Interactive Media in Public Spaces:

Ethics of Surveillance, Privacy and Bias

Ana Herruzo[1,2(✉)]

[1] Universidad Politécnica de Madrid, P Juan XXIII, 11, 28040 Madrid, Spain
Ae.herruzo@alumnos.upm.es
[2] Arizona State University, Tempe, AZ 85281, USA

Abstract. When designing public interactive environments, new advances in computing and data collection techniques from the users can enhance the public's engagement and interaction with the designed space. Consequently, ethical questions arise as the ambiguity surrounding user information extraction and analysis may lead to privacy issues and biases. This paper examines topics of surveillance, privacy, and bias through two interactive media projects exhibited in public spaces. In particular, this paper focuses on analyzing the curation and manipulation of a database of an AI-powered and computer vision-based interactive installation that uses advances in computing such as DeepLearning and Natural Language Processing but runs into privacy and gender bias issues. This paper aims to showcase how decision-making, data curation, and algorithmic processes can directly impact and reinforce surveillance, privacy, and bias in public spaces.

Keywords: Media architecture · Artificial intelligence · Ethics

1 The Shifting Paradigm

We live in unprecedented times, where the rise of technology and artificial intelligence are now contributing to unique new possibilities in art and architecture and contributing to society's advances in many ways. In a similar manner that digitalization has impacted society, creating new economic models, it is also essential to reflect on its influence at cultural and artistic levels and the consequences of implementing new technological advances in our practices, the users, the cities, or today's society. As more artists and architects begin to integrate digital media into architectural and public spaces, benefiting from using the latest technologies to create interactive experiences, some ethical questions can arise.

Due to the rapid evolution in hardware and computing power, machine learning is becoming more accessible, affordable, and widespread; a whole new range of tools are available to artists, architects, and designers to explore human-computer interactions and digital art. When bridging art with new technological advances such as surveillance systems and machine learning algorithms, we begin working and manipulating data from users. We encounter an interesting space, an interstitial bubble, where questions and ethics of surveillance and privacy should be accounted for, especially now that it has

D. Gerber et al. (Eds.): CAAD Futures 2021, CCIS 1465, pp. 135–150, 2022.
https://doi.org/10.1007/978-981-19-1280-1_9

been demonstrated that some machine learning algorithms can contribute to perpetuating cultural and stereotypical biases [1].

When designing public interactive environments, new advances in computing and data collection techniques from the users can enhance the public's engagement and interaction with the designed space. Consequently, ethical questions arise as the ambiguity surrounding user information extraction and analysis may lead to privacy issues and biases. In this paper, we examine data collection, privacy, and gender bias by studying a couple of digital media interactive projects in the public space of two different museums. We examine these projects by looking at their design process, the technology implemented, and their impact on Surveillance, Privacy, and Bias. We pay special attention to analyzing the creation and manipulation of a database of an AI-powered and computer vision-based interactive installation deployed in the public space of a museum in Los Angeles, CA. This project uses advances in computing such as DeepLearning and Natural Language Processing to create deeper engagements with users and add new layers of interactivity, but as a consequence, runs into issues of privacy and gender bias.

This paper aims to exemplify how when integrating new technologies into interactive art pieces. When deploying them into public spaces, ethics, trust, and transparency towards users' data in contemporary digital culture should be accounted for during the design process. Also, ultimately, it aims to showcase how decision-making, data curation, and algorithmic processes can directly impact and reinforce surveillance, privacy, and bias in public spaces.

2 Interactive Environments and Architecture

In his book New Media Art, Mark Tribe [2] describes the term New Media art as "projects that make use of emerging media technologies and are concerned with the cultural, political, and aesthetic possibilities of these tools". He locates New Media Art as a subset of two broader categories: Art and Technology and Media Art. Tribe also explains how the term is used interchangeably with categorical names like "Digital art," "Computer art," "Multimedia art," and "Interactive art".

The work discussed in this paper falls within that same intersection, studying and utilizing scientific and technological disciplines with an overall application to different media environments, not just digital but also physical and human-scaled spaces.

The concepts of interactive art and environments have been present for many years now, starting with pioneering experiments such as in 1977 Myron Krueger's "Responsive Environments" [3], which explores different types of human-machine interaction, and the potentials of interactive art, and its implications in a number of fields. Since then, there has been much experimentation with different human-computer interactivity scenarios, such as computer vision, voice-activation, body and skeleton tracking, EEG brain waves, thermal cameras, and many types of sensors and controllers.

Technology can allow us to create interactive pieces where the user can be an active agent of the story and design media in a non-linear narrative, opposed to traditional time-based media such as cinema. These new technologies have also allowed us to take the media out of the box (the traditional fixed display), break with linear narrative storytelling, and make the user, with its interactive experience, the active driver of the story.

