Excavating "Excavating AI":
The Elephant in the Gallery

Michael J. Lyons
Ritsumeikan University

Abstract

Two art exhibitions, "Training Humans" and "Making Faces," and the accompanying essay "Excavating AI: The politics of images in machine learning training sets" by Kate Crawford and Trevor Paglen, are making substantial impact on discourse taking place in the social and mass media networks, and some scholarly circles. Critical scrutiny reveals, however, a self-contradictory stance regarding informed consent for the use of facial images, as well as serious flaws in their critique of ML training sets. Our analysis underlines the non-negotiability of informed consent when using human data in artistic and other contexts, and clarifies issues relating to the description of ML training sets.

Keywords: digital ethics, informed consent, training sets, ai, machine learning, affective computing

In this essay, I will present a case study and critical analysis involving two art exhibitions and an accompanying essay by contemporary artist Trevor Paglen and media studies scholar Kate Crawford. "Training Humans," held at the Milan Osservatorio, Fondazione Prada, from September 2019 to February 2020, exhibited collections of images of human faces used for training computer vision systems. An associated exhibition, "Making Faces," was organized by Prada Mode Paris in January 2020 to coincide with the opening of the Paris Haute Couture event. An essay by Crawford and Paglen, "Excavating AI: The politics of images in machine learning training sets," was self-published online to accompany the exhibitions [6].[1]

According to the Fondazione Prada web page, the exhibition explored:

> . . . *two fundamental issues in particular: how humans are represented, interpreted and codified through training datasets, and how technological systems harvest, label and use this material.*

---

[1] Henceforth, the abbreviations C&P, TH, MF, EAI will refer respectively to the authors, the two exhibitions, and the essay.
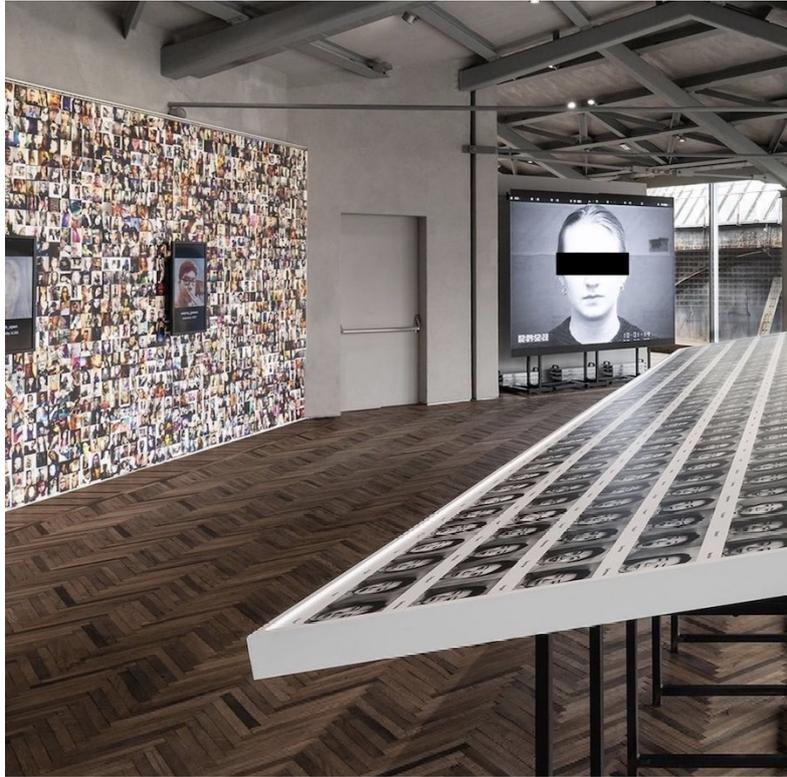
Figure 1: "Training Humans" exhibition, Milan Osservatorio.

