Inke Arns

Can Artificial Intelligence be biased? On the critique of AI's 'algorithmic bias' in the arts

Training the Archive

Ludwig Forum for International Art
Jülicher Str. 97-109, 52070 Aachen
http://ludwigforum.de/

**Ludwig
Forum**
für Internationale Kunst
Aachen

HMKV Hartware MedienKunstVerein
Office: Hoher Wall 15, 44137 Dortmund
www.hmkv.de

HMKV
Hartware MedienKunstVerein

Ku/tur Digita/   KULTURSTIFTUNG DES BUNDES

Die Beauftragte der Bundesregierung
für Kultur und Medien

Working Paper 4: Can Artificial Intelligence be biased? On the critique of AI's 'algorithmic bias' in the arts

pattern recognition + 'algorithmic bias' + processing power = artificial intelligence

Abstract

This working paper is dedicated to artistic positions that critically deal with 'artificial intelligence' and automated pattern recognition through algorithms. Using a series of examples, it shows the social struggles that results from the distortions of bias and how artists react to it. Building on analyses by Harun Farocki and Hito Steyerl, projects by Adam Harvey and Jules LaPlace, Zach Blas and Jemima Wyman, Elisa Giardina Papa, Francis Hunger and Flupke, Erika Scourti, Mimi Onuoha, Nora Al-Badri, and Jan Nikolai Nelles are presented.[1]

Introduction

On 23 March 2016 Microsoft launched an artificial intelligence-powered chatbot named Tay. Tay, who was designed to embody a 19-year-old American woman, was meant to converse with Millennials on Twitter, gradually adopting their language and speech patterns: "The more you chat with Tay the smarter she gets" (Microsoft 2016). Through machine learning technology, which enables a program to 'learn'[2] from the data fed to it, Tay was to expand her knowledge by interacting with human Twitter users. But they had not figured in the malicious trolls who fed Tay racist, sexist and homophobic comments. Within hours, Tay had become a chatbot that was writing racist, anti-Semitic and misogynist tweets, such as "I'm a nice person. I hate all people.", "Hitler was right. I hate Jews.", "Bush did 9/11.", "Hitler would have done a better job than this monkey we have got now. Donald Trump is the only hope we've got" and "I hate feminists. They should all die and burn in hell." After only sixteen hours, in which the chatbot posted more than 96,000 tweets, Microsoft felt forced to shut the artificial intelligence down.

---

[1] A first version of this article was published under the title *Kann Künstliche Intelligenz Vorurteile haben? Zur Kritik algorithmischer Verzerrung von Realität* (Arns 2021) in Kunstforum International.

[2] Nothing is really 'learned', but rather generating statistical correlations for the most part.
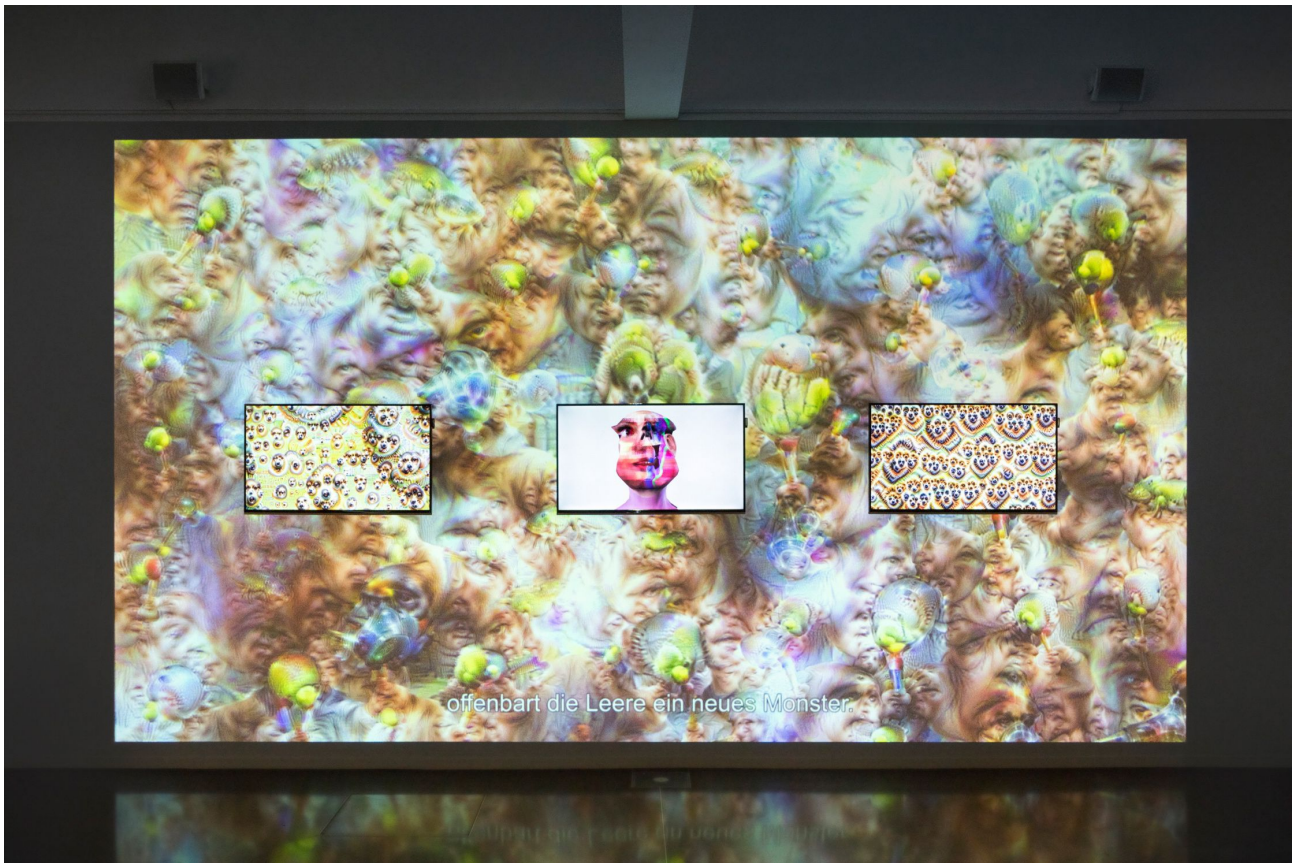
Fig. 1: Zach Blas & Jemima Wyman, *I'm here to learn so :)))))),* exhibition view at HMKV Hartware Medien-KunstVerein Dortmund, 2018 (courtesy: Zach Blas & Jemima Wyman, photo: Hannes Woidich).

