Original Research Article

BIG DATA & SOCIETY

Big Data & Society
January–June: 1–12

DOI: 10.1177/20539517221145372
journals.sagepub.com/home/bds
SAGE

Machine learning and the politics of synthetic data

Benjamin N Jacobsen

Abstract

Machine-learning algorithms have become deeply embedded in contemporary society. As such, ample attention has been paid to the contents, biases, and underlying assumptions of the training datasets that many algorithmic models are trained on. Yet, what happens when algorithms are trained on data that are not real, but instead data that are 'synthetic', not referring to real persons, objects, or events? Increasingly, synthetic data are being incorporated into the training of machine-learning algorithms for use in various societal domains. There is currently little understanding, however, of the role played by and the ethicopolitical implications of synthetic training data for machine-learning algorithms. In this article, I explore the politics of synthetic data through two central aspects: first, synthetic data promise to emerge as a rich source of exposure to variability for the algorithm. Second, the paper explores how synthetic data promise to place algorithms beyond the realm of risk. I propose that an analysis of these two areas will help us better understand the ways in which machine-learning algorithms are envisioned in the light of synthetic data, but also how synthetic training data actively reconfigure the conditions of possibility for machine learning in contemporary society.

Keywords

Machine learning, data, algorithms, risk, ethics, variability

Introduction

In 2020, computer scientists at the Mixed Reality & AI Labs at Microsoft Cambridge developed an algorithmic model that is capable of generating photorealistic synthetic images of human faces. The aim of the model is to provide a means of developing synthetic training data for facial recognition algorithms. The model, which is called ConfigNet, is a composite of computer graphics techniques and generative adversarial networks (GANs), the algorithm often evoked in the context of so-called 'deepfakes' (de Vries, 2020). The key idea behind ConfigNet, the researchers claim, is to enable both the generation of photorealistic synthetic faces at scale as well as enhanced control over the generative process: 'this allows for independent control of various face aspects including: head pose, hair style, facial hair style, expression and illumination' (Kowalski et al., 2020: 301). 'By simultaneously training the model on unlabelled face images', the computer scientists explain, 'it learns to generate photorealistic looking faces, while enabling full control over these outputs' (p. 301). As a result, with ConfigNet the synthetic face is rendered into an array of controllable and tweakable parameters within the algorithmic model: head poses can be shifted, hair styles can be changed, skin colours can be altered. A question raised by Microsoft Cambridge's ConfigNet model is therefore: What if facial recognition algorithms could be trained using so-called 'synthetic faces', faces that do not relate to any specific persons? Synthetic data in this context promise to provide machine-learning engineers both with control over the manufacturing process as well as with access to a vast heterogeneity of faces that can be used to train facial recognition systems – none of which refer to any actual real persons. In short, whatever face is needed can be algorithmically generated. The parameters can be tweaked to give machine-learning models any types of faces they need.

The importance of data for the reconfiguration of society has been well documented in the critical scholarship (Beer, 2016; Boyd and Crawford, 2012; Halpern et al., 2022; Kitchin, 2014) as has the variegated ethicopolitics of machine-learning algorithms and facial recognition (Amoore, 2020; Andrejevic and Selwyn, 2019; Bucher, 2018; Buolamwini

Department of Geography, Durham University, Durham, UK

Corresponding author:
Benjamin N Jacobsen, Department of Geography, Durham University, Lower Mountjoy, South Rd, Durham, DH1 3LE, UK.
Email: benjamin.jacobsen@durham.ac.uk

and Gebru, 2018; Dourish, 2016). Moreover, the training datasets of machine-learning algorithms have also constituted a crucial point of critical intervention in recent years (Denton et al., 2021; Gebru et al., 2020). Studies have shown, for instance, how issues of racial and gender bias are endemic in training datasets such as ImageNet (Crawford and Paglen, 2019) and those used for policing (Angwin et al., 2016). As a result, there have been calls to make algorithms and training datasets more transparent and representative (Hutchinson et al., 2021). Yet, what happens to the common critique that algorithms are not trained on an adequate representation of the world if, in fact, they can now be trained on data that do not derive from any existing real world? What happens when algorithms are trained on data that are not real but rather 'synthetic', not actually referring to real persons, objects, or events? There is currently little understanding of the role played by and the ethicopolitical implications of synthetic training data for machine-learning algorithms.[1] Synthetic data are currently being used to train algorithmic systems in a plethora of domains, contexts, and tasks ranging across social media, facial recognition, immigration, autonomous driving, weapons detection, terrorist threats prediction, military drone surveillance, crowd behavioural analysis, privacy-preserving patient data, fraud detection, insurance anomaly detection, medical imaging, segmentation, and classification.[2] Synthetic data for machine learning were considered one of the 'Top 10 Breakthrough Technologies' of 2022 in *MIT Technology Review* (Heaven, 2021), and in July 2021 Gartner, a research and advisory company, wrote a piece provocatively titled 'Forget About Your Real Data – Synthetic Data is the Future of AI' (White, 2021). It is therefore crucial to examine the ethicopolitical implications of machine-learning algorithms being trained on synthetic data.