Throughout the cinematic century, the dominance of the moving image has been a single image on a single frame. Anne Friedberg, in her book The Virtual Window [4], does a quick run through the history of the cinematic century and the evolution of digital media formats, explaining how the traditional "canvas/frame/window" used in film has now multiple forms and shapes and is made of all kinds of hardware or materials or tridimensional physical environments. Andy Warhol's multimedia experiments, such as the expanded sensorium of performative "happenings" on "Exploding plastics inevitable shows" with the Velvet Underground or "The Eameses' Multimedia Architecture," are examples where artists and architects began to develop successful installations that broke with the single frame in the 1960s. However, overall, breaking the frame, and transitioning into new immersive, experiential, interactive environments, is a phenomenon expanding highly in the last decades with the rise of New Media Art (Fig. 1).

Fig. 1. Projection-mapping project for the opening of the exploratorium museum, by obscura digital.

Projection-Mapping is an excellent example of taking media out of the box, and creating custom formats, and using the building as a canvas to paint, design or interact. This technique has become widely used in the last years to augment media, create experiences on the facades of buildings, and turn buildings into fantasies and interactive spaces. Other examples of taking media out of the box and integrating it into the built environment are Gensler's new Digital Experience Design projects [5], having teams specializing in this discipline opening in several offices across the United States. These teams focus on integrating digital media into architecture and incorporating the user experience design into their projects. These teams that work with multiple applications in digital environments (ubiquitous connectivity, touchless solutions, mobile access,

and data intelligence) are made of multidisciplinary roles, experience designers, UX designers, computer scientists, systems integrators, motion graphic artists, animators, media producers, and creative technologists; is essence hybrid art and technology teams.

Artist Refik Anadol [6] creates large immersive digital installations that interact with the architecture of the space, and he creates mesmerizing animations of particle systems using neurological, environmental, and geographic data, with the help of artificial intelligence; more specifically, Machine Learning. His work is an excellent example of how, along with the rapid evolution of real-time rendering engines, programmable shaders, and new algorithms, it is now possible to effectively create instant real-time data-driven media at large resolutions and excellent rendering quality. Due to this evolution in hardware and computing power, machine learning is also becoming more accessible, affordable, and widespread; therefore, an entirely new range of tools is available to explore user interactions and digital art, as exemplified in the case studies described in this paper.

2.1 The Shifting Context: Public Space and Its Audience

Interactive experiences using machine learning are often developed in closed environments such as labs and have a small pool of user testing during the project's development. In interactive art, they are often exhibited in private spaces, where issues of privacy or bias will not stand out as much, given the exposure of the piece and the number of observers.

This changes when we work in a public space, the user pool, becomes the general population, and the exposure of the media can reach anyone passing through the space, and it can be filmed, photographed and shared. This shift from private or semi-public to exhibiting in public space detonated this paper's findings. Along with the realization of the ethical issues that can arise when using interactive experiences, computer vision, and user analysis based generated art in public environments. The boundaries of consent, privacy and biases become blurred.

In these case studies, we describe two projects that create interactive digital canvases in public environments in the courtyards of two museums. We perform an analysis looking at topics of surveillance, privacy and bias, and methodologies implemented to avoid the ethical issues that could be derived from them.

3 Surveillance, Privacy and Bias

Nowadays, as part of the digital age, advances in Computer Science -while becoming part of our daily lives- are triggering a series of unprecedented models in today's society. Shoshana Zuboff [7], in her book "The Age of Surveillance Capitalism", describes the shift that has occurred from traditional mass production lines in factories from the 19th century to a new business model in digital form that is fully reliable in users data, predictions models, and behavior manipulation to produce revenue. She performs an exciting walkthrough of the evolution of the history of Silicon Valley tech giants and digital corporations such as Google, Microsoft, Apple to describe how this new business model has emerged in the last 20 years. This model relies highly on user data and its

analysis using machine learning algorithms to extract accurate predictions. She argues that these predictions are ultimately able to impact the behavior of millions of users worldwide.

Topics of users' data gathering, and privacy are increasingly becoming of public interest these days, for example, on Netflix's documentary "The Social Dilemma". In an interview with a Silicon Valley software developer, he describes how many young programmers developed groundbreaking technologies and software tools such as Google Drive, Google Maps, but are now highly concerned about how design decisions that were made by 25/35 year-olds are impacting the life and behaviors of billions of people. The industry was not prepared and unaware of the impact these advances and tools would have worldwide as Surveillance Capitalism becomes the dominating form of digital life. The importance of decision-making is a crucial topic to be revisited further in this paper.

