



Author
Daniel C Howe

Location
Hong Kong

Daniel Howe (US/HK) is an artist, researcher, and critical technologist whose work focuses on the interrelationships between networks, language, and politics. His collaborative practice explores the nature of algorithms and their impact on human values like autonomy, diversity, privacy and freedom. Howe is an open-source advocate and contributor to dozens of socially-

engaged software projects over the past two decades. His outputs include software interventions, art installations, algorithmically-generated text and sound, and tools for artists. He currently resides in Hong Kong where he teaches at the School of Creative Media.
www.rednoise.org/daniel

Title
Spectre, Deepfakes and the Contamination of Contamination

Tag
Deepfake
Influence Industry
Surveillance
Artificial Intelligence
Technology

Material



P. 70

Daniel C Howe

Spectre, Deepfakes and the Contamination of Contamination

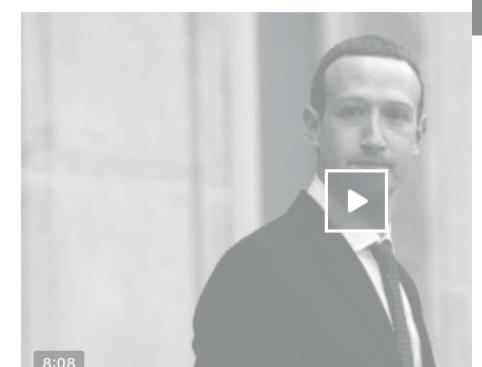
Spectre, by Daniel Howe and Bill Posters, is an immersive installation that attempts to detourn the tactics and technology used by the digital influence industry to manipulate people's behaviours and decision-making. The installation invites visitors to explore the implications of surveillance-driven capitalism and the new religion of Dataism, and to better understand their impacts on trust, privacy and democracy. Below Howe talks about the release of Spectre, the deepfakes from the project, and the events of the last year.

Soon after our prototype of the Spectre installation debuted at the Sheffield DocFest in mid-2019, videos of the deepfake characters who guide visitor through the experience spread rapidly across social media, in what was referred to at the time as 'going viral'. These synthetic celebrities – Mark Zuckerberg, Kim Kardashian, Donald Trump, Damian Hirst and others – and the glimpses into their warped psyches that the project revealed, were covered across the media landscape and even discussed in a session on deepfakes in the US Congress. Of course, we had

After manipulated Zuckerberg, Pelosi videos, lawmakers grasp for 'deepfake' answers

Beyond tech barriers, experts and lawmakers grapple with free speech questions.

By Lee Ferran and Fergal Gallagher
June 14, 2019, 6:58 AM • 10 min read



When seeing isn't believing: using AI to create 'Deepfakes'.
The latest in artificial technology allows creators to swap and meld vi

Days after an artist uploaded a digitally manipulate
appearance to show Facebook CEO Mark Zuckerberg
monologue about "controll[ing] the future," a real-1



BIDEN ADMINISTRATION CORONAVIRUS OPEN SOURCE RECODE THE GOODS FUTURE PERFECT THE HIGHLIGHT MORE

Facebook wants to stay "neutral" on deepfakes. Congress might force it to act.

Mark Zuckerberg is the newest victim of AI-manipulated media, but he won't be the last.

Sigal Samuel | Jun 13, 2019, 3:00pm EDT



Facebook needs a new approach to deepfakes and misinformation. When will it forego its supposed "neutrality" and start taking some actual positions? | Washington Post/Getty Images

FUTURE PERFECT How would you react if you saw a video of Facebook CEO Mark Zuckerberg describing himself grandiosely as "one man, with total control of billions of people's stolen data, all their secrets, their lives, their futures?"

Hopefully, you'd realize that this video — which was posted a few days ago on Instagram, a company owned by Facebook — is a forgery. More specifically, it's a deepfake, a video that has been doctored, using AI, to make it look like someone said or did something they never actually did or didn't.

The Zuckerberg video was created by artists Bill Posters and Daniel Howe together with the real startup Canny as part of a UK exhibition. It's actually not a great deepfake — you can tell that it's a voice actor speaking, not Zuckerberg — but the visuals are convincing enough. The original footage they doctored comes from a 2017 clip of Zuckerberg discussing Russian election interference.

Under the watchful gaze of the global community, each minute that the video stays up serves as a test for the social media company

Katia Shand-Baptiste | @kubared | Wednesday 12 June 2019 16:24 | comments

OO There's delicious irony in all of this. Weeks after (once again) dismissing the dangers of doctored political footage on Facebook, its founder and CEO unwittingly landed the starring role in a misleading video of his own.

Well, not technically his own, it's part of an art installation called Spectre that he had nothing to do with. But that's exactly the point of the video and other "deepfakes" like it — to make us believe that what we're watching came straight from the horse's mouth, as opposed to complex computer programming, or convincing video editing.

To bring you up to speed, towards the end of last month House speaker Nancy Pelosi became the target of an attack on her credibility. Using footage from a speaking engagement at the Centre for American Progress, the altered video convincingly made it seem as if, instead of speaking in the clear and concise manner that we're used to, Pelosi was having trouble speaking altogether, slurring and repeating her words as if under the influence.

News

Jun 13, 2019

Artists created deepfake videos of Mark Zuckerberg to test Facebook's misinformation policies.