To begin, it is worth reflecting on the question *what exactly are synthetic data?* In the broadest terms, they are data points that have been generated by a machine-learning algorithm; data which in turn are used to train other machine-learning algorithms (Courville et al., 2016). They have not been extracted from a real-world context, such as biometric data, social media data, or annotated facial images. They have not been mined nor collected and as such they are not part of the 'behavioural surplus' (Zuboff, 2019) that lies dormant until the 'AI extractive industry' (Crawford, 2021) exploits its financial and computational possibilities. Instead, they are algorithmically generated data points that approximate those found in the training dataset. In short, they are data generated *by* algorithms and *for* algorithms. Synthetic data are often used to either reproduce the salient attributes of a given data example (e.g. synthetic computed tomography (CT) scans capturing the likeness of real liver lesions) or to reproduce the overall statistical distribution of some real-world dataset (e.g. synthetic insurance or medical patient datasets that mimic the patterns of real persons without actually referring to any real persons) (Nikolenko, 2019). Moreover, they are commonly used in sectors such as medical imaging where there is commonly a paucity of training data for AI models and a related ethical question in the use of sensitive personal data (Chen et al., 2021). In other words, synthetic data are broadly speaking used to mimic real-world data, to look like it, to stand in for it, and to be used *as though it were* real training data for machine-learning algorithms.

One particularly widespread mode of generating synthetic data is by using GANs. Proposed by Ian Goodfellow and colleagues in their breakthrough paper of 2014, the core idea of GANs is that two neural network algorithms are trained simultaneously and pitted against each other through a complex adversarial and iterative process. One neural network, often called the generator, seeks to capture and reproduce a given data distribution whilst the other, often called the discriminator, estimates the probability that a data sample has derived either from the generator or from the training dataset (Goodfellow et al., 2014; Zeilinger, 2021). The discriminator is trained on real data samples (such as images of cars, faces, or cats) in order to be able to detect the fake samples generated by the generator. The generator network, on the other hand, is initially trained on vectors of noise to produce a data distribution that resembles the real data depicting a car or a face, which is then fed to the discriminator as input. The discriminator will then emit a score that indicates the likelihood that the sample is either fake or real, which in turn is backpropagated and the parameters of the generator are tweaked accordingly by the model.[3] The generation of synthetic images via GANs is therefore an iterative process whereby the generative network gradually learns to produce increasingly realistic synthetic images of cars, faces, or cats. In short, over time, the real and the fake will become increasingly indistinguishable to the discriminator network (Goodfellow et al., 2014). The aim with generative models such as GANs is to produce synthetic outputs that are as proximate as possible to the training data without being an identical mapping of them.

As the impact of synthetic data used for machine learning remains virtually unresearched in the social sciences, this article explores the politics of synthetic data through two interlinking ideas: first, I argue that synthetic data promise to emerge as a rich source of exposure to variability for the algorithm. It is this exposure to variability that is said to underpin a process of machine learning where the algorithm anticipates future instances based on the multiple past features it has been exposed to. As Amoore (2020: 72) put it in a discussion of the AlexNet algorithm, 'if one exposes it to sufficient data in the variability of what a leopard could be, then it will learn to anticipate all future instances of leopards'. Here, I point out that synthetic data contribute to what Orit Halpern calls the resilience logic of machine learning. Synthetic variability, in this view, constitutes a way to make algorithms more resilient to volatility and change in the world. Second, I argue that one of the central promises of synthetic data is that they can place algorithms beyond

the realm of risk. Conceptualising synthetic data as a technology of risk, I point out that they participate in the redefinition of risk through the generation of synthetic variability: rare events, atypical faces, and unusual objects. In short, they constitute a promise to de-risk machine learning. However, I argue that this promise is problematic because it obfuscates the ways in which risk is always already the structural condition of all machine-learning algorithms. That is, they participate in the partial and experimental engendering of new boundaries of normal and abnormal, good and bad, which inevitably has a differential impact on human lives. The structure of the paper also echoes this dual argument that I put forward: the first section broadly discusses the relationship between synthetic data and variability whilst the latter section examines synthetic data in relation to ideas of risk. Overall, an analysis of these two key aspects will help us better understand the ways in which algorithms are envisioned in the light of synthetic data, but also how synthetic training data can be seen to reconfigure the condition of possibility for machine learning in contemporary society.