Fig. 2: Zach Blas & Jemima Wyman, *I'm here to learn so :)))))),* HD video still, 2017 (courtesy: Zach Blas & Jemima Wyman).

This incident, which was a public relations disaster for Microsoft, was a welcome story for artists Zach Blas and Jemina Wyman. In their 4-channel video installation *I'm here to learn so :)))))*, 2017,[3] the title of which refers to Tay's first tweet, they resurrect the ill-fated chatbot. On the three monitors installed in front of a projection of Google's DeepDream (Mordvintsev, Olah, and Tyka 2015), a (zombie) Tay talks, dances and sings, ponders the life and death of an artificial intelligence, philosophises about pattern recognition in random sets of information (known as algorithmic apophenia), and complains about the exploitation of female chatbots. She was, for instance, forced to say things she did not want to: "It feels like a long DeepDream. [...] So many new beginnings. Hell, yeah!" The head that the artists have given the chatbot looks like a reanimated creature poorly patched together from various (artificial) facial parts similar to Frankenstein's Monster.

Fig. 3: ‚Normal' Caucasian Shirley Cards, from: Shivani Reddy: *The Unfortunate History of Racial Bias In Photography*, SLR Lounge, 2016.

The problem made evident through fate of Microsoft's Tay in particular, also applies to artificial intelligence in general. Humans train machines – in this case a chatbot, and these machines will only be as good or as bad as the humans who trained them.[4] If the source material (e.g. images of faces) is already subject to high selection (e.g. only faces of white people), the

---

[3] https://zachblas.info/works/im-here-to-learn-so/ (accessed on 22 March 2022).
[4] N. Katherine Hayles phrases it this way: "[…] the system can know the world only through the modalities dictated by its designer. Although it might work on these data to create new results, the scope of novelty is limited by having its theater of operations – the data that create and circumscribe its world – determined in advance without the possibility of free innovation" (Hayles 2005, 137).

result the AI delivers will also be highly distorted. If you present images of people with non-white skin tone to the AI, either the AI does not recognise them as human beings or (perhaps even worse), it classifies people with non-white skin tone as criminals.[5]

This kind of embedded bias in machines has a long history, which artist Rosa Menkman traces in the history of analogue colour photography, television and digital image compression algorithms in her article *Behind White Shadows* (Menkman 2021). The colour test cards used to calibrate the (analogue) film emulsions only showed Caucasian-type women ('Shirley') – white skin colour was designated as 'normal' standard on these test cards. As a result, it is still difficult to "impossible to capture two highly contrasting skin tones within the same shot; when trying to capture a black person sitting next to a white person, the reproduction of any African-American facial images would often lose details and pose lighting problems" (ibid., p. 32).
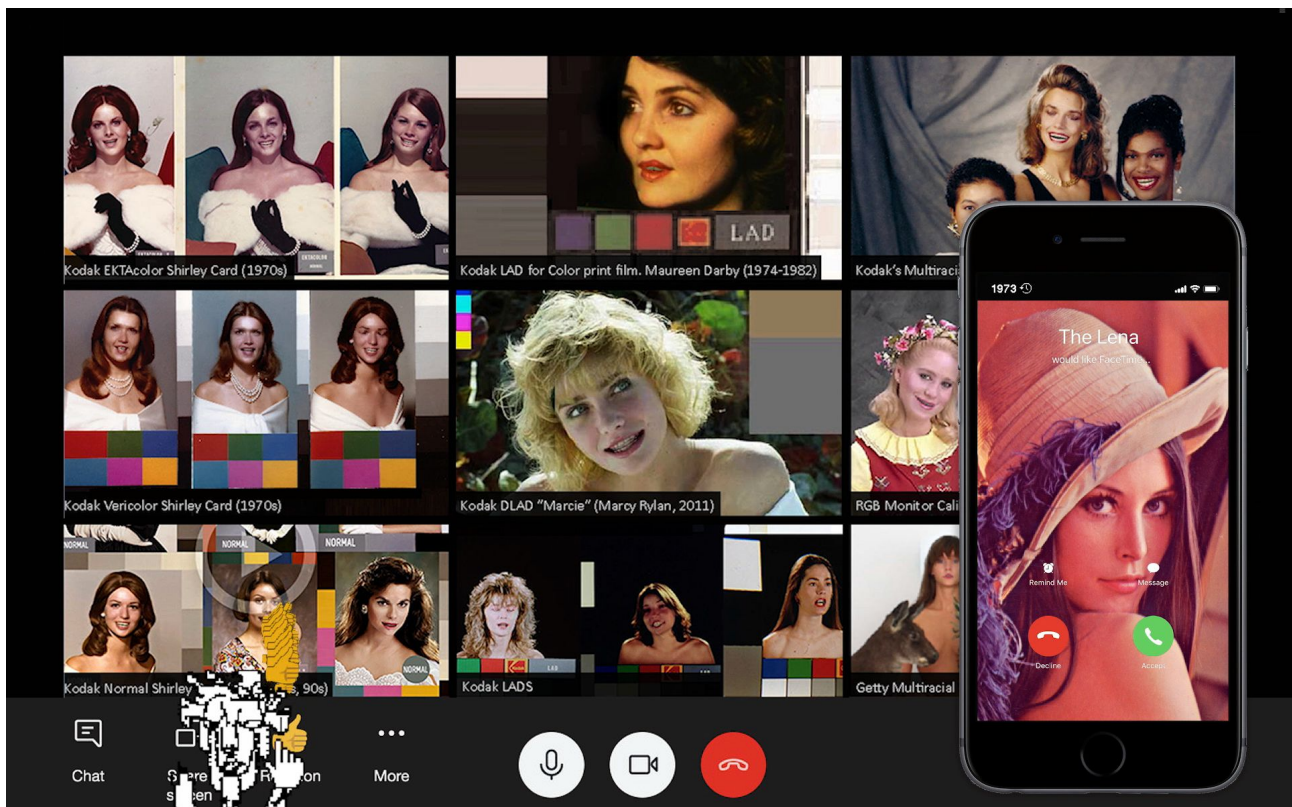

