
Pushing the Boundaries of AI Art: an Immodest Proposal

Eugenia Iofinova

Abstract

Recent advances across the field of machine learning have created a world in which the existent, publicly available models, training tools, and compute enable unprecedented access to model building and deployment. This development creates a number of well-documented novel dangers, it also presents also an opportunity for relatively under-resourced artistic cooperatives and even individual artists, to create interactive and performance art outside of the scope we conventionally see in the “AI Art” community today. Inspired by this development, this work argues for the artistic merit of expanding the scope of what we think of AI Art far past what we have so far seen exhibited in venues such as NeurIPS and in museums. We propose a movement toward art that is more directly provocative, that centers humans rather than the “AI”, and that engages with model training rather than simply inference. As a proof of concept, we describe a fictional interactive exhibit, the Penametrator, which invites the users to interact with, and contribute to the training of, a model that estimates the length of a (fully clothed) visitor’s penis.

1 Introduction

The explosion of openly available model architectures and weights [4], as well as a proliferation of publicly available image generation and LLM chat interfaces [7, 2], have, in a profound way, democratized access to deep learning technology, enabling a large — and growing — number of organizations and private individuals to build models for all sorts of applications, some of them benign, and some intentionally or unintentionally harmful [10]. Moreover, due to the free ability to distribute these models, the harms can have reach and impact on an unprecedented scale. The proliferation of these tools means that, as a community, we’ve largely lost control of how, when, and with what purpose these models are created, published, and deployed. In some cases, even regulation cannot be considered a reasonable strategy, as the “bad actors” may be the governments themselves.

While these tools have also largely democratized artistic expression, it is becoming increasingly clear that the range of AI-assisted art that we see today is rather limited in its scope. Generally, existing work consists of art that centers the AI, such as, for instance, artwork that explores latent spaces of models or their outputs. Other works use AI as a tool, either through generative models such as Stable Diffusion, or more subtly in common artist tools with features such as photo editing. Much of this work invites the users to engage with it as an outside observer, passively consuming content and being invited only to search for patterns or form conclusions. To be sure, this artwork is valuable as a way to explore AI, and much of it has substantial aesthetic value made possible through the use of the technology.

We posit that, while frequently beautiful and engaging, these pieces cover only a small portion of what art — and artistic expression — can and should be, and that it captures only a small amount of AI-human interactions. Historically, art has been used as a way to engage with societal shifts

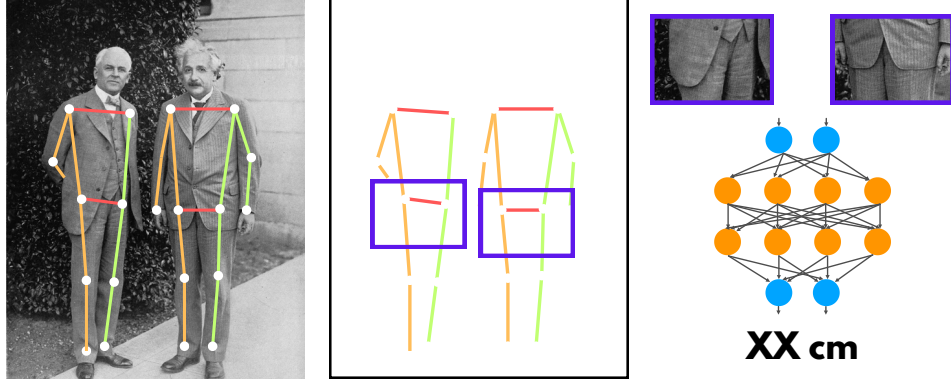


Figure 1: The functionality of the proposed exhibit. (a) Given a (captured) input image, a pose estimation is used to locate and orient their rough body positions. (b) This pose estimate is used to locate the section(s) of the image corresponding to the lower torso. (c) The extracted sections are passed to a neural network trained to (likely, poorly) predict the penis length from the images. In a public display, the predicted length would only be shown in the aggregate, and only if a minimum number of human figures are detected. Image obtained from the Wikimedia commons and distributed under the Creative Commons license.

and events that are dangerous, controversial, or just patently absurd and sort of stupid¹. By contrast, classical art exists on a range of interactivity, ranging from still, to performance, to interactive and even dangerous art (consider, for instance, the groundbreaking work of Marina Abramović[1]). The increasingly easy access to the tools needed to create AI tools means that it is now within the abilities of a single artist, or artistic collective, to create entire interactive artworks - physical or virtual - that encourage the viewer to engage with themes of AI proliferation and harms in ways that have not yet been explored.