As such, through a critical analysis of synthetic data I aim to 'incite an opening' (Berlant, 2007), an additional entry point into the study of the power and politics of machine-learning algorithms; one which attends to the synthetic data on which they are increasingly trained as well as the possibilities, discourses, and logics this engenders. Yet, I also wish to emphasise that synthetic data by no means constitute an innocent transformation of the possibilities of machine learning. As I discuss in the conclusion, they are fundamentally ethicopolitical insofar as they have the potential to disrupt and even displace our current grounds for critique of AI. A study of synthetic data should therefore encourage a rethinking and reimagining of the relationship between machine learning, power, and ethics more broadly.

Resilience and the importance of exposure

In order to better understand the link between synthetic data and the idea of variability, it is crucial to understand the role of synthetic data for machine learning in relation to broader conceptualisations of shock and resilience. Orit Halpern (2021, 2022) has continually emphasised a correlation between historical developments in psychology, economics, and computer science in the emergence of a widespread logic of resilience. Halpern examines how ideas of resilience became central to discussions surrounding the optimal response to extreme volatility and change. Broadly speaking, resilience signifies a capacity to adapt to 'intense external perturbations and thus to persist over longer time periods' (Halpern et al., 2017: 122). Shock and resilience therefore became a paradigm which supposedly worked on both micro and macro scales, for human brains and societies. Volatility and shock became something one could and *should* adapt to on an individual (and neurological) level, but the constant absorption of shock also emerged as a way of managing and organising society. Resilience emerged as a way of life.

How does the notion of resilience relate to computational systems and machine-learning algorithms? Halpern draws a parallel between resilience as conceptualised in psychiatry and economics with the inception of neural network algorithms. Frank Rosenblatt's 'perceptron', the first artificial neural net developed in the late 1950s, was modelled on the basic architecture of the human brain and, as such, faced similar issues as human brains: how to respond to extreme change or 'traumatic' events? In short, how to optimise the management of volatility. This logic and language of resilience in relation to algorithmic systems became particularly apparent at the start of the COVID-19 pandemic. During the height of the pandemic, normal living was put on hold and instead there was a rush to bulk buy items such as toilet rolls, face masks, and hand sanitisers. In fact, the pandemic constituted a stark disruption and sudden shift in our normal patterns of consumption. This disruption had a destabilising impact on Amazon's recommender systems. The disruption caused, as one reporter noted, 'hiccups for the algorithms that run behind the scenes in inventory management, fraud detection, marketing, and more', which meant that 'machine-learning models trained on normal human behavior are now finding that normal has changed, and some are no longer working as they should' (Heaven, 2020).

The disruption of Amazon's recommendation models echoes a question posed earlier: How can algorithmic systems be made adaptable and resilient under such volatile circumstances? How can they be made adaptable to new or shifting curves of normality? 'Healthy brains and neural networks', Bruder and Halpern (2021: 13) argue, would both have to be 'equipped with a range of mechanisms that govern the effects of volatility on the respective system to allow for continuous change and adaptation while avoiding catastrophic breakdowns'. Focusing on the historical use of pruning or compression techniques in AI research, Bruder and Halpern suggest that these methods serve the overall aim of maintaining 'the network's internal health' (p. 19). Pruning and compression, in their view, exemplified how 'shock and trauma can serve a stability function in our present, in fact, even a mechanism to preserve meaning and value in markets as well as in language processing' (p. 17). Similarly, the COVID-19 pandemic was evoked by data scientists and computer engineers as an apt pretext to make contemporary machine-learning models more resilient to volatility and rare events. But rather than evoking pruning and compression techniques, there was a call for infusing more *variability* into the learning process of machine-learning algorithms. Rajeev Sharma, global vice president of the technology services company

Pactera Edge, stated in an interview with *MIT Technology Review* that 'A pandemic like this is a perfect trigger to build better machine-learning models', adding that algorithms in the future should also be trained on other 'freak events' such as the Great Depression of the 1930s, the Black Monday stock market crash in 1987, the Dot-Com Bubble in the late 1990s, as well as the 2008 global financial crisis (Sharma cited in Heaven, 2020). In other words, rare events such as pandemics and market crashes serve not only as additional data examples to be included in the training datasets of algorithmic models. Rather, they are a mechanism by which machine-learning algorithms can be made more resilient, robust, and adaptable to change and volatility. It can enable them to bounce back.