Though sympathetic to C&P's overall aims, several issues troubled me about the exhibition, TH, its satellite, MF, and the analysis in EAI [6]. Briefly, the main concerns are:

- C&P's description of machine learning training sets is inadequate and flawed, containing both generic and specific errors
- C&P employed ethical double-standards by exhibiting images of private individuals without informed consent.
- C&P failed to respect clearly stated terms of use for several of the datasets.
- The EAI essay contains important factual errors and misleading statements

The present commentary explains and elaborates on these concerns to propose that C&P's flawed approach compromises their aims. This critical analysis and commentary is intended to contribute constructively to dialogue concerning the use of human data for artistic and other purposes and to correct some of the mistakes that have been instigated by the exhibitions and the essay, EAI.

Disclosure

The following factors have informed and influenced my comments:

- I am a co-author, with colleagues Miyuki Kamachi and Jiro Gyoba, of

the JAFFE image set,[2] that was exhibited at TH, MF, and discussed in EAI.

- I worked briefly on face recognition technology, as manager of a team that took part in Phase II of the FERET face recognition competition held in March 1995. I was not comfortable working on surveillance technology and soon left the field. Since that time, for various reasons, I have been opposed to the deployment of surveillance systems.
- I have a long term interest in contemporary art and have encountered some of Trevor Paglen's projects at exhibitions and in books. My opinion of his past work has been relatively positive.
- There are no competing financial interests.

Training Sets and False Taxonomy

Crawford and Paglen's two exhibitions, TH and MF, and the essay EAI, may be viewed as a critique of image taxonomy, and especially as a caveat concerning the political implications of labelling photographs of human individuals. Most spectacularly, their project popularized bizarre and demeaning labels on some of the person categories in the ImageNet database.

The foundation of C&P's analysis of computer vision training sets, however, is itself marred by a taxonomic blunder. The fundamental problem is that C&P attempt to subsume a very heterogeneous selection of datasets into the single undifferentiated category of machine learning 'training set.' The datasets exhibited at C&P are diverse with regard to origin, intended purpose, status of copyright and informed consent, terms of use, funding source etc.

In the following I will illustrate the importance of making distinctions between various kinds of image datasets. C&P exhibited facial image datasets originating in two distinct ways: sets that were carefully designed and photographed by research groups under controlled laboratory conditions, and sets consisting of images scraped in bulk from the internet. I refer to these respectively as *constructed datasets* and *scraped datasets*.

Consider how their different origin affects the public exhibition of the images. There are, of course, ethical concerns with the unauthorized public exhibition of both kinds of image datasets, but with an important distinction: the status of copyright and informed consent are precisely known for the constructed datasets, and uncertain or unknown for the scraped datasets.

In contrast with scraped training sets, constructed image sets such as JAFFE, FERET,[3] and CK[4] also have explicitly defined *terms of use*. These three sets permit use for the purpose of *non-commercial scientific research*, and allow for limited reproduction of the images in scholarly articles reporting research results.

By exhibiting the images publicly in art shows, C&P breached the terms of use for the constructed sets JAFFE, CK, and FERET. The artists and the Fondazione Prada claim that their use did constitute 'non-commercial scientific

---

[2]The JApanese Female Facial Expressions, a set of images visible on the right in Figure 1.
[3]FacE REcognition Technology dataset [31]
[4]Cohn-Kanade facial expression dataset [21], visible at the back in Figure 1.

Figure 2: "Making Faces" exhibition at Maxim's, Prada Mode Paris.

research'[5], but this entails an untenable semantic stretch.

Consider, for example, the description for the satellite exhibition "Making Faces," held as part of Prada Mode

> ... *a travelling social club with a focus on contemporary culture that provides members a unique cultural experience along with music, dining, and conversations* ...

held at the exclusive Belle Époque restaurant Maxim's featuring 'exhilarating performances,' celebrity guests, and a two-day 'food experience' curated by a Michelin-starred chef. I can only guess the budget for this luxury junket, but do know that the image sets defining the themed event cost them not a single euro cent.

The association of the Prada fashion brand with TH and MF further calls to question C&P's claim of non-commercial status. The Fondazione Prada, which hosted TH, describes itself as a 'non-profit organization with a mission to promote and encourage cultural fields such as contemporary art.' However, holding "Making Faces" as part of the ostentatious Prada Mode Paris fashion event, with the visible participation of the Prada Group CEO, points to a tighter association with corporate interests. The close participation of a private corporation in an exhibition intended to probe what is essentially a matter of public policy, *viz.* the politics of surveillance technology, should itself beg scrutiny. What vested interests might such a corporate sponsor hold? Can such activities be sincerely described as devoid of commercial interest?

