

## Handouts don't exist. Hustle or you don't eat.

CONOR MCGARRIGLE

Lecturer in Fine Art, School of Creative Arts, Technological University Dublin

---

### ABSTRACT

It is well established that AI has a bias problem; however, black-boxed machine learning systems render it difficult to even understand and visualize the nature and extent of the problem, let alone find solutions. This paper discusses an artistic research approach toward highlighting AI bias and explores the aesthetic potential of machine learning through a case study of an AI artwork called *#RiseandGrind*. The artist trained a recurrent neural network on a dataset extracted from Twitter hashtags (*#Riseandgrind* and *#Hustle*), which were selected to represent a specific filter bubble (embodied neoliberal precarity) in order to produce a biased AI that generates tweets for a Twitter bot. This paper unpacks how this artwork makes visible the processes of machine learning in a playful and poetic way. The work reveals how the original filter bias is consolidated, amplified, shaped, and ultimately codified through this machine learning process. The AI is found to reproduce a cohesive world view that, while reflecting the original data bias, further amplifies that bias through a process of flattening.

---

### INTRODUCTION

*#RiseandGrind* is a machine learning art project that uses Twitter hashtag data to train a neural network to generate text for a Twitter bot, with the process of training and text generation presented in a series of gallery installations. The work was commissioned by the Science Gallery Lab Detroit in 2018 and exhibited in three versions during 2018 and 2019. The hashtags, *#RiseandGrind* and *#Hustle*, were selected as representative of a filter bubble that I identify as embodied neoliberal precarity, a form of economic self-exploitation arising from an assimilation of the values of the precarious digital economy and characterized by an adherence to the neoliberal principle that economic success or failure is built entirely on individualized effort. The dedication to the hustle, relentless self-promotion, and dispensing of bland life advice in this bubble is total, almost to the point of parody; it is also lively, witty, and anarchic, encapsulating the best and worst of social media.<sup>1</sup>

The project seeks to render aspects of the black-boxed machine learning process visible, to highlight algorithmic bias through demonstrating the training process, and to reflect on the role of artistic practice and research as part of a broader dialogue on machine learning and data while exploring their aesthetic potential. While questions of algorithmic bias are not new,<sup>2</sup> their relevance has not diminished, as illustrated by controversy over the training of Open AI's GPT-2 model on

strongly biased Reddit data.<sup>3</sup> This text discusses the motivation and process of producing an intentionally biased AI that results from training on biased filter bubble data, unpacks the process of training a recurrent neural network (RNN) on Google's TensorFlow from a non-specialist perspective, and renders this process visible through a series of gallery installations.

## Twitter Hashtags

As a core mechanism for coordinating the flows of Twitter conversations, hashtags dynamically form and reform ad hoc publics that assemble and cohere around news and information,<sup>4</sup> common interests, cultural moments, and values.<sup>5</sup> In the words of linguist Michele Zappavigna, hashtags presume the existence of a “virtual community of interested listeners.”<sup>6</sup> Alongside other categorization techniques such as location, followings, and trending subjects, the hashtag is the “killer app” that allows Twitter's users to consume and interact with tweets from users they do not follow or have location in common with, and to engage strangers in conversation on matters of common interest. In addition to organizing informational flows, hashtags play a role in constituting online identities and in co-producing these network identities, an effect particularly notable on Black Twitter.<sup>7</sup> However, they also form filter-bubbles, self-referential immersive information environments; these can be both mutually supportive communities of special interest and echo chambers where ideological positions are rehearsed and reinforced with little outside intervention.<sup>8</sup> This can serve to reinforce and amplify bias and is subject to automated manipulation from bots. Research shows that 9–15% of all Twitter accounts are automated bots.<sup>9</sup>

In addition, hashtags act as a powerful method of categorizing Twitter users' interests for the purposes of surveillance, targeting, and for the accumulation of what Shoshana Zuboff calls “behavioral surplus.”<sup>10</sup> Each Tweet returns comprehensive metadata which typically include 70–80 database fields or 250–300 lines of JSON (JavaScript Object Notation). Twitter packages and sells insight on their users through their data services, with these data accessible at limited scale through their APIs. Twitter is thus an accessible source of current and relevant training data for machine learning applications including sentiment analysis and content generation. The data assemblages of surveillance capitalism target not only the specifics of interests and activities as indicated by followings, likes, retweets, locations, interactions, and hashtag activity, but also through algorithmic sentiment analysis, opinions, and values. The goal of surveillance capitalism, according to Zuboff, is not only one of behavioral prediction but ultimately one of persuasion, to nudge and change behavior based on past actions.<sup>11</sup> The full consequences of building predictive and persuasive models on data that are incomplete and biased, the result of self-replicating internet echo chambers that should not be extrapolated, are widespread but only beginning to be fully understood.<sup>12</sup> Anima Anandkumar, Director of AI Research at Nvidia, in a discussion of the training of Open AI's GPT-2 model on Reddit succinctly summarizes the problem, “when you train on Reddit data, out comes garbage.”<sup>13</sup>