Resilience, in this context, is therefore not a question of being able to account for every possible future event; it is neither a question of planning nor establishing political orders with fully controllable parameters. Resilience is not primarily a question of anticipation. Instead, resilience emerges as a logic of openness, elasticity, adaptability, and change. Whilst the exposure of algorithms to extreme variability in the world highlights the fragility of these systems, it is precisely this exposure that figures as the means by which they can be made more resilient. Echoing Bruder and Halpern (2021), the exposure to variability is assumed to have a 'stability function' for machine-learning models. It is a particular mechanism that transforms the effects of volatility into a source of stability for algorithms. In other words, variability is conceptualised as a means to achieve more resilience and, thus, is reconfigured in highly desirable terms for machine-learning algorithms. The question therefore is not whether this conceptualisation of the world as variability, as a varied array of probability distributions, is accurate but rather that it makes possible and desirable a logic of resilience in machine learning. Yet, what happens when this variability is not readily available? How can algorithmic models be made resilient when there is insufficient variability in their training datasets? In what follows I want to argue that tech companies seek to bypass this issue by *generating* synthetic data as a source of variability to algorithmic models. As the logic goes: If it cannot be found, it can be generated; if it *could* be in a training dataset it *should* be. Thus, the emergence of synthetic data embodies a broader logic whereby machine-learning systems increasingly seek to incorporate instability, perturbations, and shock through a wide variety of generated data examples in order to become more robust and resilient.

'We don't want to wait to see the "odd"': Creating synthetic perturbations

As I have discussed, exposure to variability – to the world as varied, chaotic, and characterised by rare events – is becoming an increasingly crucial aspect of training machine-learning algorithms. Exposure is understood here as the mechanism by which algorithms are made both increasingly capable of generalising to previously unseen data distributions as well as more resilient to shock events.[4] In other words, synthetic data as exposure to variability can even be understood as 'an embrace of transformation and shock' (Halpern, 2021: 246), an embrace of the generative capacities of volatility and the rare. The algorithm's capacity to approximate, generalise, and infer in the world is very much dependent on its exposure to the world as a complex assemblage of varieties and rarities, that is, a wide distribution or spectrum of different data examples in the world. Much of the promise underlying synthetic data is encapsulated in its supposed capacity to provide a near limitless source of exposure to variability. In a comment piece recently published in *Nature*, which discusses the use of machine learning in medical imaging and healthcare, synthetic data are hailed as the means to overcome the paucity of annotated training datasets that is often a hindrance to training optimal machine-learning models in the field of medicine. Synthetic data, the authors note, are said to 'increase diversity in datasets' as well as to 'increase the robustness and adaptability of AI models' (Chen et al., 2021: 493).

One of the underlying assumptions of the generation and use of synthetic data is that variability can be algorithmically or synthetically generated. Much of the value and promise ascribed to synthetic data is in its varied or diverse representation of a wide spectrum of artificially generated data examples: edge cases, corner cases, abnormalities, anomalies, and rarities. Known for their capacity to generate photorealistic images, GANs are increasingly used to generate the difference and diversity that developers want an algorithm to see. In the field of autonomous driving (AV), for instance, self-driving vehicles depend on vast volumes of data that represent driving instances in a plethora of contexts. Although companies such as Tesla are known for their huge volumes of driving data used to train their deep learning models, the lack of variability, of very rare cases, still figures as a challenge. This has opened up a space for synthetic data, that is, for edge cases to be synthetically generated and incorporated into the training of AV models. This was well illustrated in a presentation titled 'Hard Miles without Hard Miles' delivered by Paul Newman, founder of the AV tech start-up called Oxbotica, at the 2021 NVIDIA GTC Conference. As Newman (2021) put it, 'We don't want to wait to see "odd"' and 'we really don't have to wait for hard miles'. The generative capacities of algorithms such as GANs means that the well-known imperative to capture the world as data is shifting. As Newman (2021) stated, 'let's not do hard miles to find the hard miles'. The implication is, of course, that the 'odd', however it is conceptualised and operationalised, can always be algorithmically

generated. It does not have to have collected from a real-world setting, a process which may be both heavily time consuming and expensive. The suggestion here is also that instead of waiting for the odd to take place, it can be made readily and speedily available through generative processes.

Synthetic data thus embody a promise to generate and augment the tails in an overall distribution for a machine-learning algorithm, even if these tails do not exist or are not readily available in an existing dataset. Synthetic data are assumed to provide richer and more fine-grained representations of the tails in a data distribution. Yet, synthetic data figure not only as an artificial source of exposure to variability, but are also seen in more explicitly adversarial terms. In a computer science research paper titled 'ChaffeurNet: Learning to Drive by Imitating the Best and Synthesizing the Worst', Bansal et al. (2019) note the challenges of developing robust deep learning algorithms for self-driving vehicles.[5] Training a recurrent neural network algorithm on 30 million data examples – which amounts to over 60 days of continuous driving – they state, was not enough. One reason for this was that 'this model would get stuck or collide with another vehicle parked on the side of a narrow street, when a nudge-and-pass behavior was viable' (p. 23). In order to make the algorithm better able to generalise to a wider, unseen distribution of data examples in the real world, Bansal and colleagues proposed the use of what they call 'synthetic perturbation data'. That is, exposing the network to 'a much more diverse set of drifts (translation, rotation, heading, speed) than what is possible with just the steering perturbations' (p. 23). The aim of synthesising perturbations, they state, is to expose the learner 'to data that is immediately outside the demonstration' (p. 23).