More concretely, entry to TH required purchasing a ticket, a commercial transaction. There is an exhibition catalogue (containing JAFFE images) for sale[6] via the museum bookshop. It is not unreasonable to suppose that the exhibitions may have boosted sales of Paglen's photobooks or otherwise helped to move merchandise.

---

[5]Letter from the Fondazione Prada, Aug 6, 2020.
[6]Quaderni Fondazione Prada #26: Training Humans, ISSN 2421-2792

Financial transactions aside, TH/MF created a media spectacle [38], attracting considerable public attention, often considered a boon by artists and academics alike. The bottom line is that one cannot honestly describe TF or MF as 'scientific research' and both took place in contexts that were far from non-commercial.

There is an uncanny resemblance to processes described by Shoshana Zuboff in her analysis of *surveillance capitalism* [43]. Datasets mined online, a "surplus resource" obtained at no cost, were repackaged as commodities for consumption by the contemporary art world and popular media, yielding considerable rewards, financial and otherwise. C&P might argue that their work is intended to render the public a beneficial service. Whether or not such a rationale is justifiable, it strangely echos the alibis of the behemoth surveillance capitalists.

Informed Consent and Digital Ethics Malpractice

Scientific researchers who conduct experiments involving humans are required to provide their subjects with knowledge sufficient to make an informed, voluntary decision to participate or not [1]. This principle, known as *informed consent*, also concerns any personal data, including photographs, acquired in research: subjects must be informed about how their personal data will be used and disseminated, and voluntarily provide agreement [10].

All three constructed datasets were assembled in projects regulated by the human subjects ethics committees of their respective institutions, requiring the informed consent of the private individuals shown in the images and videos. All three sets prohibit redistribution. The FERET and CK pages ask potential users to download, sign, and submit an application form before receiving access to the images. JAFFE, until recently, relied on an honour code, asking users to read and agree to conditions of use before downloading.[7] These policies are intended, in part, to protect the privacy of the depicted individuals, according to the conditions under which they gave informed consent.

Whatever one may think about C&P's mishandling of terms of use for the constructed image datasets in TH/MF, shirking the requirement for informed consent is a more serious matter. The authors of JAFFE, CK, and FERET obtained informed consent contingent on the terms of use. When they downloaded the datasets, C&P certainly did not acquire the right to modify those terms to suit their ends. To exhibit these images in public, C&P should have first obtained permission from the persons depicted. In an interview with Gaia Tedone,[8] Paglen seems to be aware of ethical complications associated with exhibiting photographs of private persons without their knowledge, and Crawford adds:

> . . . we were really careful to create the entire exhibition as a thing where
> if you see a face that you know or yourself that you can actually choose to

---

[7] Until now, there were no serious abuses of this policy in more than twenty years. Access is now restricted and vetted.

[8] "What are the ethics of exhibiting datasets of faces?" Interview by Gaia Tedone [38].

Figure 3: An astonishing tweet.

have it removed and this is a freedom and a sense of agency that doesn't normally exist . . .

This strongly resembles the promise IBM makes for it's non-consensual "Diversity in Faces" dataset [36]. Unfortunately, informed consent makes no sense *ex post facto*—to have any meaning at all informed consent must be obtained *beforehand*.

There are similarities with two types of digital ethics malpractice identified by Luciano Floridi [16]. Slyly twisting the principle of informed consent to fit one's convenience is an instance of *ethics shopping*. Offering to redress the appropriation of personal images, after the fact, is an instance of *ethics bluewashing*. To posit, however, that what is undeniably non-consensual use endows the victim with exceptional 'freedom and agency,' is a mind-bending masterpiece of sophistry.[9] Scraped datasets contain images of varying copyright status, but researchers navigate this barrier by invoking the principle of *fair use* or *fair dealing*. For photographs and videos of private individuals, however, fair use and even the permissive terms of a creative commons license, refer only to copyright and do not satisfy the requirement for informed consent [36, 32]. In EAI, C&P underline the issue, when they comment on

. . . the practice of collecting hundreds of thousands of images of unsuspecting people who had uploaded pictures to sites like Flickr . . .