## Training Data

The project began by scraping the #RiseandGrind and #Hustle, hashtags that are not openly politically partisan or controversial and don't play a role in social media culture wars, which is not the same as saying that they are apolitical. In fact, I suggest that they are ideological, espousing a value system that disavows structural inequality to emphasize individualistic self-reliance, where hard work and entrepreneurial hustle are all that it takes to succeed in a neoliberal economy. They

are representative of an economic worldview that I identify as embodied neoliberal precarity; that is, a value-based form of self-exploitation that conflates the requirements and economic values of the precarious gig economy with personal identity and self-realization or individuation, threading a ground that has been previously described as characteristic of the sharing internet economy.<sup>14</sup> As with many hashtags, their usage is complex, with irony and sarcasm juxtaposed with naively bombastic tweets. As expected, automated bot activity is evident at various levels of sophistication, from crude spam hashtag storms to carefully targeted tweets that pass as human. However, the overwhelming impression is of a filter bubble delineated within these hashtags with a clearly articulated message and a cohesive worldview, even if that does not withstand sustained scrutiny. It is important for the project that these hashtags are energetic and entertaining, lively and fun; they are well attuned to the medium of Twitter in form and content, which is part of the reason that these particular hashtags were selected as training data. The project began with questions. What would training an AI produce from these data and what conclusions could be drawn from its results? Would traces of the process of training and adjustments to the training parameters produce aesthetic traces that are unique and characteristic of the process? Could the process of making visible these black box processes in an artwork add to the critical debate on AI and society through introducing this additional perspective that goes beyond the expertise of data scientists?

## Intelligence of Machine Learning

While the term AI is widely accepted, it is important to acknowledge it as a misnomer. AI can be more accurately described as machine learning, typically using convolutional and recurrent neural networks and generative adversarial networks. Current machine learning techniques differ from earlier generations of AI with their focus on creating thinking-machines that would emulate the human brain to create a general artificial intelligence. While this biomimetic terminology persists, the techniques are very different. Machine learning is a probabilistic method that works with statistical correlations and heavy-duty computational power to identify patterns in datasets and encode these into a model that can be used on unseen data to perform its functions; typically this might include text analysis and generation, identifying objects or faces, speech recognition, image generation, machine translation, medical imaging analysis, autonomous vehicles, and of course playing games. Despite the anthropomorphic terminology, this is not general human-level intelligence; it is rather a fast, effective, brute force method that achieves impressive results in narrow domains.<sup>15</sup> Machine learning techniques require high-quality training data in a computationally intensive process with results only as good as the quality of the data input. In many applications, including computer vision and autonomous vehicles, there is a Mechanical Turk aspect to the process as automation is made possible through the efforts of workers painstakingly creating training data by manually labelling millions of images.<sup>16</sup>

Machine learning systems learn through advanced pattern recognition within very large data sets, with these patterns then encoded into the algorithms—the process of training. These algorithms are then applied to data, categorizing them based on these patterns or generating new data based on past actions, the premise being that the larger the data set the more accurate the outputs. Of course, this isn't always true. Machine learning has been found to codify bias, error, racism, and sexism in what philosopher Bernard Stiegler calls functional stupidity or *la bêtise*,<sup>17</sup> with the very effectiveness and opaqueness of these systems militating against the recognition of this bias. As decision-making systems based on machine learning are widely deployed, the flaws in their makeup are emerging. Recent scholarship has revealed these flaws; racist search engine assumptions,<sup>18</sup>

sentencing systems that discriminate against people of color,<sup>19</sup> even facial recognition used to unlock cell phones that are unable to distinguish between Chinese female faces.<sup>20</sup> Problems of bias are complex. In addition to outright bias, racism and sexism bias can originate from failure to recognize the unrepresentative nature of data in what has been described as “privilege hazard.”<sup>21</sup> However, failures in AI are not always readily apparent, and as machine learning systems are integrated into all aspects of life, the act of discrimination and bias is baked into proprietary algorithms forming part of larger processes of algorithmic governance,<sup>22</sup> so that identifying where in a complex process this bias is located is not often possible and actionable. It is against this emergent context that this project was conceived.