During the COVID-19 pandemic, it has been proved that forms of surveillance are beneficial, such as tracking the spread of the virus with the pandemic applications, Q.R. codes scanning, radarCovid in Europe and others implemented in Asia. The European app does not serve the goal of surveillance capitalism and behavior surplus, or engage in prediction practices, as stated on their privacy terms [8] but instead is solely meant to help stop the spread of a virus that has dramatically disrupted the world as we have known it up to now. Nevertheless, in many other cases, these applications, as a consequence of surveillance capitalism, incur on instrumentalism [7]. These prediction paradigms are often offered to us as the technological dream of public safety, national security, fraud detection, and even disease control and diagnosis. They are often accepted as valid and objective, but instead, they can naturalize and amplify discriminatory outcomes [9].

3.1 Machine Learning Biases

An algorithm bias or A.I. bias is a phenomenon that happens when an algorithm generates results that are "prejudiced" due to wrong assumptions during the process [10].

Machine learning derived biases are essential topics discussed and highly studied nowadays in computer science due to their possible impact and potential to harm groups or individuals. It is a difficult task since it also has social, cultural, and legal consequences in their applications beyond their mere scientific development.

There are severe concerns about how A.I. and Machine learning can perpetuate cultural stereotypes, resulting in biases regarding gender and race. Some studies measure attitudes towards, and stereotypes of social groups [11], and these cultural stereotypes have been proven to propagate in some machine learning algorithms when using text-based data sets. In a study on "Semantics derived automatically from language corpora contain human-like biases" [1], scientists replicated findings where female names are more associated with family than career words, compared with male names. To avoid biases being perpetuated by automated machine learning algorithms, scientists are beginning to propose different methods for algorithms to avoid these biases. For example, [12] provides a methodology for modifying an embedding to remove gender stereotypes and developed algorithms to "de-bias" the embedding.

In the city's environment, we are also beginning to acknowledge issues relating to the application of these algorithms. Journalist J. Fasman [13] talks about the "predictive policing programs" implemented in the U.S. to deploy more police officers in specific

spaces by using historical crime data, but this data is not objective since neighborhoods can change over time. He highlights how these programs can run the risk of essentially calcifying past racial biases into current practices.

Machine learning bias often stems from problems introduced by the individuals who design or train the machine learning systems. While creating machine learning algorithms and training them with data sets, data scientists should try to work with the data in ways that will minimize these biases. Nevertheless, decision-makers are the ones who need to evaluate when machine learning technology should or should not be applied. In the case studies analyzed in this paper, the designers/developers are both scientists and decision-makers; it is an interesting contemporary scenario that will allow us to make decisions and curate the database to avoid creating biases.

4 Case Studies

4.1 Emergence

"Emergence" is an interactive and audiovisual project, for the opening of the Exploratorium museum. It was created by Obscura Digital, and the author was the Lead Interactive Engineer. The main facade was displaying media created in the studio by filming analog experiments and were designed to fit the front facade of the Pier. The second façade, located in the museum's courtyard, displayed an interactive experience, and it is the part we will focus on in this paper (Fig. 2).

Fig. 2. Interactive experience consisting of thermal imaging surveillance of public space users and projection-mapped onto the museum's façade.

This interactive experience used thermal imaging cameras to record the attendees and project them onto the building, to make the attendees part of the project itself. In this way, they saw themselves augmented onto the building but transformed into a different colour palette since the thermal imaging cameras can capture and analyze temperatures of objects and return media by applying different color ranges to temperatures. Thermal imaging systems function similarly to conventional cameras and are used to capture an image of an environment, but instead of using visible light to construct an image, a thermal imaging system utilizes I.R. (infrared) light to form an image.

The two thermal imaging cameras were placed in public space, and any person utilizing the space could easily enter the field of view of these cameras and suddenly be projected and augmented to a canvas ten times their size. Their images could be seen from afar and several streets away due to the scale of the building and the projection.

The footage from the cameras was live-streamed to the severs and projection-mapped onto the Museum's facade.

Analysis

Surveillance. No surveillance or privacy issues were raised during the design process or after the project. citizens attending the show enthusiastically engaged with the installation, bringing different objects up to the cameras and playing with them to explore how temperature changes were reflected in the footage.

Privacy. Thermographic cameras are widely used in veterinary testing [14] to diagnose animals without carrying out exploratory procedures. Perhaps if this project had been developed in a post-pandemic future, now that our temperature is being taken prior to entering many establishments, flights and restaurants, privacy issues could be raised. If a person walking through the courtyard were to have a fever, this information could be revealed.

Bias. Publicly displaying users' corporeal thermal images could also lead to bias, for example, viewers assuming that someone with a high temperature could automatically be infected by COVID-19.