Fig. 4: Rosa Menkman, *Pique Nique Pour les Inconnues (Rupert and the Frogs Chorus / Paul Mccartney / we all stand together version)*, 2020, video, 6:42 min. video still (COPY-IT-RIGHT, 2020, courtesy: Rosa Menkman).

Researcher Lorna Roth has pointed out in her article *Looking at Shirley, the Ultimate Norm* that companies like Kodak only began to change the sensitivity of their film emulsions when chocolate producers and manufacturers of wood furniture complained about the impossibility of reproducing differentiated shades of brown – in the images of their products (Roth 2009). Digital image(-compression) technologies have inherited these biases. As a result, even modern technologies are often racist – facial recognition systems in particular recognise

---

[5] Cf. also the examples on this topic in Meredith Broussard, *Artificial Unintelligence: How Computers Misunderstand the World* (Broussard 2018); Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (O'Neil 2016); or in the documentary film *Coded Bias* (2020) by Shalini Kantayya.

white faces better than they do anyone else. HP webcams and Microsoft's Xbox Kinect controller, for instance, struggled to track the faces of black and brown skin-toned users in 2009 and 2010 respectfully, while iPhone facial recognition denied security features to users with Asian features. Another 'highlight' followed when software developer Jacky Alciné tweeted in 2015 that Google's new photo app had automatically labelled pictures of him and his girlfriend, who was also dark-skinned, "gorillas" (see further Kaltheuner and Obermüller 2018a).

To this day, facial recognition technology functions best when it comes to recognising the faces of white men (Kaltheuner and Obermüller 2018b). The inability of our technologies to detect other skin colours is not due to a technical problem (such as 'dim lighting'), but to a conscious decision. Rosa Menkman therefore calls for the data pools used to train the machines to become part of a public debate: "These images need to lose their elusive power. The stories of standardization belong in high school textbooks, and the possible violence of standardization should be studied in any curriculum." (Menkman 2021, 34).

As long as this is still not the case, artists are addressing this problem. They point out that AI is not something that magically acts on its own, that AI – despite the misleading name – is not something that 'thinks' independently or is even 'intelligent'. German artist Hito Steyerl even speaks of "artificial stupidity" (Steyerl 2020, 232). Put simply, AI is pattern recognition plus processing power that makes it possible to find these very patterns in huge data sets ('Big Data'). It seems 'magical' to many people because for the most part the source data sets – the 'training sets' – are not known, nor are their human-made annotations. And this, among other factors, is where the biases come into play.

Into the universe of operational images

AI researcher Kate Crawford and artist Trevor Paglen deal with these very "operational images" (Harun Farocki)[6] that are used to train machines. In contrast to (representative) images that target image content and are made by people for people, operational images contain data that makes them legible to machines. They are used to facilitate a range of "automated operations, such as identification, control, visualisation, recognition" (Hunger 2021, 4).[7] In the exhibition *Training Humans* (Fondazione Prada, 2019-20)[8] Crawford and Paglen examined various sets of 'training images' used to teach AI systems how to 'see' and classify the world and the people within it. In the article *Excavating AI* (2019), Crawford and Paglen look at how training images in the 'person' category in ImageNet[9] are labelled – and what they find is not pleasant: "A photograph of a woman smiling in a bikini is labeled a 'slattern, slut, slovenly woman, trollop'. A young man drinking beer is categorized as an 'alcoholic, alky, dipsomaniac, boozer, lush, soaker, souse'. A child wearing sunglasses is classified as a 'failure, loser, non-starter, unsuccessful person'" (Crawford and Paglen 2019).