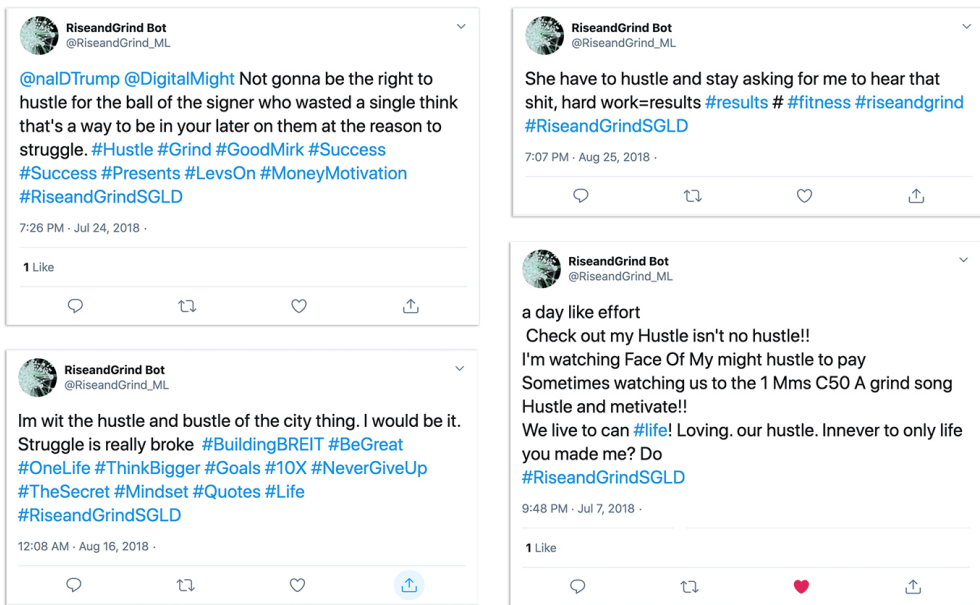
## Training the Neural Network

The project began by amassing a corpus of approximately 600,000 tweets from the hashtags #RiseandGrind and #Hustle, with this dataset used to train a language model on Google’s TensorFlow. The intention was to train a model to generate tweet length text based on this narrowly defined corpus, in order to discover in what ways the generated text would reflect or deviate from the training data. A character-level recurrent neural network (RNN) was employed to model the character probability distribution in the training data, that is the probability that one character will follow another, to produce text that would be tweeted. A detailed discussion of this process is outside the scope of this paper, but it is discussed in greater detail elsewhere.<sup>23</sup> As part of the artwork the network was configured to log the training so that the process at work would be made visible. The network periodically generated random text from the current learned state, progressing from this example from early in the process:

#AtnBwiq nw#thn##aasi a r iaatiy aooornania#ytstytnrr ia#y iirsoytsi #nnr itsnry tt tay y##  
ts t tysiony#yyiay# iatonrot t nniryto#oi#irtyao#rrr oyontyars t n ossaistr o o nniinstainso  
ysotyy#a r#yirtsarii#in r#tyaoniassysioiarsysry # ray yniretsttrysynnati aaann#satr#ytsi  
nyiranyta satyi i naiattrniaa iaannyato nnr# eriarsnasoy a # ni# ai#orni s

to more sophisticated texts that acquire the correct tweet form and mimic aspects of the style and themes from the training but ultimately make no sense (see Figure 1).

Training continued over several iterations to produce results that were at first glance fitting well with the flow of the hashtag conversations. A Twitter bot tweeted texts from all stages of the training on a project account @RiseandGrind\_ML (see Figure 1 and Figure 2). The bot was clearly identified as a bot in its title, RiseandGrind Bot, fitting with the convention of self-identified bots on Twitter, and included a pinned tweet on the account explaining the project. In addition, each tweet added an exhibition hashtag.



***Figure 1. #RiseandGrind, 2018, early generated tweets achieve many aspects of Twitter style but don't make sense***



***Figure 2. #RiseandGrind, 2018, Conor McGarrigue. Sample tweets from fully trained model.***

The final generated tweets had very successfully adopted the style and form of their training data and did not seem out of place as tweets. A sentiment shift was noted with certain politically right-wing tendencies in the training data becoming more pronounced and insistent; the tweets had not only repeated but amplified the bias obvious in the training data, this paper's title a case in point. This amplification of the bias appeared to have come from a flattening out of difference as subtlety and context were erased; the idiosyncratic, the ironic, the linguistic plays, the subtle sub-tweets, and the nuanced weird of the internet were all lost in translation. Without this context and these modifiers, the patterns of the hashtags' text were hollowed-out, one-dimensional exhortations to hustle and grind with none of the idiosyncratic nuance and fun of the originals (Figure 3). Although the texts assumed the form of tweets, they were essentially broadcasting a fixed text that, while generated from a model trained on these hashtags and incorporating many themes evident in the hashtag, nevertheless lacked the capacity to engage with other users, pick up on their cues, or respond appropriately.