The Washington Post

Democracy Dies in Darkness

Get one year for \$29

Sign in



Sections



Privacy





#SpectreKnows

Daniel C Howe

Spectre Deepfakes and the Contamination of Contamination

71

little idea how our understanding of the term ‘viral’ would be transformed in the months to come. As COVID-19 spread across the world and upcoming exhibitions of the installation –in the Americas, Europe and Asia – were cancelled, the deepfakes from the project, a relatively small part of the Spectre experience, became embroiled in a global conversation about AI and media manipulation that still continues. Yet synthetic media, which has long existed in aural and photographic forms was, even then, only one element of the manipulation equation. New tactics and technologies, whether psychologically micro-targeted advertising, professional political manipulators, or toxic social media bubbles have had the potential to impact cultural and political spheres in powerfully negative ways, and were largely absent from the debate at the time of Spectre’s release. Yet with the Brexit referendum close at hand, and the 2020 US election only months off, the stakes were high. The weaponization of targeted disinformation in social media had already been shown to have impacted prior elections (the US presidential election in 2016, for example) as well as public debate on critically important issues like climate change. In a recent study, researchers estimated that on a typical day in the weeks leading up to the Paris Climate Agreement (IPCC 2018), upwards of 25% of all Twitter posts on topics related to climate change were made by automated accounts or bots, defined as “tools designed to generate, disseminate, and promote positions on social media”. Of these the majority were found to be “skeptical or denying of the wide scientific consensus on climate change” (Marlowe 2020).

Considering the magnitude of these issues and the degree to which misinformation can derail public debate, it is not only Spectre deepfake character studies that are relevant to the question of algorithmic manipulation. I focus here instead on a far less-discussed element of the project, specifically the “manipulation game” that forms the overarching narrative for the immersive experience. For visitors to the installation, who have logged in with a social media account and then agreed to the project’s terms-of-service, the game begins when they are tasked with selecting, from a series of (surreptitiously captured) profile-style photographs, a previous user of the system to be the target of an as-yet-unknown series of manipulations. Users are then presented with a series of gamified tasks which focus either on generating additional data about their target, or later, using existing data to influence the target’s beliefs and behaviors. The tasks increase in complexity and potential discomfort, beginning with simple demographic assessments such as age and gender, to evaluating the financial, political and sexual inclinations of the target, to helping facial recognition algorithms more precisely identify emotions in the captured images of their faces. The final task which users (those who have not abandoned the installation – the vast majority do not) is to create a series of social media advertisements, via an interactive design tool, in order to change the target’s intended vote on either the Brexit decision or the 2020 US election, depending on geographic location. Users are tasked to design these advertisements specifically for their target, based on a range of personal data points that the system has accumulated, including the target’s OCEAN personality score, a five-category system believed (and used by marketers (Seimiene

2012)) to accurately model a human subject’s patterns of thoughts, feelings, and behavior (Diener 2019)¹. As in the well-known Milgram experiments (McLeod 2017) the subject of primary interest is not the target here, but rather the visitor, who, by making judgments and selecting strategies, incrementally reveal themselves to Spectre’s cognitive psychological quantification. This fact becomes apparent to visitors as they realize their participation has already enabled their own subjugation to future users of the system, who may choose to target them. By engaging with Spectre, much like the modern-day systems of platform capitalists that it takes as inspiration, you empower it to use you – to forever task others with the relentless penetration of your persona. In this sense Spectre targets the fundamental property of immune systems: the ability to discriminate “self” from “other” so that contaminating bacteria or viruses can be attacked. We might even consider the system to be “autoimmunological” in a sense, exhorting participants to attack their own cells and tissues as if they were foreign threats.

Of course, metaphors of autoimmunity have been discussed extensively in critical theory (though often derided by actual immunologists), notably in the later work of Derrida, developing out his earlier preoccupation with ‘contamination’². He writes “to protect its life, to constitute itself as unique living ego, to relate, as the same, to itself it is necessarily led to welcome the other within... it must therefore take the immune defenses apparently meant for the non-ego, the enemy, the opposite, the adversary and direct them at once for itself and against itself.” (Derrida 1993) An apt description of how the vast systems of computational disinformation, driven by the Zuckerbergs of the world, function (and equally apt in the context of Spectre’s manipulation game). As Derrida notes, such systems are only quasi-suicidal. The contamination of the self is achieved by the self, as other, to an extent that self and non-self can no longer be distinguished. As Peta Mitchell writes, “Autoimmunity in this sense, then, reveals itself as the limit of the law of contamination—indeed as the contamination of contamination.” (Mitchell 2017) The systems of late-stage surveillance capitalism, in digesting the datafied corpses of its happily consenting users, mutate and reproduce in what can perhaps best be described as a metaphorical act of tireless auto-necrophilia.

Sources:

- Derrida, Jacques. 1994. *Specters of Marx*. Vol. 2006, p177. New York: Routledge
Diener, Edward, and Richard E. Lucas. 2019. “Personality traits.” *General Psychology: Required Reading*, p278.
Noba Project.
Marlow, Thomas, Sean Miller, and J. Timmons Roberts. 2020. *Twitter Discourses on Climate Change: Exploring Topics and the Presence of Bots*. SocArXiv
McLeod, S. A. 2017. *The milgram shock experiment*. Simply Psychology. <https://www.simplypsychology.org/milgram.html>. Feb 5, 2017.
Mitchell, Peta. 2017. “Contagion, virology, autoimmunity: Derrida’s rhetorics of contamination.” *Parallax* 23:1 (2017): 77–93.
Seimiene, Eleonora. 2012. “Emotional connection of consumer personality traits with brand personality traits: Theoretical considerations.” *Economics and management* 17.4 (2012): 1472-1478.

¹ It is important to note that the data that the user is provided in these manipulation activities is real data about the target, collected as the installation runs and distilled via a large-scale deep-learning model developed by researchers at the Psychometrics Centre at Cambridge University and trained on over a million human subjects.

² In fact, Derrida’s shift towards immunological tropes in his writings of the 1990s is often considered a marker of his so-called ‘ethico-political’ turn.