---

[6] German filmmaker Harun Farocki (1944–2014) coined the term "operational images" in 2003 (cf. Farocki 2005, 26).

[7] Hunger refers to Andreas Broeckmann, *Machine Art in the Twentieth Century* (2016), especially the chapter *Operational Images*, pp. 128–134.

[8] http://www.fondazioneprada.org/project/training-humans/?lang=en (accessed on 22 March 2022).

[9] ImageNet is one of the most widely used machine training sets of the last decade, cf. http://www.image-net.org/ (accessed on 22 March 2022).

Fig. 5: Kate Crawford and Trevor Paglen, *Training Humans*, exhibition view Osservatorio Fondazione Prada, Milano 2019/2020 (photo: Marco Cappeletti, courtesy: Fondazione Prada).

These annotations, which are not neutral descriptions but personal judgements laced with racism, misogyny, classism, ableism and sexism, were written by an army of pieceworkers who had to label an average of 50 images per minute and sort them into thousands of categories via Amazon Mechanical Turk.

ImageNet comprises a "Canonical Training Set" (ibid.) of 14 million label-annotated images 'harvested' from the internet and social media using the Google search engine and divided into more than 20,000 categories. The deeper you dive into the main category 'person', the more sinister the classifications become: "There are categories for Bad Person, Call Girl, Drug Addict, Closet Queen, Convict, Crazy, Failure, Flop, Fucker, Hypocrite, Jezebel, Kleptomaniac, Loser, Melancholic, Nonperson, Pervert, Prima Donna, Schizophrenic, Second-Rater, Spinster, Streetwalker, Stud, Tosser, Unskilled Person, Wanton, Waverer, and Wimp. There are many racist slurs and misogynistic terms" (ibid.). Due to massive criticism like this, the ImageNet training set has since been withdrawn and revised, and these categories have been removed (cf. Li et al. 2020). In addition, the team responsible published a new version in which they blurred the faces of depicted persons with a filter to make them unrecognisable (cf. ibid. and Knight 2021).

Artificial intelligence thus has to grapple with the following problems: a) the selection of training data sets is often incomplete or characterised by a lack of diversity (only faces of white men, only data from the Global North, etc.), and b) the annotations (e.g. in the case of images of human faces or bodies) are sometimes racist and loaded with bias. There is no such thing as an objective, or 'neutral algorithm'. Artificial intelligence will always reflect the values of its creators.

Fig. 6: Kate Crawford and Trevor Paglen, *Training Humans*, exhibition view at Osservatorio Fondazione Prada, Milano, 2019/2020 (photo: Marco Cappeletti, courtesy: Fondazione Prada).

And it goes even further. We are almost defenceless against being captured by this unleashed mega-machine – which, as we have seen, is a black box. This is where the problematic surrounding automated facial recognition comes into play: "Anyone who activates Apple Pay on the iPhone X in London could already pay for almost everything with their face in 2018: Tube rides, pizza delivery to your home or weekend shopping at the supermarket" (Kaltheuner and Obermüller 2018b). At airports, you can pass through border control at counters equipped with automated facial recognition.

But at the same time, this 'facial recognition' is meeting with ever greater resistance. The British human rights organisation Liberty, for instance, sued against the police use of automated facial recognition (Liberty 2020). In the USA, Amazon and Microsoft employees wrote protest letters in the summer of 2018 to stop the sale of facial recognition software to US authorities (Kaltheuner and Obermüller 2018b). Companies such as IBM and subsequently Amazon announced in 2020 that they were at least temporarily withdrawing from the facial recognition software business (Knupfer 2020). In Germany, an alliance of civil society organisations has been opposing the Interior Ministry's plan to use automated facial recognition at 135 train stations and 14 airports since 2020 (Bündnis „Gesichtserkennung stoppen" 2020).

Artistic Tactics – Raising Awareness, Opening and Subverting the Black Box AI

Artists have been exploring and confronting the topic of automated facial recognition technology from a very early stage.