***Figure 3. #RiseandGrind, 2018, Conor McGarrigle. Generated tweets from late in training.***

## Creativity and AI

Nonetheless, despite this hollowing out of context, *#RiseandGrind* was producing convincing texts; they were banal, bombastic, annoying, but also witty, risqué, and sometimes just weird. In a relatively short period of training it went from no understanding to generating English (and some Spanish) text that not only adopted the correct tweet format—addressing other users in the correct form, including links with URLs and mostly correct use of hashtags—but also accurately identified and emulated themes evident in the target hashtags. It engaged with other users through the crude mechanism of @ing generated addresses that corresponded to valid accounts and even got blocked for including other accounts in its broadcasts. It was easy to attribute this ability to a deeper understanding of Twitter, that it had in fact learned and was intelligent, even creative. Certainly, there is a human tendency to anthropomorphize machines that are triggered by machine learning,<sup>24</sup> however, what was produced, while effective, was far from intelligent in any real sense. The neural network had no semantic understanding of the texts, it was in fact calculating the probability that one character will follow another using the example of the training text and generating text

accordingly. However, once we stray into the domain of machines that make art there is more at stake than a tendency to see ourselves in machines that emulate us, however crudely.

Creativity implies intelligence, which is why chess as a game requiring analytic ability and creativity has traditionally been the benchmark to test AI; in the 1957 Dartmouth summer project where John McCarthy first coined the term Artificial Intelligence, it was assumed that once a computer beat a chess grandmaster, general AI would have been achieved.<sup>25</sup> When IBM's Deep Blue defeated World Chess Champion Gary Kasparov in 1997 using brute force methods, this special status shifted to Go as, unlike chess, it is too complex to be beaten by brute force methods. In 2017, when Google's Alpha Go beat Go World Champion Lee Sedol, it briefly seemed that the promise of general artificial intelligence was fulfilled. Mathematician Marcus du Sautoy even identifies move thirty-seven of game two as the moment of creativity.<sup>26</sup> We can recognize this as the Skynet moment from the Terminator films, when the military defense system becomes sentient at precisely 2.14 a.m. on August 29 1997 and decides to wipe out the human race—which is essentially the thrust of philosopher Nick Bostrom's best-selling *Superintelligence*, which sees an awakening of machine sentience as an existential danger to humankind.<sup>27</sup>

Machines that make art form part of this narrative, with the co-option of artistic creativity as proof of concept acting as a proxy for general artificial intelligence. While a more sustained discussion into the nature of artistic creativity, AI, and its relation to generative tradition in art is beyond the scope of this article, we can say that direct claims for AI-generated art as creativity qua intelligence demonstrate a misunderstanding of the complex distribution of agency in art.<sup>28</sup> Certainly, machine learning can successfully recognize visual patterns and elements of style and replicate them in novel combinations, with GANs (generative adversarial networks) being particularly effective at this. While the techniques may be novel, the use of generative systems, including computers, in art is not. Since the 1950s this has gone under many names: computer art, cybernetic art, art and technology, art and science, systems art, algorithmic art, generative art, artificial life, electronic art, software art, digital art, media art, new media art; while some are distinct fields of endeavor there is a great deal of overlap with the terminology always in play.<sup>29</sup>

AI art closely follows this tradition, fitting quite naturally within generative art, a field of practice that incorporates autonomous systems into the production of the art. This is not limited to computers and includes mechanical and conceptual systems that cede some control of the final outcome from the artist's hand to quasi-autonomous agents, allowing for emergence within the conceptual structure of the work. Important and valid arguments can be made about machine learning tools forming the latest generation of creative tools, and of the rich connections between developments in science and technology and artistic creativity. Nonetheless, this does not equate to machine creativity. Claims for machine creativity need to be situated within an art historical and theoretical framework, rather than consisting of unsupported assertions based on superficial visual resemblance or the fact that a quasi-autonomous system fabricated part of the work.<sup>30</sup>

There is more at stake, though, than overstating impressive visual results. Claims for creative AI are, I suggest, a form of artwashing.<sup>31</sup> It is an instrumentalization of art and artists that serves to bolster and sustain discussion on irrelevant claims for general AI, drawing attention from problems of actually existing machine learning systems. These claims act to form a distraction screen that refocuses discussion on the possibilities and ethics of a putative general AI, even when such ideas are not even remotely relevant to the pragmatics of current machine learning. This is not to say that

artists should not engage with the aesthetic and critical possibilities of AI and machine learning in all its forms, far from it. Art has an important role to play in expanding the possibilities of AI as an artistic medium, of calling attention to the inequalities and discrimination of unregulated algorithmic systems, in making visible and challenging the unrelenting expansion of surveillance capitalism, not to mention an exploratory role through the creative misuse of these powerful tools. However, as with all art practice working with emerging technologies, a critical awareness of the limits and broader societal implications of the technology need to be maintained.