As a concrete, though as yet hypothetical example of such an artwork, we propose the Penametrion: an interactive exhibit that estimates the length of a visitor’s penis, if extant, by using photo and/or video capture of their (fully-clothed) body. The prediction would be done using a machine-learning classifier trained on data collected through voluntary submissions solicited through internet channels, and, optionally, through the donated data of the visitors.

The technical details of how such an AI system could be built are fairly straightforward, and discussed in Section 3. In Section 4, we discuss the range of themes explored in the (fictional) exhibit. We further note that the specific exhibit, the Penametrion, is but one sample avenue to explore these ideas. The goal of describing it in detail is to convince the readers of the technical feasibility of the overall approach, as well as to invite readers to engage with their own ideas for using model-buiding and deployment as tools of creative expression to explore the risks and rewards of machine learning and artificial intelligence. Regardless of the reader’s feelings about the specific exhibit, we invite the reader to consider whether these themes are ones that are generally present in AI artwork that has been exhibited at conferences such as NeurIPS or in museum settings, and whether they should be.

2 (Hypothetical) Exhibit description

We envision the exhibit as consisting of two parts. In the public area, a display equipped with a camera can automatically detect the lower torsos of museum visitors and superimpose approximate bounding boxes over a live video feed of museum visitors in the view area of the camera. If enough (e.g., over ten) such areas are detected, the display can additionally show the model’s prediction of the average penis size of the people captured in the feed. Adjoining this exhibit would be a semi-private interactive area that visitors (likely above a certain age) can enter one-by-one to contribute their training data to the model if they so choose. Optionally, these two components may be aug-

¹Consider, for instance, that the four Tesla car models are named S, 3, X, and Y, or the allegations by Scarlett Johansson that OpenAI tried to hire her to voice their assistant[6].

mented with an informational display that is outside the private area and can be seen by several visitors at once.

The semi-private data donation area would consist of a camera and a display screen mounted on a wall. To participate in the exhibit, the visitor is invited to step into a circle drawn on the floor several meters in front of the camera. The camera then take a photo of the user's (clothed) lower torso, and then the display shows the user the photograph, as well as the predicted length of their penis (0, if the model predicts that the visitor does not have one). The visitor can then leave the exhibit, in which case the photograph and prediction are destroyed, or contribute the photo, as well as their estimate of the correct length of their penis, to the model's training database.

In a real-world setting, the informational area could then consist of several information boards, describing the exhibit, the functionality of the AI model that does the prediction, as well as the data collection process. A further display board could describe the real and potential use of AI tools to find out private data on individuals in everyday life, including such items as state surveillance and deepfake pornographic content.

3 Technical Implementation and feasibility

A properly assembled and curated artistic exhibit requires a tremendous amount of work, much of it in the area of layout architecture and graphic design; the successful creation of Penametrone would certainly require the careful design of the measurement interface, as well as the supporting content. These details, while crucial, are arguably outside the scope of AI model building and will not be discussed here. Instead, we focus on how the length estimation could work.

3.1 Data collection

To train a model using machine learning, it generally helps to start with some data, which in this case would, ideally, be manually labeled. The collection of such data - of whatever quality - is the largest technical challenge of such a project. A feasible approach here might be voluntary donation. By posting online a tool for volunteers to upload a photo of their clothed lower torso, along with a measurement, it would be possible to collect data from willing contributors. These contributions could be solicited through the artists' personal networks, or through online channels such as X and Reddit. Depending on the number of submissions, the data quality auditing could be done manually. Note that some measurement error is expected; however, the submitted photographs may be verified that they do, in fact, show a human lower torso, and the measurement may be verified to be within some reasonable bounds.