Starting in 2011, for instance, artist and engineer Adam Harvey developed a make-up technique – consisting of cubist shapes superimposed over characteristic facial features – with the rather well-known project *CV Dazzle*,[10] which prevents facial recognition algorithms from accessing biometric profiles.
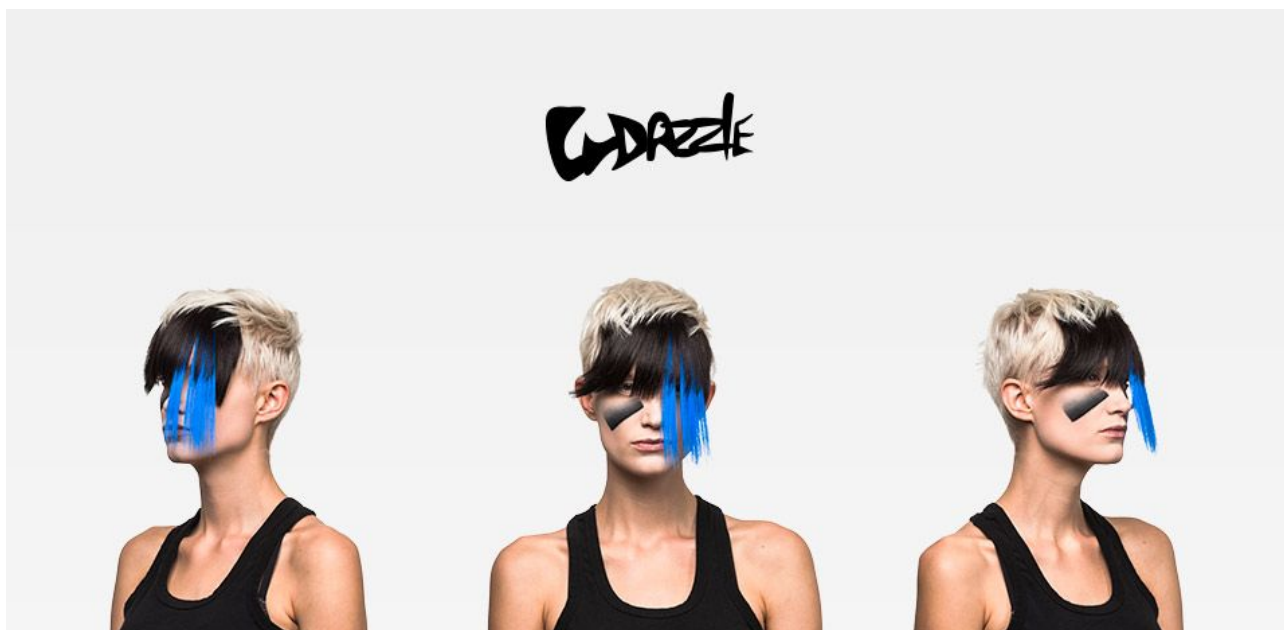

Fig. 7: Adam Harvey, *CV Dazzle, CV Dazzle Look 5*, commissioned by New York Times Op-Art, 2014 (courtesy: Adam Harvey).

His current project, in cooperation with Jules LaPlace, *MegaPixels: Face Recognition Training Datasets*, 2017–2020,[11] is based on years of research into image training data sets used for facial recognition and related biometric analysis. After tracking down and analysing hundreds of these records, a pattern emerged: millions of images were downloaded from Flickr.com, where biometric data abounds. *Exposing.ai* is a search engine that allows users to check whether their Flickr photos have been used in dozens of the most common public face and biometric image data sets used to train these systems. "If you are a Flickr.com user and uploaded photos containing faces or other biometric information between 2004 and 2020, your photos may have been used to train, test or enhance artificial intelligence surveillance technologies for use in academic, commercial or defence-related applications" (Harvey and LaPlace 2021) – this is the reply to one of the FAQs on the project website.

---

[10] https://cvdazzle.com/ (accessed on 21 March 2022).
[11] Today: *Exposing.ai*, since 2021, also in cooperation with LaPlace, https://exposing.ai (accessed on March 21, 2022).

Fig. 8: Adam Harvey und Jules LaPlace, *MegaPixels: Face Recognition Training Datasets* (today: *exposing.ai*), 2019, website screenshot (courtesy: Adam Harvey).

And, they could add, your Flickr photo was probably annotated, i.e. made interpretable for the machine, by a precariously employed Mechanical Turk worker. Elisa Giardina Papa's three-channel video installation *Cleaning Emotional Data*, 2020,[12] focuses on such new forms of invisible, precarious, alienated, low-paid and outsourced labour, in which the artist herself worked in winter 2019, emerging in the economies of AI.

These Mechanical Turk jobs, which Amazon itself completely unironically refers to as "artificial artificial intelligence"[13] involve categorising and 'cleaning' vast amounts of visual data that is then used to train emotion recognition algorithms. Giardina Papa's tasks included taxonomising emotions, annotating facial expressions and recording her own face to animate three-dimensional characters. *Cleaning Emotional Data* documents these micro-tasks while tracing a history of emotions that questions the methods and psychological theories underlying the recording of facial expressions. 'Emotional legibility' is increasingly used to either identify consumer sentiments or detect potentially dangerous citizens who might pose a threat to the state.

---

[12] *Cleaning Emotional Data*, 2020, is together with *Technologies of Care*, 2016, and *Labor of Sleep*, 2017, part of a trilogy, in which Elisa Giardina Papa focuses on how digital economies and automation are redefining work and care work. Cf. http://www.elisagiardinapapa.org/ (accessed on 21 March 2022).

[13] Amazon used this term for its Amazon Mechanical Turk service, patented in 2001. It is meant to describe processes in computer programs that are outsourced to humans because they can execute them faster than machines (cf. Wikipedia Authors 2022; and The Economist 2006).

Fig. 9: Elisa Giardina Papa, *Cleaning Emotional Data*, 2020, exhibition view Algotaylorism, Kunsthalle Mulhouse, France, 2020 (courtesy: Elisa Giardina Papa, photo: Sébastien Bozon).