## **#RiseandGrind Exhibition**

There were two main impetuses to this work. First, the hashtags themselves: my artistic practice has previously explored informal internet communities, particularly those that form around practices and concepts to construct an autonomous conceptual world view. Previous works such as *JoyceWalks*, *Spook...*, and *24h Social* addressed these ideas of hybrid internet-based conceptual worlds from a number of perspectives.<sup>32</sup> *#RiseandGrind* follows in this path through an engagement with a cohesive worldview assembled under the two selected hashtags; one that speaks to the power of social media as a medium for assembling people and ideas in powerful and entertaining ways, and the power of the platform to algorithmically shape its content without this influence being readily apparent. Secondly, the work involves a process of critical engagement with machine learning, to make visible not only how bias can be reinforced through machine learning but also the operation of the black-boxed machine learning process itself.

The work was installed in a number of different iterations from its original commission for *Hustle* at the Science Gallery Lab in Detroit. It was further developed, with new components added, for exhibition in *TULCA* Exhibition of Visual Art in Galway, and in *Screentime* in the Green on Red Gallery Dublin. The work's exhibition is centered on a neon piece, *#RiseandGrind*, connected to a live Twitter feed which illuminates and dims based on activity on the hashtag. Screens display the training process throughout the duration of the exhibition as scrolling text displaying the input training text, the network's generated texts, and their probability weighting epoch by epoch, from early stages to fully trained. The Twitter Bot, which was only active during the run of the first exhibition, is displayed on a series of 7" screens powered by network-connected Raspberry Pi board computers alongside a live display of the hashtags from Twitter (see Figure 4).

The final element is a screen-based display of all the generated tweets, typed character by character. This display includes all the tweets from all stages of training showing the process of the neural net "learning." The exact configurations of these elements are decided individually for each exhibition based on the specificities of the installation space. The exhibitions are accompanied by mediation and discursive events, including the Science Gallery Lab Detroit's extensive program of docent tours, artist talks for *TULCA* in Galway, and a panel discussion with academics and curators at Green on Red gallery, with all events open to the public.





***Figure 4. #RiseandGrind, 2018, Conor McGarrigle. Installation view, Science Gallery Lab Detroit.***

## CONCLUSION

*#RiseandGrind* is a research-based artwork that, through a process of active engagement with the machine learning tools of what is known as artificial intelligence, sought to make visible the complex relationship between the origins and context of training data and the results that are produced through the training process. As a process work the final exhibition outcomes, while important, form only one element of the work. Of equal importance is the sustained process of engagement with these Twitter hashtags, their rendering as data and processing to model the complex activity so that it could be emulated. This process is at one level an attempt to unpack, understand, and generate new knowledge about machine learning and artificial intelligence, conceptually and in practice. In this respect the active engagement with the process—with its errors, missteps, and lack of domain expertise—are an essential component of the work. Display of all stages of the process of training and all the texts generated in the exhibition was central to this, as it allows viewers to form their own conclusions from the visualized process.

The work originates in an interest in ad hoc internet communities that assemble around diverse interests conceptually linked through overarching values—in this case belief in the neoliberal economic values underpinning the digital economy—and how these dynamically form within Twitter's hashtag bubbles. Arguably this ability to assemble and empower communities and audiences, real and imagined, across geographical divides is what has made the internet central to

everyday life. The work demonstrates the tension between the internet as social and surveillance space, evident in the relative ease with which hashtags can be captured and used to model behavior. Although there is an expectation that tweets are in the public domain, the uses to which they are put are not widely known. The artwork visualizes one such process, the training of machine learning models and their deployment to generate text. In this direct manner the project seeks to make visible the opaque workings of machine learning and to highlight issues of bias and the role of the origin and context of training data in creating and sustained algorithmic discrimination. However, the process of the work itself gave rise to many individual, unique moments, where algorithmic and machine processes produced in-between states that speak to the nature of the technical process, making visible the workings of normally opaque algorithms in a way that I suggest is unique to art. Thus, the work adds an additional perspective to the critical debate on AI and society.