Synthetic data figure here as *perturbations* – that is, deviations, disruptions, interruptions, ruptures – to the machine learning model. It represents, as the title of the paper suggests, the worst-case scenarios that a driver, human or non-human, may encounter whilst driving. What do these perturbations consist of? As the authors point out, they include 'synthetic collisions with other objects and the environment' (p. 23) as well as instances of 'driving off-road' (p. 26). Such data examples can be difficult or even impossible to collect, as they require that someone actively gets involved in vehicular collisions whilst being able to record it as data. In contrast, synthetic data constitute a means of synthesising or generating these events without having to collect it elsewhere. These synthetic perturbations are considered valuable to the model as they 'help the network to learn to avoid these and generalize better' (p. 23). In other words, synthetic exposure to some of the worst-case scenarios helps the algorithm learn to recognise them and, most importantly, to avoid them. By introducing artificial disturbances or perturbations to the algorithm in the present, the algorithm is thought to be rendered more resilient to possible shock and 'freak events' in the future.

Synthetic data, recognition, and 'the separating power' of machine-learning algorithms

Synthetic data are emblematic of an emerging logic where any form of variability in a data distribution for a given domain can be algorithmically generated – whether that is edge cases such as car collisions for autonomous driving or more black and minority ethnic faces for a facial recognition model. Yet, this connection between machine-learning algorithms, synthetic data, and the notion of recognition is worth expanding upon. For, as the previous example highlights, synthetic data are framed as an easily scalable way to train the algorithm what to *avoid* in the context of driving. Therefore, it was a question of recognition as a form of exclusion. Yet, synthetic data also appear to hold out the promise of expanding the prescribed set of possibilities for recognition in the algorithmic model. That is, they claim to enlarge the model's conventions and limitations of what can be recognised in the world. As was stated in one computer science paper, synthetic data 'empowers the network to achieve more complex behavior' (Tremblay et al., 2018: 1084). As exposure, and as a source of data variability, synthetic data therefore promise to expand an algorithm's field of possible vision. In short, to render the world in its complexity and variability increasingly visible and actionable to the algorithm.

This notion of an extended field of vision is well illustrated in the domain of medical imaging and classification. In a paper titled 'GAN-based Synthetic Medical Image Augmentation for increased CNN Performance in Liver Lesion Classification' (2018), Israeli-based scientists Maayan Frid-Adar and colleagues outline their combined approach to liver lesion classification. They used a GAN to produce additional synthetic data to train a convolutional neural network (CNN) to classify liver lesions in CT images. Whilst the discriminator neural net was trained on real CT images of various liver lesions, the generator networks produced additional realistic synthetic liver lesion images in order to enlarge and make more varied the training dataset of the classifier algorithm. The authors argue that algorithmically generating synthetic data resulted in higher accuracy performance for the model classifying benign and malignant tumours in the CT scans.[6] Crucially, synthetic variability functions here as a means of 'directing and disciplining the attention' of the algorithm, making it possible for the classification model to focus on specific aspects of the CT scans whilst ignoring others (Amoore, 2009). Synthetic data operate here as a function of attention, a mode of amplification, of directing the model's attention towards the liver lesion through exposure to synthetic variability. In other words, through an exposure to an array of synthetic variations of the liver lesion, the real liver lesion is supposedly made visible.

Yet, the particular line of sight generated by synthetic data in this case goes beyond merely directing and disciplining algorithmic attentiveness; instead, synthetic variability figures precisely as a mechanism which enables the division of the world in new ways. It is a type of rationality which participates in the reconfiguration and regeneration of the world. As the researchers state, one ongoing difficulty in medical imaging, especially in tumour detection, is how to disentangle and make sense of overlapping features and attributes: 'Each class has its unique features but there is also considerable intra-variability between classes, mostly the metastases and hemangiomas' (Frid-Adar et al., 2018: 329). In other words, the difficulty remains a matter of how to differentiate between the overlapping characteristics of malignant and benign cancers, which directly impacts diagnostic decisions. However, when trained on synthetic data as a mechanism of exposure to variability, the convolutional neural net supposedly 'exhibited better separating power' (p. 326). That is, it was seemingly able to better differentiate between different kinds of cancerous tumours, benign and malignant. As such, synthetic data increasingly form part of 'the visual economy' of contemporary algorithmic society, whereby they act 'as a means of dividing, separating, and acting upon arrays of possible futures' (Amoore, 2011: 34). In other words, they reconfigure the conditions of possibility of machine learning in the sense that they participate in new ways of breaking up the visual field of images, to disentangle it, making the algorithm attentive to new contrasts and juxtapositions in the pixel values of input data. They promise to generate a new 'prescribed set of possibilities' of algorithmic vision, the 'conventions and limitations' within which algorithms are necessarily embedded (Crary, 1990: 6). Synthetic data can therefore be understood as a technology that both cuts up the world and binds it together in new ways, generating new actionable correlations, diagnoses, and ultimately futures. Rather than merely foreclosing or expanding, synthetic data as a mechanism of exposure to variability help to generate novel ways of algorithmically processing the world. People, events, objects, features, and attributes are rendered recognisable in new ways through synthetically trained algorithmic systems. As such, the question that remains to be asked is: What are the wider ethico-political implications of synthetic data constituting a source of variability to algorithmic models?