Kate Crawford has tweeted about the lack of informed consent for the use of private images in scraped training sets (see Figure 3). Remarkably, however, Crawford does not seem to have noticed that she did not have the informed consent of the woman whose face she was tweeting. The irony is acute—this JAFFE image was used, without permission or informed consent, as the icon for the EAI web page: it was embedded automatically with every post of the EAI site to social media. In effect, the EAI site acted as a machine that caused

---

[9]Watch the video to experience the full impact.

Figure 4: An unsupported claim.

anyone who shared it on social media to unwittingly breach informed consent and the JAFFE terms of use.[10]

Social Media Fallout

Am I taking issue with a mere technicality? For the JAFFE images, at least, the non-consensual public exhibition had several unwelcome consequences. Soon after TH opened in Sept. 2019, photos of the JAFFE images, by professional photographers and museum visitors with smartphones, began to multiply.[11] Photos showed up on social network sites like Instagram, Twitter, and Facebook; in news reports and online magazines; in videos on YouTube and Vimeo; on the Fondazione Prada website; on the EAI web page, and in many other places. The proliferation of such photos escalated after MF. A photograph clearly showing the face of one JAFFE woman, now a successful professional with a public persona, appeared for sale at Getty Images with a price tag attached. When I alerted that woman to the situation, she was shocked and dismayed. The JAFFE volunteers certainly did not consent to such indiscriminate dissemination of their photographs.

I noticed the widespread proliferation of JAFFE images in August 2020, long after TH/MF had closed. Too late for that golden promise of exceptional 'freedom and agency.' My attempts to undo the damage caused by the TH/MF media spectacle proved time-consuming, frustrating, and ultimately impossible. I was able, at least, to convince Getty Images to take down the offending photograph, but only after a multi-day effort that needed several emails and intercontinental phone calls.

Views of the JAFFE and CK faces and were broadcast as part of a Netflix science-lite documentary 'Connected' in an episode titled 'Surveillance.' The

---

[10]As of Aug 30, 2020 the site icon had changed, possibly in response to my request. The screenshot in figure 3 is from Aug 23, so the image seems to have been used this way for more than 11 months. I have masked the faces.

[11]Judging by the dates of the posts—I did not see these until many months later.

production company, Zero Point Zero, had contacted me to ask for, and did not receive, my permission to show the images. They did not heed the negative response, nor has Zero Point Zero admitted that this was a serious breach of journalistic ethics. To add insult to the injury, while discussing the images with show host Latif Nasser, Kate Crawford fabulates a narrative attributing military origins and funding to the research (see Figure 4). Months have passed, but Crawford has still not responded to my request to provide an explanation for this groundless claim.

Ceci n'était pas un 'Training Set'

The blurb at the TH exhibition web page states:

> "Training Humans" explores two fundamental issues in particular: how humans are represented, interpreted and codified through training datasets, and how technological systems harvest, label and use this material.[12]

We have seen that JAFFE and some other image sets were not 'harvested' by a 'system' but carefully designed and constructed by researchers. The quote implies another fallacy: that JAFFE is primarily a 'training dataset.' C&P's claim that JAFFE was intended to be a 'training dataset' is pure fiction. As it turns out, JAFFE began with the scientific aim of modelling data on facial expression perception by humans.

C&P say that TH/MF took two years of research to prepare. It appears, however, that they did not find the time to read the documents attached to JAFFE: a "README_FIRST.txt" file, that contains an explicit caveat regarding the image labels, and an article describing how the image set was assembled and how it was used [23]. The article's title alone, "Coding Facial Expressions with Gabor Wavelets," does not mention recognition, classification, or learning. That should have acted as a clue. Had they looked at the article, C&P might have been puzzled to see that no machine learning algorithms were trained, no images classified, and no recognition rates reported.