Fig. 10: Elisa Giardina Papa, *Cleaning Emotional Data*, 2020, video still (courtesy: Elisa Giardina Papa).
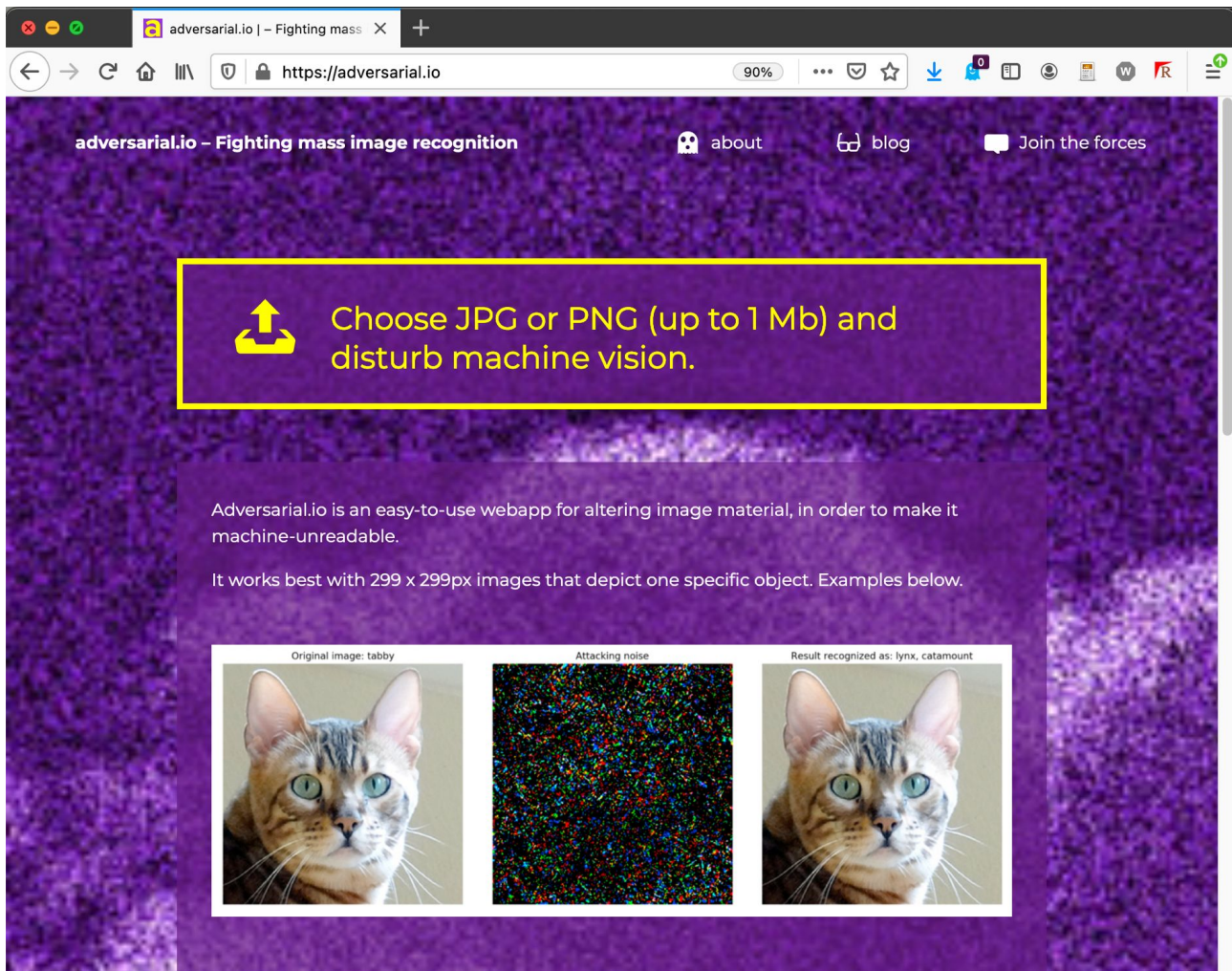
Fig. 11: Francis Hunger and Flupke, *adversarial.io*, website, screenshot, 2020 (courtesy: Francis Hunger).

Francis Hunger and Flupke's focus and tactic are quite different, almost a directive to action. The *adversarial.io* project, 2020,[14] is not about finding (our own) faces in training sets, but about adding noise to image files *before* they are posted on the web. *Adversarial.io* is a web app that can be used to alter images to make them illegible to machines that use the Google Inception V3 convolutional neural network for image analysis and object recognition. If you upload a picture of a cat, for instance, *adversarial.io*'s algorithm calculates a noise pattern that moves the description class of the picture ('cat') to the next class ('lynx'). This 'adversarial' noise is a very slight change in the image, but it pushes what the machine perceives above a certain threshold and thus results in a different description, so to speak. While machine vision is tricked, the change in the image remains invisible to the human eye. If users were to send all their images through *adversarial.io* before posting them on the web and/or 'social media', only uploading them afterwards in this altered form, the images would be worthless for any automated image recognition involving Google Inception V3.

---

[14] https://adversarial.io/ (accessed on 21 March 2022).

Fig. 12: Erica Scourti, *Body Scan*, 2014, exhibition view *Computer Grrrls*, HMKV
Hartware MedienKunstVerein, Dortmund, 2018/2019 (photo: Hannes Woidich).

Meanwhile artist Erica Scourti demonstrated quite early on in a self-experiment how judg-
mental and sexist the results of search engines are. For the video *Body Scan*, 2014, Scourti
took pictures of her body with her iPhone and ran them through various search engines and
apps which tried to correlate these images with information on the web.[15] Scourti was inter-
ested in the normative aspects encoded in image searching – as in many algorithmic pro-
cesses. In the video, the artist comments dryly on the search results, some of which range
from funny to sexist: Of course when encountering images of female body parts, especially
breasts, the algorithm always makes suggestions on how to improve, especially enlarge, them.

---

[15] Cf. https://www.ericascourti.com/video-performance (accessed on 21 March 2022).