3.2 Model architecture and training

Real-world inference could be done through a two-step process. First, a pose estimation tool such as Google's MediaPipe for pose landmark detection [5], which uses a custom model similar to a MobileNetV2[8] in its architecture, could be used to crop the user's image to the relevant area. Then, a secondary neural network - likely also a MobileNet - would be used to predict the penis length from the cropped image. Due to the simple architecture and the lack of corrections for issues such as height differences in the users, such a model likely would not be very accurate under the best circumstances - however, we expect the errors from the model architecture choices to be much smaller than the errors from the data collection methodology.

4 Themes and directions

The overall goal of the Penametrone is to encourage visitors to interact with the risks and limitations posed by the rise of artificial intelligence in an uncomfortable, yet playful way. By pushing far outside the bounds of the ways people conventionally engage with AI-supported products, the visitors are encouraged to form a fresh perspective on the potentials of the technology. At the same time, the actual process of building and deploying the AI supporting the exhibit remains entirely conventional, reinforcing the point that the issues regarding classifier quality, data collection, and consent are present across the spectrum of AI-assisted tools and products; thus, the questions and concerns

that arise naturally when engaging with the Penametrone are equally valid when applied to other AI-assisted tools. Overall, the visitors would be encouraged to view AI technology in general as limited not by the skill, but rather the creativity and daring of the engineers, as well as the complaisance of the surrounding society and power structures.

Additionally, the Penametrone directly engages with the ambiguity of what it means for a harmful model to exist. Observe that the construction of such a model, albeit likely quite low-quality, is very easy to accomplish - thus blurring the lines between AI tools that exist only in someone's imagination, those that exist as proofs of concepts only (such as the Penametrone), and those that are massively deployed. In interacting with this exhibit, the visitor may realize that they have no way of knowing whether such a tool already exists in a medium where it may be downloaded and used by their friends and acquaintances, or whether it could start to exist tomorrow. In other words, with AI tools, in a very general sense, we don't really know what's real.

4.1 Privacy and consent

The advances made in the field of artificial intelligence, and specifically in large neural networks, have enabled astonishing technological progress in a variety of areas, and have spurred a push for data collection in all areas of people's private lives, including tracking their shopping habits, medical history, private correspondence, search history, geographical movements, and biometric data such as facial features and fingerprints.

The massive data collection, along with increasing ability of corporations, governments, and sometimes even individuals, to infer or guess private information about individuals from the mined data, raise serious concerns regarding privacy and consent. Generally, consent seems to be an unsolvable problem, since even data collected in public spaces as individuals do everyday tasks can increasingly be used for such inference, e.g., through facial recognition or geotagging. Additionally, much of the inferred information exists in an uncanny valley where the inferred attributes are of high enough quality to be useful in a business sense, yet the data collection and model building process is riddled with biases and errors in data collection, leading to frequent errors in the inferred traits, sometimes with serious real-world consequences.

The ability of AI to infer private information about a specific individual from the combination of that person's public data and other individuals' private data increasingly blurs the line between what is public and what is not. AI-assisted tools already exist today that can create realistic pornographic content from facial and clothed-body images of nonconsenting individuals[3]. Compared to these, the Penametrone is actually quite mild, while still designed to elicit discomfort. Ironically, by inviting visitors to contribute their own data to the model's training algorithm, the exhibits would give visitors a sense of empowerment largely absent from user's everyday interactions with AI models.

Note further that were such a tool, of high quality, to exist - as perhaps it does - it could have very threatening implications, for instance for detecting transgendered people, similarly to existing work that claims to be able to detect people who are homosexual[9]. This is a real social risk, and one that even a provocative exhibit must handle with care. In the actual proposal, we take pains to avoid 'outing' anyone by not publicly displaying any individual's inferred penile length, but only an average one across a large group of individuals.