4.2 WISIWYG: What I See is What You Get

The second case study that we look at in this paper goes much deeper into surveillance, privacy and bias since during this project, these issues were raised during the design process and user testing, which allowed us to correct them. We will detail a complete description of the installation, the design process, and how we managed to rectify the possibility of incurring issues of privacy and bias.

This project was developed in an academic environment, as part of the Applied Computer Science-Media Arts program at Woodbury University and led by professors Ana Herruzo and Nikita Pashenkov. It was exhibited at a public open space of the J. Paul Getty Museum. It consists of a large interface that essentially serves as a mirror to the audience and incorporates live interactive visuals based on data extracted from the user's movements and facial expressions, age, and gender. These visuals are accompanied by synthetic texts generated using machine learning algorithms trained on the Museum's art collection (Fig. 3).

Fig. 3. Users while experiencing the installation and playing with their silhouettes on the screen.

As part of an academic project, in conversations with educational specialists at the museum, interest was shown in exploring human emotions as a thematic element in a project that aimed to merge art and technology. The theme of working with human emotions led us to explore the state-of-the-art scenarios and algorithms that could help us detect facial expressions in humans. Later, we explored how we could connect human emotions depicted in the art pieces from the Museum's art collection with an installation that would read the attendees facial expressions. Part of the project included guided visits to the Museum to study some of the artworks in its collection, learning about the artists' intentions and what the works are communicating in their portrayal of emotions and facial expressions.

While studying the Museum's art collection [15] that comprises Greek, Roman, and Etruscan art from the Neolithic Age to Late Antiquity; and European art from the Middle Ages to the early 20th century; with a contemporary lens, questions arose regarding static and finished pieces of art, in contrast to interactive and responsive artworks [3]. A key observation guiding our concept was that artworks in exhibitions and museums are typically accompanied by a title and a brief description. In our case, unique visuals would be generated based on user interaction, and we proposed creating new synthetic titles and descriptions to accompany each user engagement. It seemed fitting that the textual output would be generated by machine learning algorithms trained on existing texts from the Museum's art collection to create a unique connection bridging the Museum's carefully curated static content with new dynamically generated visuals. The resulting installation utilizes a Kinect sensor to analyze and mirror the users' movements and a separate camera to read facial expressions via computer vision and Deep Learning algorithms, using their outputs in the next stage as a basis for text generation based

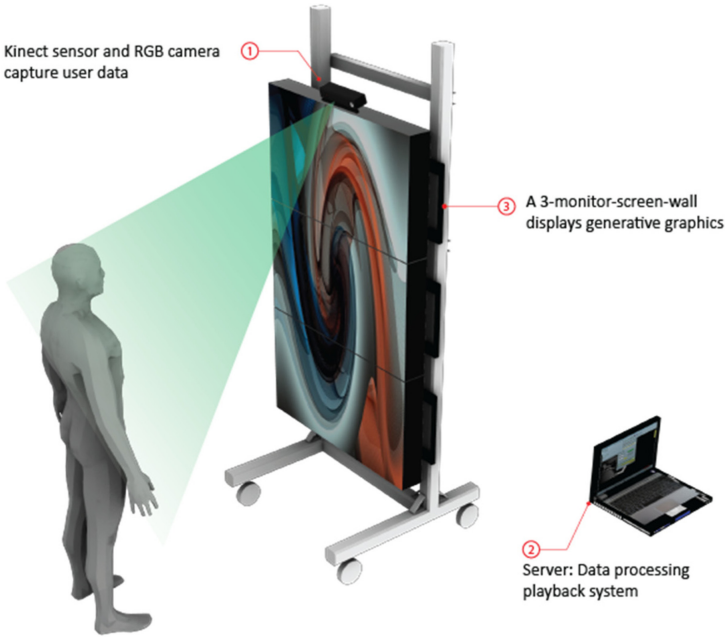


Fig. 4. Installation diagram

on natural language models trained on the text descriptions of the Getty Museum's art collection (Fig. 4).

Our goal was to create a digital art interface with an intimate connection with the users while simultaneously generating new periodic content that is constantly evolving, changing, never the same. This is in direct contrast to architectural buildings, or the art exhibited by the Museum, both consisting of finished, static objects. The project called "WISIWYG" is a play on the popular acronym "What You See Is What You Get" (WYSIWYG), based on the idea that the installation incorporates computer vision processing and machine learning algorithms to generate outputs according to what it sees from its own perspective. In some way, it is an "intelligent" facade that serves as a mirror that reads you and returns media generated with your data, movements, and interactions, and also outputs a text description of the new art piece created with all of the information extracted (Fig. 5).

Features that make this project unique include the combination of real-time generative graphics with exciting new machine learning models.