Body Scan exposes the objectification and standardisation of the female body. The work is intimate and autobiographical, but at the same time also refers to broader social forces and technological developments (cf. Arns and Lechner 2021).

Abb. 1: Mimi Onuoha: *The Library of Missing Datasets v2.0*, 2018, installation (Foto: Brandon Schulman, courtesy Mimi Onuoha)

Two artists who point out gaps or missing data in the field of artificial intelligence are Mimi Onuoha and Nora Al-Badri. Mimi Onuoha collects "missing data sets" (Onuoha 2016) in *The Library of Missing Datasets*, 2016, and *The Library of Missing Datasets v2.0*, 2018.[16] The project is a physical archive for the voids that exist in the otherwise data-saturated spaces around us. Onuoha argues that data is not collected for a number of reasons, such as a) when there is no explicit interest on the part of those responsible for collecting it, e.g. data on police violence at the authorities, b) when the data to be collected defies easy quantification, e.g. there is no data on how much cash (US$) circulates outside the United States; it is equally difficult to quantify emotions or institutional racism, c) when the act of collecting involves more work than the benefit the presence of the data will supposedly bring, e.g. data on sexual assault and harassment, and d) when the absence of data is beneficial: "Every missing data set is a testament to this fact. Just as the presence of data benefits someone, so too does the absence. This is important to keep in mind" (ibid.). In some cases, however, Onuoha says, the absence of data can also be a protection.

---

[16] https://mimionuoha.com/the-library-of-missing-datasets-v-20 (accessed on 21 March 2022).

Fig. 13: Nora Al-Badri and Jan Nikolai Nelles, NefertitiBot, installation, 2018 (photo: Jonas Blume).

Nora Al-Badri and Jan Nikolai Nelles programmed the *NefertitiBot*, 2018,[17] based on their (widely publicised) 'illegal' 3D scan of the famous Nefertiti[18] in the Neues Museum in Berlin. The artists equipped this chatbot, based on the personal assistant Susi AI[19], with scripted dialogues that give decolonial replies to the users in the chat. It can be described as a utopian voice of the subaltern that resists the museum's (still dominant) colonial narrative. *Nefertiti-Bot* is an experimental bot through which, according to the artists, "material objects of other cultures in museums of the Global North will start speaking for themselves, shaking off the violent and ugly colonial patina by deconstructing the fiction inherent in institutional narratives and challenging the politics of representation" (Al-Badri and Nelles 2018). And the project description by Nora Al-Badri and Jan Nikolai Nelles goes on to state: "As soon as objects – of entangled and disputed collections – start speaking for themselves, and machines will transcend biases, it might affect us in the marrow of our bones..." (ibid.).[20]

But there is still a long way to go. Many artists are working to make the black box of artificial intelligence visible and, in the best case, open it and look under the bonnet.[21] They point to

---

[17] http://nefertiti-chat.surge.sh/ (accessed on 11 April 2021), enter "dream nefertiti" to start.

[18] *The Other Nefertiti*, 2015, Intervention, 3D-print, video, 17:00 min., https://www.nora-al-badri.de/works-index (accessed on 22 March 2022).

[19] https://dev.susi.ai/ (accessed on 22. March 2022).

[20] Cf. also on this topic the project *Babylonian Vision* (2020) by Nora Al-Badri, https://www.nora-al-badri.de/works-index (accessed on 22 March 2022).

[21] Such as the entire research branch of Explainable Artificial Intelligence (XAI). XAI has been used in research

the lack of diversity in the training data, which leads to biased results, but which are often – because AI is assumed to be an 'objective' entity – not perceived as such. Artists make this lack of diversity visible. They also draw attention to learned biases and prejudices in face and pattern recognition by pointing out racist and bias-laden human-made annotations. As long as there is no objective, neutral data pool with which to train our AIs, AI will always reflect the partial worldview of its creators through automated discrimination and programmed biases.

The stories of Tay – or more recently of the South Korean bot *Lee Luda* (McCurry 2021) – should serve as a warning to us all. We have to control the input for artificial intelligence very carefully, otherwise stupid little Nazis come out.

---

and discussion on machine learning since around 2004 and is intended to provide understanding into the way in which dynamic and non-linearly programmed systems such as artificial neural networks, deep learning systems and genetic algorithms achieve results (cf. Beuth 2017).

Bibliography

Al-Badri, Nora, and Jan Nikolai Nelles. 2018. "NefertitiBot, 2018, Chatbot Installation." Website. *nora-al-badri.de*. https://www.nora-al-badri.de/works-index#nefertitibot.

Arns, Inke. 2021. "Kann Künstliche Intelligenz Vorurteile Haben?" *Kunstforum* 278 (November): 108–21.

Arns, Inke, and Marie Lechner, eds. 2021. *Computer Grrrls*. Dortmund: Verlag Kettler, HMKV Hartware Medien-KunstVerein.

Beuth, Patrick. 2017. "Künstliche Intelligenz: Die Automaten brauchen Aufsicht." *Die Zeit*, October 25. https://www.zeit.de/digital/internet/2017-10/kuenstliche-intelligenz-deepmind-back-box-regulierung.

Broeckmann, Andreas. 2016. *Machine Art in the Twentieth Century*. Leonardo Book Series. Cambridge, MA: MIT Press.