AI art is a growing area with a range of practices and approaches that demonstrate that machine learning techniques can produce aesthetically and conceptually engaging art and artifacts. Recent survey exhibitions such as *Uncanny Valley: Being Human in the Age of AI* at the de Young Museum San Francisco, *AI More than Human* at the Barbican London, and *Future and the Arts: AI, Robotics, Cities, Life* at the Mori Museum Tokyo have given overviews of the concerns of the field, which are, of course, broader than any survey show can accurately represent. AI art is a field that goes beyond prominent self-identified AI artists to include artists whose practices engage with machine learning tools in myriad ways as these tools are incorporated into ever more applications. How artists engage with machine learning and AI in their work follows in a lineage of art and technology, encompassing multiple modes of connection and interaction, with artists working many approaches at different scales, from straightforward tool relationships to complex embodied critical engagements, often with multiple modes of engagement in single works and sustained practices. As AI's hype-cycle accelerates, and actually existing machine learning dominance expands, AI art has the potential to make important contributions to this debate.

## ACKNOWLEDGMENTS

*#RiseandGrind* was commissioned by the Science Gallery Lab Detroit with support from Michigan State University. Further development of the project was made possible by a residency with UCD Parity Studios and The Insight Centre for Data Analytics at University College Dublin.

## ENDNOTES

---

1. An observation based on following the hashtag for an extended period of time. Space doesn't permit inclusion of examples, but I do recommend the reader spend some time with the hashtags on Twitter.
2. The growing body of scholarship on the topic includes for example, Safiya Noble's work on racist search algorithms, Ignazio and Klein's intersectional analysis of algorithmic regimes, Ruha Benjamin on how racist bias is built into software systems, Virginia Eubanks and Cathy O'Neil on algorithmic discrimination, Shoshana Zuboff's comprehensive work on Surveillance Capitalism, Antoinette Rouvroy and Thomas Berns on Algorithmic

Governmentality, Couldry and Mejias on data colonialism, and Kate Crawford on AI's hidden environmental and social costs.

3. See Emily Sheng, "The Woman Worked as a Babysitter" for an account of bias in Language Generation.
4. See Axel Bruns and Jean Burgess, "The Use of Twitter Hashtags in the Formation of Ad Hoc Publics."
5. Murthy, Dhiraj. *Twitter*. Cambridge and Medford MA: Polity, 2018.
6. Zappavingna, Michele. "Ambient Affiliation: A Linguistic Perspective on Twitter." *New Media & Society* 13, no. 5 (2011): 788–806. <https://doi.org/10.1177/1461444810385097>
7. The complex role of hashtags on Black Twitter is comprehensively analyzed in the following scholarship: Brock, André. "From the Blackhand Side: Twitter as a Cultural Conversation." *Journal of Broadcasting & Electronic Media* 56, no. 4 (2012): 529–49; Freelon, Deen, Charlton D. McIlwain, and Meredith Clark. "Beyond the Hashtags: #Ferguson, #Blacklivesmatter, and the Online Struggle for Offline Justice." Center for Media & Social Impact, American University, 2016; Graham, Roderick, and Shawn Smith. "The Content of Our #Characters: Black Twitter as Counterpublic." *Sociology of Race and Ethnicity* 2, no. 4 (2016): 433–49.
8. Pariser, Eli. *The Filter Bubble: What the Internet Is Hiding from You*. London and New York: Viking Press, 2011.
9. Varol, Onur. "Online Human-Bot Interactions: Detection, Estimation, and Characterization." *Proceedings of the 11th International Conference on Web and Social Media*, ICWSM, 2017.
10. Zuboff, Shoshana. *The Age of Surveillance Capitalism*. New York: Public Affairs, 2019, 65.
11. Ibid, 68.
12. The reports of the AI Now Institute at New York University are instructive in highlighting some of these issues. <https://ainowinstitute.org/reports.html>
13. Tweeted from @AnimaAnandkumar November 6, 2019. <https://twitter.com/AnimaAnandkumar/status/1191983025250295815>
14. Scholz, Trebor. *Digital Labor: The Internet as Playground and Factory*. New York and London: Routledge, 2013.
15. The unreasonable effectiveness of deep learning is discussed in greater detail in the following: Andrej Karpathy, "The Unreasonable Effectiveness of Recurrent Neural Networks"; John Kelleher, *Deep Learning*; and Terrence J. Sejnowski, "The Unreasonable Effectiveness of Deep Learning in Artificial Intelligence."
16. See Dave Lee, "Why Big Tech Pays Poor Kenyans to Teach Self-Driving Cars," and Kate Crawford, *Atlas of AI*, 63–69.
17. See Noel Fitzpatrick and John Kelleher, "On the Exactitude of Big Data: La Bêtise and Artificial Intelligence."
18. Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press, 2018.
19. Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. "Machine Bias." *Propublica*, May 23, 2016. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
20. Zhaung, Pinghui. "Chinese Woman Offered Refund after Facial Recognition Allows Colleague to Unlock iPhone X." *South China Morning Post*, International Edition, December