Let's look more closely at C&P's mistaken account of JAFFE:

> *The intended purpose of the dataset is to help machine-learning systems recognize and label these emotions for newly captured, unlabelled images. The implicit, top-level taxonomy here is something like 'facial expressions depicting the emotions of Japanese women.'*

C&P wrongly project onto JAFFE a purpose that was not what we had in mind when we designed and photographed the image set in 1996. Likewise, we have never claimed, anywhere, that the photos represent felt emotions—we describe the images unambiguously as posed facial expressions. EAI continues:

> *...there's a string of additional assumptions ...that there are six emotions plus a neutral state; that there is a fixed relationship between*

---

a person's facial expression and her true emotional state; and that this relationship between the face and the emotion is consistent, measurable, and uniform across the women in the photographs.

Neither have we nor has Ekman, ever claimed that there are only six emotions. This is an elementary misconception about Ekman's work that a basic familiarity with the introductory facial expression literature could have helped C&P avoid, here, and in several spoken presentations. Lisa Feldman Barrett, a prominent Ekman critic cited by C&P, has conducted experiments using five facial expressions plus a neutral face [18]. Should we conclude, using C&P's reasoning, that Barrett believes there are only five emotions?

From, for example, the critical historical account given by Ruth Leys [22], which EAI cites, C&P could have learned that Ekman's proposal refers to six universally recognized *basic facial expressions* (BFE) plus any number of culturally variable facial expressions [11]. Ekman has written elsewhere that there may be thousands of facial expressions, in addition to his basic six [13]. C&P list several further assumptions that we supposedly made about felt emotion. Unfortunately, these are also made without any justification.

Why we Made JAFFE

What, then, was the intended use of the JAFFE dataset? The project had two non-technological, scientific aims:

- to test the psychological plausibility of a biologically inspired model [26, 25, 27, 29] for facial expression representation.
- to explore the relationship between categorical and dimensional paradigms in facial expression research [24, 23, 28].

Very briefly, emotions have been characterized as *categorical* by Darwin [9], Tomkins [39], and Ekman [12], and as *dimensional* by Wundt [41], Schlosburg [35], Russell [33], and Barrett [3]. Using our "Linked Aggregate Code" [29] and non-metric multidimensional scaling (nMDS), a statistical technique, we discovered evidence for low dimensional structure derived from visual aspects of the BFE images. Without referring to labels or semantic evaluations, we recovered arrangements of the BFE resembling the "affective circumplex." Effectively this suggested a way to bridge the competing categorical and dimensional models whose roots both date to the beginnings of scientific psychology (Wundt and Darwin). The discovery was confirmed a few years later by Dailey and Cottrell, using a different, but related, approach [8].

How JAFFE Became a 'Training Set'

In the interests of an open data policy, we began to provide the JAFFE images and semantic ratings to other scientific researchers.[13] After attending one of my talks, Zhengyou Zhang, a computer vision researcher, initiated a collaboration

---

[13]C&P did not acknowledge that both JAFFE and CK datasets have been widely used in experimental and computational psychology, neuroscience, and other areas not related to machine learning.

on automatic facial expression classification [42]. That was the first use of JAFFE as a 'training set.' Subsequently, many pattern recognition studies have used JAFFE as a benchmark for comparing classification algorithms. JAFFE may have *become* a 'training set,' but we did not create it with this intention, as C&P claim in EAI.

Most such studies use JAFFE not only for training but also for validating and testing algorithms. C&P's presentation naïvely implies that an algorithm is trained using a dataset like JAFFE, then it is unleashed on the real world. Indeed the machine learning studies using JAFFE are typically academic 'toy world' studies that are not deployable in real-world applications.

C&P continue:

> *The JAFFE training set is relatively modest as far as contemporary training sets go. It was created before the advent of social media, before developers were able to scrape images from the internet at scale.*

Their narrative (plainly wrong, again) is that we would have scraped the social networks for facial images had these been available. However, photos with uncontrolled lighting, pose, camera, of subjects, wearing makeup, and jewelry, would not have served our scientific purposes. On the other hand, it would not have been difficult to photograph more volunteers if we had needed to—but the small number of JAFFE posers was sufficient for the original study. Had we set out to build a facial expression 'training set,' we would and certainly could have photographed a larger and more diverse group of people.

Again, JAFFE was not intended for training machine learning algorithms. By using JAFFE as the 'anatomical model' for their exposition of training set 'taxonomy,' C&P based their discourse on a taxonomy of datasets that is itself mistaken.