4.2 Inclusion/exclusion, Sexism and limited usability

By focusing on an organ only present in half of the general population, the exhibit excludes a large subgroup from meaningfully interacting with it. This echoes real-life issues with AI tools, which routinely exclude non-English speakers, people with disabilities, and dark-skinned individuals from meaningful interactions with the tools. At the same time, the AI-assisted tools, which are created primarily by men, are frequently criticized for hyper-sexualizing women, for instance in image generation, or in voices chosen for AI assistants. This exhibit, which centers around a largely meaningless measurement of a primarily male private organ, both flips this tendency around and makes it ridiculous. Visitors who do not have a penis would still be able to interact with the exhibit, both by inviting the software to predict the size of their (nonexistent) one and optionally adding their data to the training database, and by engaging with the supporting displays to learn about the process of designing, training, and deploying the model.

4.3 Model quality; Bias and errors in the data collection process

The Penametrator classifier will be built entirely on self-reported data (see Section 3) and the participation of exhibit visitors. As such, we expect substantial errors and a general bias toward overestimation in the submitted data. This is intentional: the same issues are present in all data collection. In the case of the exhibit, we expect the bias and errors to be exaggerated to the point of rendering the results clearly ludicrous. Each visitor will then have their own opportunity to either improve the training data by being maximally accurate with their data contribution, or to exaggerate their own dimensions, thus driving down the model quality for everyone.

5 Conclusion

In this paper, we proposed a novel paradigm for thinking of newly available tools for AI researches and developers as tools of creative expression that enable individuals with limited access to resources to create interactive artworks exploring the themes of AI and society. As an example, we describe the Penametrator, a fictional, yet feasible exhibit that (badly) estimates the length of a visitor’s penis, and discussed the artistic merits of such a project. There has not, to our knowledge, been any artwork that encourages the visitor to interact with the process of training a model (or creating data for doing so), nor work that invites the visitor to, for a short time, engage with the AI in a way that feels uncomfortable or unsafe. This (immodest) proposal stems largely from our belief that art should not always be ‘safe’, but rather, is a valuable tool to push and question social boundaries. This is especially true for the case of art engaging with Artificial Intelligence, which is, in itself, a largely unsafe technology that routinely pushes social and political boundaries.

References

- [1] Marina Abramović and Ulay. Rest Energy, 1980.
- [2] Anthropic. Introducing Claude, 2023.
- [3] David Chiu, Yvonne Mere, Sara Eisenberg, Ronald Lee, Karun Tilak, Miguel Gradilla, and David Louk. People of the state of California vs. SOL ECOM, INC., BRIVER LLC, ITAI TECH LTD., DEFIREX OÜ, ITAI OÜ, AUGUSTIN GRIBINETS, and DOES 1 through 50, 2024.
- [4] Hugging Face. Hugging Face, 2016.
- [5] Google. MediaPipe Pose Landmark Detection API, 2024.
- [6] Pranshu Verma Nitasha Tiku and Gerrit De Vynck. Scarlett johansson says openai copied ‘her’ voice after she said no. *The Washington Post*, 2016.
- [7] OpenAI. Introducing ChatGPT, 2022.
- [8] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. *arXiv preprint arXiv:1801.04381*, 2019.
- [9] Yilun Wang and Michal Kosinski. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and Social Psychology*, 2018.
- [10] Laura Weidinger, Jonathan Uesato, Maribeth Rauh, Conor Griffin, Po-Sen Huang, John Mellor, Amelia Glaese, Myra Cheng, Borja Balle, Atoosa Kasirzadeh, Courtney Biles, Sasha Brown, Zac Kenton, Will Hawkins, Tom Stepleton, Abeba Birhane, Lisa Anne Hendricks, Laura Rimell, William Isaac, Julia Haas, Sean Legassick, Geoffrey Irving, and Iason Gabriel. Mobilenetv2: Inverted residuals and linear bottlenecks. *ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, 2022.