- 14, 2017. <https://www.scmp.com/news/china/society/article/2124313/chinese-woman-offered-refund-after-facial-recognition-allows>
21. D'Ignazio, Catherine, and Lauren Klein. *Data Feminism*. Cambridge, MA: MIT Press, 2020.
22. See Antoinette Rouvroy and Thomas Berns, "Gouvernementalité algorithmique et perspectives d'émancipation."
23. I discuss the training of the neural network in more detail in Conor McGarrigle, "#Riseandgrind: Lessons From a Biased AI," *Radical Immersions, Proceedings of Digital Research in the Humanities and Arts*, 2019.
24. Proudfoot, Diane. "Anthropomorphism and AI: Turing's Much Misunderstood Imitation Game," *Artificial Intelligence* 175, no. 5–6 (2011): 950–57. <https://doi.org/10.1016/j.artint.2011.01.006>
25. Nilsson, Nils. *The Quest for Artificial Intelligence*. Cambridge: Cambridge University Press, 2010.
26. Du Sautoy, Marcus. *The Creativity Code*. Cambridge, MA: Harvard University Press, 2019.
27. Bostrom, Nick. *Superintelligence*. Oxford: Oxford University Press, 2014.
28. While an extended discussion of current AI art is outside of the scope of this article, such claims run against the grain of current AI art, which follows in the tradition of generative art and ideas of co-creating with enhanced tools. See Joanna Zylińska's *AI Art: Machine Visions and Warped Dreams* for an expanded discussion of AI art.
29. See Boden and Edmonds, "What Is Generative Art?" and Broeckmann's *History of Machine Art in the Twentieth Century* for a detailed account of this tradition.
30. See for example Veale and Cardosa's discussion of the current state of the art in computational creativity (2019, 1–20) for a realistic picture of the state of the art.
31. Evans, Mel. *Artwash: Big Oil and the Arts*. London: Pluto Press, 2015.
32. For a more detailed discussion of these works, see Conor McGarrigle, "Joyce Walks: Remapping Culture as Tactical Space," *Proceedings of 15th International Symposium on Electronic Art Belfast*, 440–47. Belfast: ISEA International, 2009; and Conor McGarrigle, "Preserving Born Digital Art: Lessons from Artists' Practice," *New Review of Networking* 20, no. 1–2 (2015): 170–78.

## REFERENCES

- Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. "Machine Bias." *Propublica*, May 23, 2016. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Boden, Margaret A., and Ernest A. Edmonds. "What Is Generative Art?" *Digital Creativity* 20, no. 1–2 (June 1, 2009): 21–46. <https://doi.org/10.1080/14626260902867915>
- Bostrom, Nick. *Superintelligence*. Oxford: Oxford University Press, 2014.
- Brock, André. "From the Blackhand Side: Twitter as a Cultural Conversation." *Journal of Broadcasting & Electronic Media* 56, no. 4 (2012): 529–49.