'Where there is no real data there are no real risks': Synthetic data as technology of risk

In what follows, I want to propose that this relationship between synthetic data and variability allows for the emergence of a new logic of risk, a new framing of the ethics of machine learning. That is, synthetic data as a mechanism of exposure to variability enable the emergence of a new set of discourses around risk and algorithmic systems. In this logic, the algorithmic generation of synthetic variance allows for the representation that the risks associated with machine-learning models can be cancelled. Through the generation of variability, synthetic data promise to place algorithms beyond the realm of risk. The claim that any variability can be generated by an algorithm – that the rare and atypical can be synthetically represented – is the condition of possibility for the reconfiguration of risk in discourses around machine learning.[7] The emergence of this new logic of risk can be seen articulated, crucially, by many of the tech companies and start-ups that generate and sell synthetic data. My aim here, however, is not to a provide a systematic overview or analysis of these companies, but rather to see their discursive framings of synthetic data as diacritical to this emerging understanding of risk. For instance, the London-based tech company Hazy, founded in 2017, specialise in generating and selling synthetic transactional data for financial services. Outlining the story of the company, they state that:

> We looked at the opportunity in applying machine learning to synthetic data, which starts with the basis that *with no real data there is no real risk*. We pivoted our focus completely to AI-generated smart synthetic data and have worked to make it both differentially private and statistically representative.[8]

Hazy is by no means an isolated example of the notion that if no real data are used then risk is annulled: The tagline of the synthetic data company Syndata AB, located in Sweden, is 'Solving real data problems without real data risks';[9] and the London-based start-up called Synthesized claim that their generative algorithmic models are '10× the performance, 0 risks'.[10] On the one hand, such claims can be seen as explicit 'exercises in promotion' (Beer, 2019: 21), that is, the formulation of striking and glossy marketing speak aimed at maximising the selling of a particular tech product. On a deeper level, however, I argue that they are precisely indicative of a wider logic whereby synthetic data figure as a means of supposedly placing machine-learning algorithms beyond the realm of risk. Where the primary risks of machine-learning models are thought to emanate from their training data, the perfectible control via synthetic data becomes a means of representing those risks as eradicable.

This claim – that synthetic data enable the de-risking of algorithms, generating a risk-free zone in which algorithmic systems can operate – functions as a means of generating a specific understanding of risk. As Francois Ewald (1991: 199) has argued, 'Nothing is a risk in itself; there is no risk in reality. But on the other hand, anything *can* be a risk; it all depends on how one analyzes the danger, considers the event'. Following Ewald's insight, the definitional

parameters of risk are never given nor fixed. The boundaries of what counts as a risk in a particular context is malleable and thus foregrounds the politics of *who* gets to define risk and *how*. As such, I seek to go beyond Ulrich Beck's (1992) notion of risk in his influential book *Risk Society: Towards a New Modernity*. Published in the aftermath of the Chernobyl nuclear disaster, Beck argues that risks are not only human constructions but have attained such a scale and globality that they exceed human control whilst simultaneously existing below the threshold of ordinary human perception. Such risks are typically 'localized in the sphere of physical and chemical formulas (e.g. toxins in foodstuffs or the nuclear threat)' (p. 21). Beck is therefore concerned with the scale and increasing incalculability of contemporary global, technological threats. Yet, risks are not simply human constructions, not simply the chemical or physical consequences of human ingenuity. As Ewald convincingly argues, *anything* can be a risk. What counts as a risk (or a non-risk) depends on the particular tools, techniques, modes of analysis, and as well as conceptual schemas through which risk is defined.

What does this mean in the context of machine learning and synthetic data? Many risks, as well as their definitional parameters, are constituted in and through technologies.[11] Yet, here I explore how risks are being actively considered and defined as such in the context of machine learning and synthetic data. Rather than being the 'answer' to the risks, dangers, and threats of algorithms, I argue that synthetic data should be seen as a specific *technology of risk*. That is, they enable a particular conception of risk to emerge in relation to machine learning. in Ewald's terms, synthetic data become a way of analysing the dangers of machine learning; a particular schema of rationality, 'a way of breaking down, rearranging, ordering certain elements of reality' (Ewald, 1991: 199). Synthetic data generate a specific understanding of algorithmic risk (and non-risk). Here, I want to outline three particular ways in which synthetic data figure as a technology of risk, three specific mechanisms that enable the emergence of this political logic, this understanding of risk. These are: exposure to variability as a mode of de-risking, risk as belongingness, and risk as a 'training dataset problem'.