Mind-Reading Machines

I doubt that many of the engineers and computer scientists who use JAFFE for machine learning research have ever thought much about the psychology of human emotion. Most are probably not aware of the vast and complex literature relating to facial expression. I first noticed this in the collaboration with Zhang. During the preparation of the joint publication, Zhang was eager to claim that the results proved the neural network could recognize facial expressions with greater accuracy than humans—a misunderstanding of the semantic ratings experiments. I had to veto this mistaken claim from the co-authored article [42].

Regarding C&P's fabulation that we aimed to build a machine that reads minds from faces, I have only never discussed 'mind-reading machines' except to express my profound skepticism of such projects [28].

Critique of Ekman's Work

On the TH web page, we find the following oversimplified and factually wrong description of Paul Ekman's work:

> *Based on the heavily criticized theories of psychologist Paul Ekman, who claimed that the breadth of the human feeling could be boiled down to six universal emotions,*

C&P also present facile, and even derisive-sounding accounts of Ekman's work in public talks that are misleading for the unsuspecting.[14] They are not wrong that Ekman's views are contested. That is nothing new: when Ekman entered the field in the 1960s, he encountered opposition from the then dominant social-constructionists Margaret Mead and Ray Birdwhistell [22]. From the 1980s, there has been a prolonged debate involving Ekman and James Russell, Alan Fridlund, and others [34, 17].

Despite the impression conveyed by C&P, however, it is not accurate to describe Ekman's work as discredited. A survey conducted by Ekman in cooperation with his opponent James Russell, and discussed by Alan Fridlund, found acceptance of some of Ekman's main views among a majority of researchers working on the psychology of affect [14, 7]. Lisa Barrett, now leading the critique of Ekman's work and, more generally, the 'standard model,' can be interpreted as recognition of a need for a radical paradigm shift in emotion research [4]. This important critical work has gained attention and interest, but it cannot be said to have established a new standard paradigm [2].

Whether or not we accept Ekman's views,[15] we should recognize that throughout his long career, he has tested his ideas experimentally, published findings in peer-reviewed articles, and engaged in vigorous open debates with opponents. With this in mind, C&P's uninformed and biased[16] portrayal of Ekman's views is regrettable.

Summary: Errors, Ethics, Constraints, and Creativity

> *In disputes upon moral or scientific points, let your aim be to come at truth, not to conquer your opponent. So you never shall be at a loss in losing the argument, and gaining a new discovery.*
>
> *Arthur Martine, 1866 [30]*

C&P's essay "Excavating AI" frames their analysis of 'training sets' in terms of a grand archaeological metaphor, to signify their method of

> *...digging through the material layers, cataloging the principles and values by which something was constructed ...*

In the title of this document, I reuse the verb 'to excavate' more modestly: I intended to dig in and examine how C&P's analysis of 'training sets' holds

---

[14]For example, this talk, Haus der Kulturen der Welt, Berlin, Jan 12, 2019

[15]My own views combine dimensional/categorical and nativist/cultural aspects.

[16]In the sense that C&P mention only Ekman's harshest critics, seemingly without fully understanding the criticism, while neglecting other viewpoints [19, 5, 15, 37].

up to scrutiny. Though my comments are not exhaustive,[17] I have uncovered faulty analysis, elementary errors, misunderstandings, and questionable reasoning. By choosing JAFFE as the 'anatomical model' for their exposition of training set 'taxonomy,' without having made sufficient effort to understand what it is, C&P rashly compromised the core of their discourse. This key error amounts to faulty taxonomy, appropriately enough. By attacking a distinguished scholar, apparently without having studied his writings, C&P raised doubts about their own scholarship. C&P wrote:

> *. . . when we look at the training images widely used in computer-vision systems, we find a bedrock composed of shaky and skewed assumptions*

not realizing that shaky and skewed is a fitting description of the 'Excavating AI' essay itself. To be sure, computer vision is a technically challenging field. The literature of facial expression research is complex and confusing for the uninitiated. Perhaps C&P underestimated the difficulties of crossing disciplines. EAI offers a compelling narrative—for readers not knowledgeable or critical enough to recognize the fallacies.

