

# Indiscretions of a Contemporary Artist

## Reflections on Trevor Paglen's (ab)use of the JAFFE dataset

Michael J. Lyons, Ph.D.

Ritsumeikan University

September 20, 2021

On September 9, 2021, the American contemporary artist Trevor Paglen targeted me in an angry outburst. Two tweets to his 14,600 followers and, notably, my employer, labelled me as a *harasser* and *phrenologist*. Paglen accused me of “*preventing [him] from showing work,*” trying to “*shut down [his] publications,*” and “*misrepresenting [myself] to [his] galleries*” amongst other misdeeds. All of the accusations are ridiculous. This commentary examines the context and possible causes of Paglen’s strange tantrum.

To understand the situation, we have to look back to 1996, when I and two colleagues at the Psychology department, Kyushu University, Miyuki Kamachi and Jiro Gyoba, created JAFFE (Japanese Female Facial Expressions), a set of about two hundred photographs intended for use in our research on facial expression perception. When our project concluded in 1998, we made JAFFE available to other scientists for non-commercial research.

A few years ago, without informing us, Trevor Paglen and Kate Crawford obtained JAFFE and began to use the images in their presentations criticizing machine learning research. Use escalated in 2019 when they exhibited more than one hundred JAFFE images in a high-profile art show at the Prada Foundation in Milan. The *Training Humans* exhibition opened in September 2019 and concluded in February 2020. JAFFE images were prominently on display for the entire duration of the show.

Paglen and Crawford’s use of JAFFE images for their planned media spectacle transpired without our knowledge or the consent of the women depicted. When we created JAFFE, the volunteers gave informed consent for non-commercial scientific research using their facial images. However, they did not agree to have their faces flaunted in art exhibitions. To add to the injury, a web essay accompanying the exhibition, *Excavating AI: The politics of images in machine learning training sets*, gave a fallacious account of our work. The essay contained several fabrications, including the fictional claim that the intended use of JAFFE is for training machine learning systems.



Figure 1. Two tweets by Trevor Paglen, 2021/9/9. I have masked the faces.

I first heard about *Training Humans* several weeks after it opened. A Netflix producer contacted me asking for permission to display the JAFFE images in a program covering the exhibition, which I did not grant. I attempted to contact Trevor Paglen to inform him that *Training Humans* had violated the JAFFE terms of use. Paglen did not reply for nearly two weeks and then only with a vague message. In a subsequent exchange, neither Paglen nor Crawford addressed my concerns. Though dissatisfied with their response, I decided not to interfere with the exhibition, hoping it would be an isolated incident.

Busy with my work, I forgot about Paglen and Crawford until July 2020, when I noticed there had been a second exhibition. *Making Faces* was held at *Maxim's* restaurant as part of the Prada Mode Paris fashion event in January 2020. In poor taste, Paglen and Crawford exhibited JAFFE images as looping animations. Further search revealed a viral spread of JAFFE images on Instagram, Twitter, Facebook, YouTube, Vimeo, Netflix — *everywhere*. Paglen and Crawford had permitted gallery visitors to share smartphone captures of the JAFFE images. They had invited journalists, professional photographers, and a Netflix crew to document *Training Humans* and *Making Faces*. JAFFE images appeared in the exhibition catalogue, blog posts, and magazine articles. A JAFFE image served as the icon for the *Excavating*

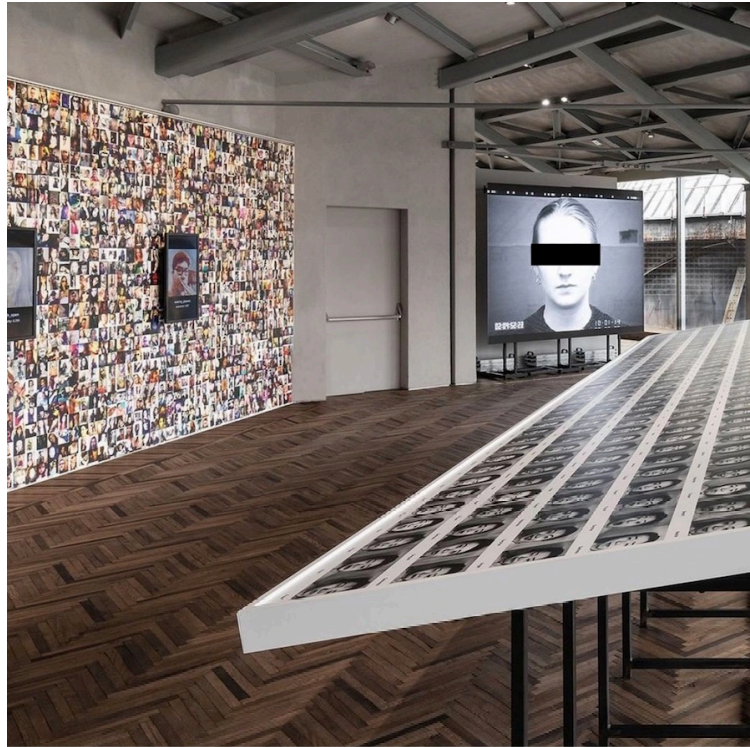


Figure 2. JAFFE on display at the *Training Humans* exhibition at the Prada Foundation, Milan.