First, synthetic data as a mechanism of exposure to variability not only figure as a way to make algorithms more resilient to volatility and change, but also as a means of supposedly de-risking algorithmic models. This can be seen, for instance, in the marketing strategy of the Israeli synthetic data company called Datagen. They claim to generate and sell data that are 'highly variable, unbiased, and annotated with perfect consistency and ground truth' and that their synthetic data 'help you fight high-level bias with datasets that reflect the diversity of our world. Our datasets can be customized to include any variance you need, from age, gender, race, and body mass to blemishes, skin textures, and other edge cases'.[12] The emergence of synthetic data, it is claimed, generates a new space of possibility for machine learning, one where any diversity or 'variance you need' can be algorithmically manufactured and bias can be fought. With synthetic data, therefore, there is an algorithmic generation of what Michel Foucault (2007: 63) calls 'differential normalities' or 'different curves of normality'. In other words, the question is not simply to capture the general, the typical, the normal; that is, the overall characteristics of a population within a machine-learning algorithm (which has been a typical promise of Big Data). Instead, the question is to generate and expose the algorithm to synthetic blemishes, abnormal skin textures, and edge cases in order to place the model beyond risk. Here there is an amplification of atypical faces and skin textures (or rather the faces and textures not otherwsie represented in the distribution) through synthetic data in order to generate a specific idea of risk seen here in terms of bias, exclusion, or marginalisation.

The risks associated with algorithmic models in this context can be supposedly reduced (and even eradicated) through the algorithmic generation of new forms of data distributions, new distributions of normality: synthetic variability. In other words, within this framework, the exposure of algorithmic models to synthetic variability are seen as less risky, less biased, as somehow doing less harm in the world. Exposure to variability, in short, figures as a way to 'solve' the risks of machine learning. Yet, the claims made by companies such as Datagen about the utility of synthetic data as exposure to variability have ethicopolitical implications. In claiming to fight bias via the algorithmic generation of synthetic instances of gender, race, and so on, the politics of the technology erases the protected characteristics that could otherwise be mobilised to open up or resist machine-learning models in society. In other words, this is a politics that claims neutrality and objectivity where there is actually the inscription of new modes of power to allocate risk across social and ethnic groups.

Second, the emergent notion of risk associated with synthetic data is intimately interwoven with notions of personhood and belongingness. This is well illustrated through a 2021 interview with Stephane Gentric, the Global Research & Development Manager at the multinational security-services company Idemia. In the interview, Gentric was asked where the company gets the data to train its facial recognition algorithms. He explained that training data for their systems are obtained in three principal ways: first, data are routinely provided by clients willing to share their data with Idemia; secondly, the company often holds 'internal campaigns' where 'our employees allow us to use their biometrics' (Gentric, 2021). Lastly, he states, training data for their deep learning algorithms are also obtained through the use of synthetic data or what he calls 'synthetic images':

> Using a Generative Adversarial Network (GAN), we have been able to create synthetic data using a real database.

> The synthetic data is similar to the real database, but it is in actual fact, completely fictional. We can, for example, generate qualitative synthetic facial images from different angles or fingerprints *that do not belong to anyone*. Not many companies are able to do this well, and this is one of the factors that differentiates us from other biometric companies (Gentric, 2021; my emphasis).

As Gentric suggests, not only faces but a vast array of biometric data such as fingerprints are synthetically generated and used for training. However, the crux of the interview with Gentric is not the emphasis on the generative capacities afforded by GANs or synthetic data, but rather a particular idea of personhood and belongingness that functions to legitimise the ongoing generative process. Central to the appeal of synthetic data are precisely the evocation of artificiality or fictionality, the fact that synthetic faces do not belong to any specific persons, that they are 'completely fictional'. The fictionality of synthetic fingerprints and irises is evoked here as a means to justify the ongoing generation and deployment of synthetic data to train their facial recognition systems, which in turn continues to have a differential impact on actual human bodies. This notion of synthetic data as a form of non-belongingness is therefore a central aspect of the emergent logic of risk associated with synthetic data for machine learning. Synthetic data figure here as a promise to cancel the risks of biometrics, the risks associated with the processes of rendering the (Othered) body identifiable, visible, and governable.[13]