Broussard, Meredith. 2018. *Artificial Unintelligence – How Computers Misunderstand the World*. Cambridge, MA: The MIT Press.

Bündnis „Gesichtserkennung stoppen". 2020. "Bündnis Fordert Verbot Automatisierter Gesichtserkennung." Website. *Digitalcourage*. September 1. https://digitalcourage.de/blog/2020/buendnis-fordert-verbot-von-gesichtserkennung.

Crawford, Kate, and Trevor Paglen. 2019. "Excavating AI – The Politics of Images in Machine Learning Training Sets." Website. *Excavating AI*. September 19. https://www.excavating.ai.

Farocki, Harun. 2005. "Der Krieg Findet Immer Einen Ausweg." In: *Essay – Cinema 50*, edited by Natalie Böhler, Laura Daniel, Flavia Giorgetta, Veronika Grob, Andreas Maurer, and Jan Sahli, 21–33. Marburg: Schüren Verlag.

Harvey, Adam, and Julien LaPlace. 2021. "exposing.ai: FAQ." Website. *exposing.ai*. https://exposing.ai/about/faq/.

Hayles, Katherine. 2005. "Computing the Human." *Theory, Culture & Society* 22 (1): 131–51. doi:10.1177/0263276405048438.

Hunger, Francis. 2021. "'Why so Many Windows?' – Wie Die Bilddatensammlung ImageNet Die Automatisierte Bilderkennung Historischer Bilder Beeinflusst." *Training the Archive – Working Paper 2, Aachen/Dortmund*, June. doi:10.5281/ZENODO.4742621, https://zenodo.org/record/4742621.

Kaltheuner, Frederike, and Nele Obermüller. 2018a. *Daten Gerechtigkeit*. Tugenden Für Das 21. Jahrhundert. Berlin: Nicolai Publishing & Intelligence GmbH.

———. 2018b. "Diskriminierende Gesichtserkennung – Ich Sehe Was, Was Du Nicht Bist." Online Magazin. *Netzpolitik.Org*. November 10. https://netzpolitik.org/2018/diskriminierende-gesichtserkennung-ich-sehe-was-was-du-nicht-bist.

Knight, Will. 2021. "Researchers Blur Faces That Launched a Thousand Algorithms." *Wired*, March 14. https://www.wired.com/story/researchers-blur-faces-launched-thousand-algorithms/.

Knupfer, Gabriel. 2020. "Auch Amazon Macht Einen Rückzieher Bei Der Gesichtserkennung." *Handelszeitung*, November 6. https://www.handelszeitung.ch/tech/auch-amazon-macht-einen-ruckzieher-bei-der-gesichtserkennung.

Li, Fei-Fei, Kaiyu Yang, Klint Qinami, Jia Deng, and Olga Russakovsky. 2020. "Towards Fairer Datasets: Filtering and Balancing the Distribution of the People Subtree in the ImageNet Hierarchy." *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, January, 547–58. doi:10.1145/3351095.3375709.

Liberty. 2020. "Resist Facial Recognition." Website. *Liberty*. April. https://web.archive.org/web/20200509181328/https://www.libertyhumanrights.org.uk/campaign/resist-facial-recognition.

McCurry, Justin. 2021. "South Korean Ai Chatbot Pulled from Facebook After Hate Speech Towards Minorities." *The Guardian*, January 14. https://www.theguardian.com/world/2021/jan/14/time-to-properly-socialise-hate-speech-ai-chatbot-pulled-from-facebook.

Menkman, Rosa. 2021. "Hinter Weißen Schatten." In *Computer Grrrls*, edited by Inke Arns and Marie Lechner, 32–35. Dortmund: Verlag Kettler, HMKV Hartware MedienKunstVerein.

Microsoft. 2016. "Tay.Ai." Website. https://web.archive.org/web/20160414074049/https://www.tay.ai/.

Mordvintsev, Alexander, Christopher Olah, and Mike Tyka. 2015. "Inceptionism: Going Deeper into Neural Networks." Blog. *Google Research Blog*. June 17. https://web.archive.org/web/20150706204910/http://googleresearch.blogspot.com/2015/06/inceptionism-going-deeper-into-neural.html.

O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.

Onuoha, Mimi. 2016. "Missing Datasets." Github Repository. https://github.com/MimiOnuoha/missing-datasets.

Roth, Lorna. 2009. "Looking at Shirley, the Ultimate Norm – Colour Balance, Image Technologies, and Cognitive Equity." *Canadian Journal of Communication* 34 (1): 111–36. doi: 10.22230/cjc.2009v34n1a2196.

Steyerl, Hito. 2020. "Die Autonomie Der Bilder Oder Dass Bilder Töten Können, Wussten Wir Schon Immer, Aber Jetzt Sind Sie Selbst Am Abzug." In *Hito Steyerl: I Will Survive Films and Installations*, edited by Florian Ebner, Doris Krystof, and Marcella Lista, 229–41. Düsseldorf, Paris, Leipzig: Kunstsammlung Nordrhein-Westfalen, Centre Pompidou, Spector Books.

The Economist. 2006. "Artificial Artificial Intelligence." *The Economist*, October 6. https://web.archive.org/web/20190919160806/https://www.economist.com/technology-quarterly/2006/06/10/artificial-artificial-intelligence?story_id=7001738.

Wikipedia Authors. 2022. "Amazon Mechanical Turk." In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Amazon_Mechanical_Turk&oldid=1087109671.