I began this commentary with an assertion of sympathy for C&P's aims. That has not changed. Surveillance technologies must be monitored and regulated via open, democratic policies. Corporations must defer to the primacy of human rights. A good starting point is expressed clearly in article one of the Nuremberg Code [1]. Informed voluntary consent is essential and non-negotiable when dealing with human data.

By exhibiting images from constructed datasets such as JAFFE and CK, without first obtaining informed consent, Crawford and Paglen demonstrated what is, for my tastes, an insufficient level of respect for this fundamental human right. The flaws in EAI may be disappointing, but the failure to observe the necessity for informed consent reveals an egregious ethical double standard—the elephant in the gallery at "Training Humans" and "Making Faces." If Crawford and Paglen are willing to overlook informed consent, why should they expect anyone to do otherwise?

Is there no way to investigate the aesthetics of facial image training sets without violating informed consent? Creativity is said to thrive in the presence of constraints. Well, here is a constraint related directly to what really is at stake—human agency, freedom, dignity. Surely informed consent is something worth working *with*, not against?

Kate Crawford and Trevor Paglen have yet to acknowledge the serious errors contained in "Excavating AI" or their self-contradictory stance regarding informed consent. Neither have they admitted the negative consequences of their actions. I hope they will eventually realize the importance of doing so.

Michael Lyons is Professor of Image Arts and Science at Ritsumeikan University

---

[17]A full discussion on the 'echos of phrenology' trope is beyond the scope of this commentary. Briefly, EAI does not sufficiently acknowledge the existing, well-documented controversy over the revival of physiognomy [20, 40]. Overall, I am skeptical of the attempt to impose a grandiose narrative.

References

1. The Nuremberg Code (1947). *BMJ: British Medical Journal 313*, 7070 (1996), 1448–1448.

2. ADOLPHS, R., MLODINOW, L., AND BARRETT, L. F. What is an emotion? *Current Biology 29*, 20 (2019), R1060–R1064.

3. BARRETT, L. F. Are emotions natural kinds? *Perspectives on psychological science 1*, 1 (2006), 28–58.

4. BARRETT, L. F. *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt, 2017.

5. COWEN, A., SAUTER, D., TRACY, J. L., AND KELTNER, D. Mapping the passions: Toward a high-dimensional taxonomy of emotional experience and expression. *Psychological Science in the Public Interest 20*, 1 (2019), 69–90.

6. CRAWFORD, K., AND PAGLEN, T. Excavating AI: The politics of images in machine learning training sets. [Online; accessed 27-August-2020].

7. CRIVELLI, C., AND FRIDLUND, A. J. Facial displays are tools for social influence. *Trends in Cognitive Sciences 22*, 5 (2018), 388–399.

8. DAILEY, M. N., COTTRELL, G. W., PADGETT, C., AND ADOLPHS, R. Empath: A neural network that categorizes facial expressions. *Journal of cognitive neuroscience 14*, 8 (2002), 1158–1173.

9. DARWIN, C. *The expression of the emotions in man and animals.* John Murray, 1872.

10. EDITORIAL. Time to discuss consent in digital-data studies. *Nature 572*, 7767 (2019), 5.

11. EKMAN, P. An argument for basic emotions. *Cognition & emotion 6*, 3-4 (1992), 169–200.

12. EKMAN, P. Facial expression and emotion. *American psychologist 48*, 4 (1993), 384.

13. EKMAN, P. *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life.* Times Books, 2003.

14. EKMAN, P. What scientists who study emotion agree about. *Perspectives on psychological science 11*, 1 (2016), 31–34.

15. ELFENBEIN, H. A., AND AMBADY, N. On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychological bulletin 128*, 2 (2002), 203.

16. FLORIDI, L. Translating principles into practices of digital ethics: Five risks of being unethical. *Philosophy & Technology 32*, 2 (2019), 185–193.

17. FRIDLUND, A. J. *Human facial expression: An evolutionary view.* Academic Press, 1994.

18. GENDRON, M., ROBERSON, D., VAN DER VYVER, J. M., AND BARRETT, L. F. Perceptions of emotion from facial expressions are not culturally universal: evidence from a remote culture. *Emotion 14*, 2 (2014), 251.

