial Gaps in a Fine Arts Exhibition. In *Proceedings of the 12th International Conference on Computational Creativity (ICCC '21.*

[8] Francis Galton. 1886. Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland* 15 (1886), 246–263.

[9] Gabriella Giannachi. [n. d.]. Into the Space of the Digital Museum. In Moving Spaces, Enacting Dance, Performance, and the Digital in the Museum. https://edizionicafoscari.unive.it/en/edizioni4/libri/978-88-6969-535-3/into-the-space-of-the-digital-museum.. Accessed: 2023-4-26.

[10] Nathalie Heinich and Michael Pollak. 1989. Du conservateur de musée à l'auteur d'expositions : l'invention d'une position singulière. *Sociol. Trav.* 31, 1 (1989), 29–49.

[11] Tin Kam Ho. 1995. Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition*, Vol. 1. IEEE, 278–282.

[12] Anna Laura Jones. 1993. Exploding canons: The anthropology of museums. *Annu. Rev. Anthropol.* 22, 1 (Oct. 1993), 201–220.

[13] P Kompatsiaris. 2017. *The Politics of Contemporary Art Biennials: Spectacles of Critique, Theory and Art. Routledge Advances in Art and Visual Studies.* Routledge, New York.

[14] Joasia Krysa. 2014. *Can Machines Curate. Digital Art: Fractures, Proliferative Preservation and Affective Dimension.*

[15] Brady Lund, Ting Wang, Nishith Reddy Mannuru, Bing Nie, Somipam Shimray, and Ziang Wang. 2023. ChatGPT and a new academic reality: Artificial intelligence-written research papers and the ethics of the large language models in scholarly publishing. (March 2023). arXiv:2303.13367 [cs.CL]

[16] Zhihan Lv, Wen-Long Shang, and Mohsen Guizani. 2022. Impact of Digital Twins and metaverse on cities: History, current situation, and application perspectives. *Appl. Sci. (Basel)* 12, 24 (Dec. 2022), 12820.

[17] L Manovich. 2017. Automating aesthetics: Artificial intelligence and image culture. *Flash Art International* 316 (2017), 1–10.

[18] Leland McInnes, John Healy, and James Melville. 2018. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. (Feb. 2018). arXiv:1802.03426 [stat.ML]

[19] Darío Negueruela Del Castillo. 2017. The City of Extended Emotions.

[20] Karl Pearson. 1901. LIII. On lines and planes of closest fit to systems of points in space. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* 2, 11 (Nov. 1901), 559–572.

[21] Alex H Poole. 2016. The conceptual landscape of digital curation. *J. Doc.* 72, 5 (Sept. 2016), 961–986.

[22] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. 2021. Learning transferable visual models from natural language supervision. (Feb. 2021). arXiv:2103.00020 [cs.CV]

[23] Rene Ranftl, Katrin Lasinger, David Hafner, Konrad Schindler, and Vladlen Koltun. 2022. Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer. *IEEE Trans. Pattern Anal. Mach. Intell.* 44, 3 (March 2022), 1623–1637.

[24] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Bjorn Ommer. 2022. High-resolution image synthesis with latent diffusion models. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (New Orleans, LA, USA). IEEE.

[25] I Sample. 2023. *Science journals ban listing of ChatGPT as co-author on papers. The Guardian.* St Martin's Press.

[26] Robert E Schapire. 2013. Explaining adaboost. In *Empirical inference.* Springer, 37–52.

[27] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. 2019. ESRGAN: Enhanced super-resolution generative adversarial networks. In *Lecture Notes in Computer Science.* Springer International Publishing, Cham, 63–79.

[28] Lvmin Zhang and Maneesh Agrawala. 2023. Adding conditional control to text-to-image diffusion models. (Feb. 2023). arXiv:2302.05543 [cs.CV]

[29] Terry Yue Zhuo, Yujin Huang, Chunyang Chen, and Zhenchang Xing. 2023. Exploring AI ethics of ChatGPT: A diagnostic analysis. (2023).

[12] See https://yhsong.com/.