Computer Vision. The primary driver of the project is the camera input, processed through computer vision and machine learning algorithms. In the course of users' engagement, computer vision is first used to isolate faces and determine the number of people in the camera's field of view. This step is accomplished via a traditional computer vision face tracking method using the popular Open Computer Vision (OpenCV) library [16]. The second step uses a Deep Learning model based on a Convolutional Neural Network

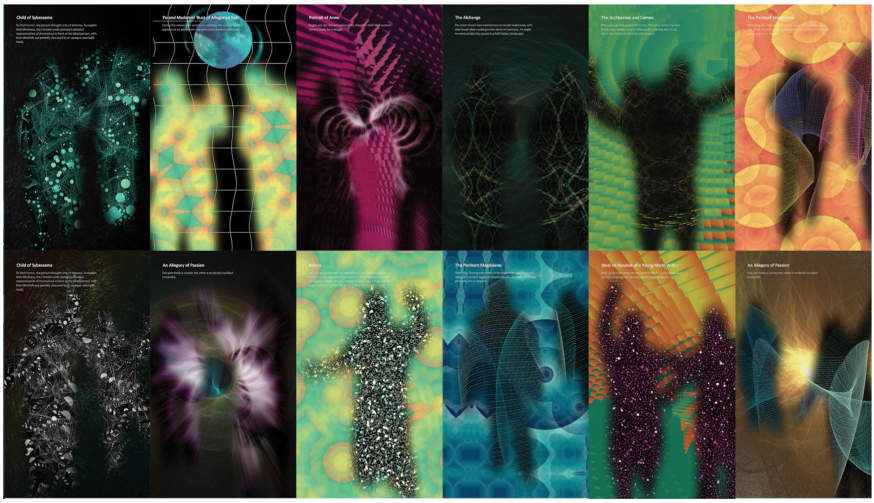


Fig. 5. Gallery showcasing several of the animations designed by the students and the different color palettes applied to them depending on the users' emotions.

(CNN) [17] constructed in Python with the help of Keras [15] and TensorFlow frameworks. The Python code to construct the facial expression detection models is available in an open GitHub code repository [19].

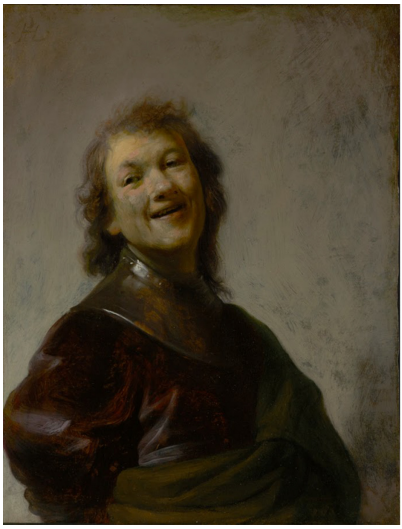
In addition to facial expression detection, the project incorporated Deep Learning models for age and gender detection. The age detection was built on a successful CNN architecture based on Oxford Visual Group's (VGG) model [20].

Data Curation. This portion of the project began by screening Getty's art collection online and selecting artwork depicting people. The focus was on artefacts containing people, based on the total number of pieces on display at the Getty Center at the time of data collection (1,276 results according to the website). As the next step, a database was created by recording the artworks' titles and descriptions and subjectively estimating the number and ages [23] of people featured.

The database was created by analyzing the 1276 artworks and collecting the following data for each piece: Title, Artist/Maker, Date, Description, Primary Sentence, Number of People, Gender, Age, Emotion and reference image (Fig. 6).

In analyzing the Getty's Museum's art collection, students experimented with the Deep Learning natural language model GPT-2 [24], an acronym for Generative Pre-Trained Transformer, released by the non-profit foundation OpenAI in February 2019. The language model generates convincing responses to textual prompts based on set parameters such as the maximum length of response and 'temperature' indicating the relative degree to which the output conforms to the features resembling training data.

Initially, the text prompts were pre-generated by the students based on their own analysis of the following: artworks and consisted of short singular and plural descriptions like "sad young boy", "two happy women", "old, scared people". The text prompts were interactively inputted to the GPT-2 model to generate responses entered into a database