*AI* website and auto-displayed each time someone shared the link. Getty Images was selling a photograph showing the recognizable face of a JAFFE volunteer who is now a university dean. Unfortunately, Paglen and Crawford had not informed us about the extent to which they had propagated JAFFE images in flagrant violation of the terms of use.

I attempted to ask posters and hosts to remove the images. Unfortunately, damage control proved time-consuming, frustrating, and only partly successful, so I redirected my efforts and wrote a formal commentary on the incident. *Excavating “Excavating AI”: The Elephant in the Gallery* identifies the problematic ethics of the *Excavating AI* project and the mistaken account of our work.

I asked Paglen and Crawford to refrain from further use of the JAFFE images, but they continued. In early June 2021, the Springer Nature journal *AI & Society* published *Excavating AI*, its text nearly unaltered. JAFFE images appeared in the article and on the journal’s home page. After I contacted the publisher, they temporarily removed *Excavating AI* from the *AI & Society* website. Springer Nature’s Research Integrity Team opened an investigation. In August, they concluded with the decision to remove JAFFE images from *Excavating AI*. Meanwhile, I prepared a second essay, *Excavating AI Re-excavated: Debunking a Fallacious Account of the JAFFE Dataset*, unravelling Paglen and Crawford’s misleading straw man arguments.



Figure 3. JAFFE looping animation, Making Faces exhibition, Prada Mode Paris. Mask added by me.

While Paglen acts as though he is entitled to make *fair use* of the images, he is not exempt from the JAFFE terms or the requirement to obtain consent to exhibit the women's faces. Moreover, the *critique of data* requires careful study and a thorough understanding of the documentation associated with that data. As my commentaries explain, Crawford and Paglen's account of JAFFE consists of poorly informed fiction and fallacy.

I have promoted my critical commentary on Twitter. Excepting Paglen, no one has objected to me sharing my work, and feedback has been positive. Many readers have told me that my analysis improved their understanding of dataset reuse, informed consent, facial expression, and affective computing.

Let us return to Paglen's aggressive tweets. Paglen claims that I am a *harasser*. I have asked Paglen, Crawford, and publishers to refrain from further use of the JAFFE images. I have contacted his galleries a few times to ask that they remove JAFFE images from the internet. I always use my real name and have not *misrepresented* myself — a bizarre claim. Could sharing my critiques on Twitter be considered as an *attack* or harassment? Is this significant compared to the massive media spectacle Paglen and Crawford stirred up for *Excavating AI*?

When Paglen writes that I want to 'shut down critique of training data,' he seems to have forgotten that I overlooked his first violation of the JAFFE terms to allow *Training Humans* to end smoothly. Has he forgotten that I did not, in fact, attempt to interfere with the exhibition?

Why have Paglen and Crawford failed to address the concerns raised in the very first email I sent to them in November 2019, concerns that I elaborated on in my two commentaries? Does Paglen think that his angry ad hominem outburst and the hackneyed insult — *Phrenologist!* — are an appropriate response? Perhaps Paglen has no answer, and his attack is an evasive tactic?

Or does Trevor Paglen lack the discernment to consider the possibility that he is mistaken?

## References

1. Lyons M.J., Kamachi M, and Gyoba J. Coding facial expressions with Gabor wavelets. *Proceedings, Third IEEE international conference on automatic face and gesture recognition*, 1998.
2. Crawford, K. and Paglen, T. Excavating AI: The politics of images in machine learning training sets. <https://www.excavating.ai>, 2019.
3. Lyons M.J. Excavating "Excavating AI": The Elephant in the Gallery. [arXiv:2009.01215](https://arxiv.org/abs/2009.01215), 2020.
4. Lyons M.J. "Excavating AI" Re-excavated: Debunking a fallacious account of the JAFFE dataset. [arXiv:2107.13998](https://arxiv.org/abs/2107.13998), 2021.

### Glossary

#### **phrenology** (*n*)

: the study of the conformation and especially the contours of the skull based on the former belief that they are indicative of mental faculties and character

source: [Merriam-Webster online dictionary](https://www.merriam-webster.com/dictionary/phrenology)