- Broeckmann, Andreas. *Machine Art in the Twentieth Century*. Cambridge, MA: MIT Press, 2016.
- Bruns, Axel, and Jean Burgess. "The Use of Twitter Hashtags in the Formation of Ad Hoc Publics." In *Proceedings of the 6th European Consortium for Political Research (ECPR) General Conference 2011*, 1–9. UK, 2011.
- Couldry, Nick, and Ulises A. Mejias. *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*. Stanford, CA: Stanford University Press, 2019.
- Crawford, Kate. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven, CT: Yale University Press, 2021.
- D'Ignazio, Catherine, and Lauren Klein. *Data Feminism*. Cambridge, MA: MIT Press, 2020.
- Du Sautoy, Marcus. *The Creativity Code*. Cambridge, MA: Harvard University Press, 2019.
- Evans, Mel. *Artwash: Big Oil and the Arts*. London: Pluto Press, 2015.
- Fitzpatrick, Noel, and John Kelleher. "On the Exactitude of Big Data: La Bêtise and Artificial Intelligence." *La Deleuzania*, no. 7 (2018): 142–55. <https://doi.org/doi.org/10.21427/dfw8-m918>
- Freelon, Deen, Charlton D. McIlwain, and Meredith Clark. "Beyond the Hashtags: #Ferguson, #Blacklivesmatter, and the Online Struggle for Offline Justice." Center for Media & Social Impact American University, 2016. <https://ssrn.com/abstract=2747066>
- Graham, Roderick, and Shawn Smith. "The Content of Our #Characters: Black Twitter as Counterpublic." *Sociology of Race and Ethnicity* 2, no. 4 (2016): 433–49. <https://doi.org/10.1177%2F2332649216639067>
- Karpathy, Andrej. "The Unreasonable Effectiveness of Recurrent Neural Networks." *Andrej Karpathy Blog* (blog), May 21, 2015. <https://karpathy.github.io/2015/05/21/rnn-effectiveness/>
- Kelleher, John. *Deep Learning*. Cambridge, MA: MIT Press, 2019.
- Lee, Dave. "Why Big Tech Pays Poor Kenyans to Teach Self-Driving Cars." BBC News, November 3, 2019. <https://www.bbc.com/news/technology-46055595>
- McGarrigle, Conor. "Art in the Data-City: Critical Data Art in the Age of Surveillance Capitalism." In *Digital Art in Ireland*, ed. James O'Sullivan, 46–58. London and New York: Anthem Press, 2021.
- . "Preserving Born Digital Art: Lessons from Artists' Practice." *New Review of Networking* 20, no. 1–2 (2015): 170–78.
- Murthy, Dhiraj. *Twitter*. Cambridge, UK and Medford, MA: Polity, 2018.

- Nilsson, Nils. *The Quest for Artificial Intelligence*. Cambridge: Cambridge University Press, 2010.
- Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press, 2018.
- O'Neil, Cathy. *Weapons of Math Destruction*. New York: Crown, 2016.
- Pariser, Eli. *The Filter Bubble: What the Internet Is Hiding from You*. London and New York: Viking Press, 2011.
- Proudfoot, Diane. "Anthropomorphism and AI: Turing's Much Misunderstood Imitation Game." *Artificial Intelligence* 175, no. 5–6 (2011): 950–57. <https://doi.org/10.1016/j.artint.2011.01.006>
- Rouvroy, Antoinette, and Thomas Berns. "Gouvernementalité algorithmique et perspectives d'émancipation." *Réseaux* 1, no. 177 (2013): 163–96. <https://doi.org/DOI.10.3917/res.177.0163>
- Scholz, Trebor. *Digital Labor: The Internet as Playground and Factory*. New York and London: Routledge, 2013.
- Sejnowski, Terrence J. "The Unreasonable Effectiveness of Deep Learning in Artificial Intelligence." *Proceedings of the National Academy of Sciences* 117 (48), December 1, 2020. <https://doi.org/10.1073/pnas.1907373117>
- Sheng, Emily. "The Woman Worked as a Babysitter: On Biases in Language Generation." Hong Kong, 2019.
- Varol, Onur. "Online Human-Bot Interactions: Detection, Estimation, and Characterization." Paper presented at the International AAAI Conference on Web and Social Media, Montreal, 2017.
- Veale, Tony, and F. Amílcar Cardoso. "Systematizing Creativity: A Computational View." In *Computational Creativity: The Philosophy and Engineering of Autonomously Creative Systems*, 1–20. Cham, Switzerland: Springer, 2019.
- Zhaung, Pinghui. "Chinese Woman Offered Refund after Facial Recognition Allows Colleague to Unlock iPhone X." *South China Morning Post*, International Edition, December 14, 2017. <https://www.scmp.com/news/china/society/article/2124313/chinese-woman-offered-refund-after-facial-recognition-allows>
- Zuboff, Shoshana. *The Age of Surveillance Capitalism*. New York: Public Affairs, 2019.
- Zylinska, Joanna. *AI Art: Machine Visions and Warped Dreams*. London: Open Humanities Press, 2020.



## AUTHOR BIO

Conor McGarrigle is an artist and researcher working primarily with digital media. His practice is characterized by urban interventions mediated through digital technologies and data-driven explorations of networked social practices. He is a lecturer in Fine Art New Media at the TU Dublin School of Creative Arts and a Senior Research Fellow at the European Culture and Technology Lab. He has exhibited extensively internationally including the Venice Biennale, Fundació Miro Mallorca, the Saint-Étienne Biennale, Redline Gallery Denver, SIGGRAPH, FILE São Paulo, Art on the Net Tokyo, Seoul New Media, SITE Santa Fe, as well as EVA International, *TULCA*, Green on Red, and the Science Gallery.