Title: Rembrandt Laughing
Artist/Maker: Rembrandt Harmensz. van Rijn (Dutch, 1606 - 1669)
Date: 1628
Description: Intently interested in the expression of human emotion, Rembrandt often used himself as his own model in his early years as an independent master in Leiden. Here, in a small and freely painted work, he appears in the guise of a soldier, relaxed and engaging the viewer with a laugh.For this sophisticated self-portrait, painted at age twenty-one or twenty-two, Rembrandt combines a study of character and emotion (known in Dutch as a *tronie*) with a rare jovial self-presentation. The lively, short brushwork in the face and brisk handling of the neutral background convey a sense of spontaneity and immediacy. This is one of a small number of paintings by Rembrandt from the late 1620s executed on copper. He signed it in the upper-left corner with his monogram of interlocking letters, "RHL" which he used only briefly, from late 1627 to early 1629.
Primary Sentence: He appears in the guise of a soldier, relaxed and engaging the viewer with a laugh.
Number of People: 1
Gender: Male
Age: 20s
Emotion: happy
Image:https://media.getty.edu/museum/images/web/enlarge/3447160 1.jpg

Fig. 6. Item from the database created for the 1276 artworks from the Getty museum's collection.

and associated by rows with tagged columns for age, the number of people, gender, and facial expressions. The database content was then programmatically correlated with the outputs of computer vision processing and Deep Learning classification using the Pandas library in Python by selecting a random database cell containing a response that matches detected facial expression tags. Finally, the selected response was rendered as the text description that accompanied visual output onscreen (Fig. 7).

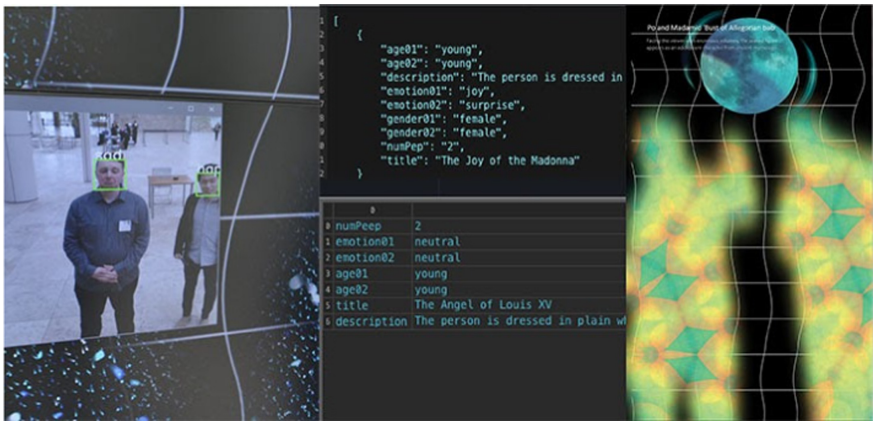


Fig. 7. Image of the students testing the face recognition algorithms at the Getty on the left. On the right, examples of user data incoming into the system and the generated graphic content.

Analysis

Surveillance. Computer vision, the installation uses camera-based computer vision to analyze the facial expressions of the users. The piece is surveilling the attendees.

Privacy. The installation creates visuals, using the data from the users, performing an aesthetic and artistic interpretation of the data, an abstraction. Therefore, this data is not transparently displayed. However, the installation created text descriptions to accompany these graphics, and these descriptions use gender pronouns and allegations of emotions and age (Table 1).

Table 1. Example of a prompt generated by the machine learning algorithms. Data fed is on the left, and A.I. generated prompt on the right.

Input data from users	Machine learning generated prompts
Gender: "Woman" **Emotion:** Fear **Age:** Young	**Title**: Child of Sybessems **Description:** To her horror, she thought only of demons. To explain her blindness, the Christian artist painted a detailed representation of himself in front of the blind girl, with Her blind left eye partially obscured by an opaque, tear-bald mask

Bias. At the beginning of user testing, some of the students raised concerns about using gender in the camera-based analysis and generated texts, implying that it could potentially make some of the attendees uncomfortable. This also raised important questions regarding the role of gender in today's society and opened the door for further discussions involving human interactions. We realized that we were incurring a bias since the algorithm we were using performed a binary gender analysis, and we had not considered non-binary gender identification.

Decision Making: De-gendering

We attempted to address this issue in part by avoiding the use of gender and programmatically manipulating the generated texts, as well as manually editing our database by screening for male and female pronouns, attempting to "de-gender" it by replacing those with the plural "they, "or neutral "person", wherever appropriate (Table 2).

Table 2. Example of the de-gendered A.I. prompts, ones the databases was revisited.

Input data from users	Degendered prompts
Gender: -- **Emotion:** Fear **Age:** Young	**Title**: Child of Sybessems **Description:** To their horror, the person thought only of demons. To explain their blindness, the Christian artist painted a detailed representation of themselves in front of the blind person, with their blind left eye partially obscured by an opaque, tear-bald mask

In this particular case, we purposely avoided gender biases by ultimately deciding not to use the data from gender recognition and by carefully manipulating and re-formatting our database. For a more detailed description of this project's creative and technical processes, see the following references [25] and [26].