19. HEAVEN, D. Why faces don't always tell the truth about feelings. *Nature 578*, 7796 (2020), 502–504.

20. HIGHFIELD, R., WISEMAN, R., AND JENKINS, R. How your looks betray your personality. *New Scientist 201*, 2695 (2009), 28–32.

21. KANADE, T., COHN, J. F., AND TIAN, Y. Comprehensive database for facial expression analysis. In *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)* (2000), IEEE, pp. 46–53.

22. LEYS, R. *The ascent of affect: Genealogy and critique.* University of Chicago Press, 2017.

23. LYONS, M., AKAMATSU, S., KAMACHI, M., AND GYOBA, J. Coding facial expressions with gabor wavelets. In *Proceedings Third IEEE international conference on automatic face and gesture recognition* (1998), IEEE, pp. 200–205.

24. LYONS, M., KAMACHI, M., GYOBA, J., AND AKAMATSU, S. Gabor wavelet representation of facial expression. *Technical report of IEICE. HIP 97*, 117 (jun 1997), 9–15.

25. LYONS, M., KAMACHI, M., TRAN, P., GYOBA, J., AND AKAMATSU, S. V1 similarity measure recovers dimensions of facial expression perception. *Investigative Ophthalmology & Visual Science 38*, 4 (1997).

26. LYONS, M., AND MORIKAWA, K. A model based on V1 cell responses predicts human perception of facial similarity. *Investigative Opthalmolgy and Visual Science 37*, 910 (1996).

27. LYONS, M., MORIKAWA, K., AND AKAMATSU, S. Gabor-based coding and facial similarity perception. *Perception 26*, 1_suppl (1997), 250.

28. LYONS, M. J. Dimensional affect and expression in natural and mediated interaction. *arXiv preprint arXiv:1707.09599* (2007).

29. LYONS, M. J., MORIKAWA, K., AND AKAMATSU, S. A linked aggregate code for processing faces. *Pragmatics & Cognition 8*, 1 (2000), 63–81.

30. MARTINE, A. *Martine's Hand-book of Etiquette: And Guide to True Politeness.* Dick & Fitzgerald, 1866.

31. PHILLIPS, P. J., MOON, H., RIZVI, S. A., AND RAUSS, P. J. The feret evaluation methodology for face-recognition algorithms. *IEEE Transactions on pattern analysis and machine intelligence 22*, 10 (2000), 1090–1104.

32. PRABHU, V. U., AND BIRHANE, A. Large image datasets: A pyrrhic win for computer vision? *arXiv preprint arXiv:2006.16923* (2020).

33. RUSSELL, J. A. A circumplex model of affect. *Journal of personality and social psychology 39*, 6 (1980), 1161.

34. RUSSELL, J. A. Core affect and the psychological construction of emotion. *Psychological review 110*, 1 (2003), 145.

35. SCHLOSBERG, H. The description of facial expressions in terms of two dimensions. *Journal of experimental psychology 44*, 4 (1952), 229.

36. SOLON, O. Facial recognition's 'dirty little secret': Millions of online photos scraped without consent. *NBC News* (2019). [Online; accessed 27-August-2020].

37. SRINIVASAN, R., AND MARTINEZ, A. M. Cross-cultural and cultural-specific production and perception of facial expressions of emotion in the wild. *IEEE Transactions on Affective Computing* (2018).

38. TEDONE, G. From spectacle to extraction. and all over again. *unthinking photography* (2020). [Online; accessed 27-August-2020].

39. TOMKINS, S. S. *Affect imagery consciousness: Volume I: The positive affects*, vol. 1. Springer publishing company, 1962.

40. WANG, Y., AND KOSINSKI, M. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of personality and social psychology 114*, 2 (2018), 246.

41. WUNDT, W. *Outlines of psychology.* Leipzig: Wilhelm Engelmann, 1897.

42. ZHANG, Z., LYONS, M., SCHUSTER, M., AND AKAMATSU, S. Comparison between geometry-based and gabor-wavelets-based facial expression recognition using multi-layer perceptron. In *Proceedings Third IEEE International Conference on Automatic face and gesture recognition* (1998), IEEE, pp. 454–459.

43. ZUBOFF, S. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power: Barack Obama's Books of 2019.* Profile Books, 2019.