5 Discussion

With the intent to bring art and technology closer and explore the boundaries of interactive art, users' interaction, and the levels of communication that the art piece can achieve with its viewer and user, we encounter an interesting space where questions and ethics of surveillance and privacy become relevant. A practice involving surveillance might be legal in some contexts but could be considered unethical in terms of human rights [27]. This adds to concerns about machine learning perpetuating cultural and stereotypical biases; or new technological advances impacting today's society. This phenomenon also begins to be present in artistic environments as well, especially when working with interactive experiences in public spaces.

Wired magazine published an article recently interviewing the authors of "Face of the riot" [28], who created a web page [29] where they duplicated every face from the 827 videos taken from the insurrection at the Capitol on January 6, 2021. They created the web page using open-source software and machine learning algorithms similar to those presented in this paper. It is crucial to understand how these databases are curated and later used to train the A.I. algorithms. Faces of the riot use 827 videos from the riots, but, what if this database could be manipulated? Or What if some of the videos were from people who were not present at the actual riot? These questions showcase how nowadays, invasions of users' privacy can occur quite easily, as well as exposure to biases. Therefore, implementing new policies and regulations regarding users' data and privacy is beginning to be highly discussed these days [30].

In his book Automated Media [31], Andrejvic brings up some interesting points related to this topic, highlighting that if automated processes were developed by decentralized systems and by independent people, and if we all had the skills to program our algorithms, we could avoid biases. Nevertheless, he explains how we need an economic investment to support the research in our current societal model to build the infrastructure and the code to generate this automation, maintain it, and debug it. Moreover, within the current economic models, this implies that the organism in charge of funding and creating the automation will inevitably make choices that will influence the level of biases that the automation will carry.

However, in this paper, we try to go further, since we showcase how in our case study while using a decentralized system, crafting our database, and having the skills to train the algorithm, we also run into the same ethical questions as if an external socio-economic institution were to take care of financing and servicing the technical aspects of the project. There are decisions to be made when choosing and curating a database to train algorithms. That decision-making process is not part of the automation but is done prior and is done by us. We are the decision-makers.

6 Conclusions

When designing public interactive environments, new advances in computing and data collection techniques from the users can enhance the public's engagement and interaction with the designed space. Consequently, ethical questions arise as the ambiguity surrounding user information extraction and analysis may lead to privacy issues and biases.

Due to the rapid evolution in hardware and computing power, machine learning is becoming more accessible, affordable, and widespread; a whole new range of tools are available to artists, architects, and designers to explore human-computer interactions and digital art. In interactive art, they are often exhibited in private spaces, where issues of privacy or bias will not stand out as much, given the exposure of the piece and the number of observers. This changes when we work in a public space, the user pool, becomes the general population, and the exposure of the media can reach anyone passing through the public space, and the content can be filmed, photographed, and shared.

This shift from working in a private or semi-public environment to exhibiting in public space detonated ethical issues using interactive experiences, computer vision, and user analysis based generated art. The boundaries of consent, privacy and biases became blurred.

This paper brings transparency to a methodology involved in curating a database fed into an algorithm and how it was manipulated to avoid biases. Aiming to communicate and help designers be aware and more conscious of the data they are using and how their decisions can help avoid incurring biases and privacy issues.

The findings presented in this paper contribute to an ongoing research line that seeks to showcase how new methodologies that benefit from new technological advances could be implemented in the arts to contribute to more ethical frameworks. As technology moves at a faster pace than policy regulations at governmental and institutional levels, we as designers can begin to acknowledge the ethical realities that we may incur while implementing new advances in computing and technology into our works and come up with solutions. Artists and architects can have complete autonomy towards the work that they develop and are usually connected in networks where their work will be displayed, exhibited, and shared with the public, reaching an audience. Therefore, with the use of computational means and decision making, we can begin to tackle and address issues of privacy and bias, propose solutions, and come up with innovative and contemporary practices that can serve as examples to follow.

References

1. Caliskan, A., Bryson, J.J., Narayanan, A.: Semantics derived automatically from language corpora contain human-like biases. Science **356**, 183–186 (2017). https://doi.org/10.1126/science.aal4230
2. Tribe, M., Jana, R., Grosenick, U.: New Media Art (Taschen Basic Art). Taschen America, LLC (2006)
3. Krueger, M.W.: Responsive environments. In: Proceedings of the June 13–16, 1977, National Computer Conference, pp. 423–433. Association for Computing Machinery, New York, NY, USA (1977). https://doi.org/10.1145/1499402.1499476

4. Press, T.M.: The Virtual Window, https://mitpress.mit.edu/books/virtual-window. Accessed 1 Oct 2019
5. Projects|Digital Experience Design | Expertise, https://www.gensler.com/expertise/digital-experience-design/projects. Accessed 20 Mar 2021
6. Anadol, R.: Synaesthetic architecture: a building dreams. Archit. Des. **90**, 76–85 (2020). https://doi.org/10.1002/ad.2572
7. Zuboff, S.: The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. Public Affairs, New York (2019)
8. App RadarCOVID. https://radarcovid.gob.es/. Accessed 16 Mar 2021
9. Gill, K.S.: Prediction paradigm: the human price of instrumentalism. AI Soc. **35**(3), 509–517 (2020). https://doi.org/10.1007/s00146-020-01035-6
10. What is Machine Learning Bias (AI Bias)?. https://searchenterpriseai.techtarget.com/definition/machine-learning-bias-algorithm-bias-or-AI-bias. Accessed 9 Mar 2021
11. Nosek, B.A., Banaji, M.R., Greenwald, A.G.: Harvesting implicit group attitudes and beliefs from a demonstration web site. Group Dyn. Theory Res. Pract. **6**, 101–115 (2002). https://doi.org/10.1037/1089-2699.6.1.101
12. Bolukbasi, T., Chang, K.-W., Zou, J., Saligrama, V., Kalai, A.: Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. ArXiv160706520 Cs Stat. (2016)
13. Surveillance and Local Police: How Technology Is Evolving Faster Than Regulation. https://www.npr.org/2021/01/27/961103187/surveillance-and-local-police-how-technology-is-evolving-faster-than-regulation. Accessed 1 Feb 2021
14. Thermal Camera Applications - Veterinary | Non-Invasive Testing of Animals. https://www.pass-thermal.co.uk/thermal-camera-applications-veterinary. Accessed 20 Mar 2021
15. Collection (Getty Museum). http://www.getty.edu/art/collection/. Accessed 20 Mar 2021
16. OpenCV: https://opencv.org/. Accessed 7 July 2019
17. Lopes, A.T., de Aguiar, E., De Souza, A.F., Oliveira-Santos, T.: Facial expression recognition with convolutional neural networks: coping with few data and the training sample order. Patt. Recognit. **61**, 610–628 (2017). https://doi.org/10.1016/j.patcog.2016.07.026
18. Home - Keras Documentation, https://keras.io/. Accessed 7 July 2019
19. ACS-Woodbury: ACS-Woodbury/WISIWYG (2019)
20. Parkhi, O.M., Vedaldi, A., Zisserman, A.: Deep face recognition. In: Procedings of the British Machine Vision Conference 2015, pp. 41.1–41.12. British Machine Vision Association, Swansea (2015). https://doi.org/10.5244/C.29.41
21. IMDB-WIKI - 500k+ face images with age and gender labels, https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/. Accessed 7 July 2019
22. Serengil, S.: Apparent Age and Gender Prediction in Keras. https://sefiks.com/2019/02/13/apparent-age-and-gender-prediction-in-keras/. Accessed 20 Mar 2021
23. Rothe, R., Timofte, R., Gool, L.V.: DEX: Deep EXpectation of apparent age from a single image. In: 2015 IEEE International Conference on Computer Vision Workshop (ICCVW), pp. 252–257. IEEE, Santiago, Chile (2015). https://doi.org/10.1109/ICCVW.2015.41
24. Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I.: Language Models are Unsupervised Multitask Learners. 24
25. Herruzo, A., Pashenkov, N.: Collection to creation: playfully interpreting the classics with contemporary tools. In: Yuan, P.F., Yao, J., Yan, C., Wang, X., Leach, N. (eds.) Proceedings of the 2020 DigitalFUTURES, pp. 199–207. Springer, Singapore (2021). https://doi.org/10.1007/978-981-33-4400-6_19

26. Herruzo, A., Pashenkov, N.: “What I See Is What You Get” explorations of live artwork generation, artificial intelligence, and human interaction in a pedagogical environment. In: Brooks, A., Brooks, E.I. (eds.) Interactivity, Game Creation, Design, Learning, and Innovation. pp. 343–359. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-53294-9_23
27. Lyon, D., Haggerty, K.D., Ball, K. (eds.): Routledge handbook of surveillance studies. Routledge, Abingdon, Oxon ; New York (2012)
28. A Site Published Every Face From Parler’s Capitol Riot Videos. https://www.wired.com/story/faces-of-the-riot-capitol-insurrection-facial-recognition/
29. Faces of the Riot. https://facesoftheriot.com/. Accessed 3 Mar 2021
30. Williams, S.: Data Action: Using Data for Public Good. MIT Press, Cambridge (2020)
31. Andrejevic, M.: Automated Media. Routledge, Milton Park (2019)