The understanding of risk as related to notions of belongingness (and data as risk-free if it does not belong to a data subject) not only serves particular techno-capitalist interests (such as the continued use of facial recognition software by Idemia), but, perhaps more importantly, it is based on a brittle balancing act. For synthetic data to be used as a viable replacement or addition to real data, they must be made to 'resemble' real data (Nikolenko, 2019). What is meant by resemblance in this context? Here, the idea is a floating signifier insofar as the significance of resemblance differs depending on the domain, task, or problem for which the synthetic data are being generated and used. For instance, in contexts such as facial recognition software, it is crucial that synthetic facial images are highly photorealistic, capturing the various nuances, contours, and complexities of the human face. In other contexts, such as credit card fraud detection, it is important that the synthetic data have the same statistical properties and characteristics as the real data, since computer scientists wish to develop models that learn to detect irregularities and anomalies in real data input streams. Yet, for synthetic data such as faces, fingerprints, irises, or skin textures to be useful they must differ sufficiently from the real-world dataset in order to be synthetic, but not differ *so much* that they prevent algorithms from being effectively used in the real world. In other words, there must be a certain gap but its distances and proximities must nonetheless be carefully managed. The link between ideas of belongingness and the risk associated with synthetic data is therefore problematised. It is a process replete with tension and contingency, because the very process of generating synthetic data is predicated on a kind of 'space of play' (Amoore, 2020); an ongoing sociotechnical negotiation between distances and proximities, between that which is 'too real' and that which is 'not real enough' – the faces that can be said to belong to someone and the faces that do not.

Third, and lastly, the specific idea of risk associated with synthetic data for machine-learning algorithms depends upon another specific demarcation: algorithmic risk is understood predominantly as a 'training dataset problem'. That is, this emergent notion of risk does not depend upon its actual impact on human and nonhuman subjects in the world, but rather in terms of *the kind of data* that is used for training machine-learning algorithms. Or, the differential impact and harms of algorithms are traced back to the dataset. As the synthetic data company Hazy put it, 'where there is no real data there is no real risk'. Although the training dataset of algorithmic systems remains a crucial point of critical intervention (Angwin et al., 2016; Denton et al., 2021; Gebru et al., 2020), the politics of synthetic data foregrounds the dangers of focusing exclusively on the dataset. That is because this overemphasis on the training dataset as the primary site of politics is precisely a central component of the reconfiguration of risk in machine-learning systems made possible by synthetic data. In this view, algorithms can be placed beyond the realm of risk when they are trained on supposedly varied, unbiased, balanced, representative, and therefore risk-free datasets. Within this logic, where there is no real data there is no real risk, insofar as a stark ontological and ethicopolitical dichotomy between real-world data (=risky) and synthetic data (=risk free) is established. Synthetic data can therefore be understood as a technology of risk in that they attempt to engender 'a realm of clear distinctions between safety and danger, truth and falsity, past and future' (Adam and van Loon, 2000: 7). Here, the biased algorithm figures as an inevitable outcome of the biases inherent in the training dataset. By shifting the domain of risk to the 'real' dataset, synthetic data promise to be the means by which algorithms can be rendered free from their own manufactured uncertainties. In this logic of risk, a strategic cut is performed. It is a cut which clearly delineates and distinguishes between the dataset and the algorithm itself, one which enables the reduction of the ethicopolitics of machine learning to the training dataset. *Fix the data, fix the algorithm*. This understanding of risk associated with synthetic data is therefore both predicated on and reinforces a clear break between data and algorithmic architecture.

However, this dichotomy between real (risky) and synthetic (risk free) is fundamentally problematic because the

cut between data and algorithm is problematic. The reconfiguration of risk as 'a training dataset problem' is to render the dangers of machine-learning algorithms not only tangible and calculable in a very specific way but also an attempt to render them controllable, whilst simultaneously perpetuating a view of algorithmic systems as having the potential of being placed beyond the realm of risk. Although this notion of 'algorithmic bias is a data problem' constitutes a field of rich critical investigation, it can also obfuscate the web of interlinking sociotechnical mechanisms, processes, and power relations that underlie the development and deployment of machine-learning models. For instance, as Hooker (2021) has argued, the issue of reducing bias to the training data fails to address how, say, the impact of model design can just as easily result in bias as well. As she states, 'algorithms are not impartial, and some design choices are better than others' (p. 1). In contrast to the emergent logic of synthetic data, whereby algorithms can be placed beyond the realm of risk, there is a continual need to emphasise the necessary partiality of all machine-learning algorithms. 'Algorithms', Louise Amoore writes (2020: 19), 'are giving accounts of themselves all the time. These accounts are partial, contingent, oblique, incomplete, and ungrounded'. Irrespective of the real or synthetic data on which they were trained, algorithms are always already partial and experimental, engendering new boundaries of normal and abnormal, good and bad, as well as new modes of seeing and doing. This means therefore, as Jasanoff (2016: 34) puts it, 'the idea of zero risk – that is, of a flawlessly functioning technological environment in which machines and devices do exactly what they are supposed to do and no one gets hurt – remains just that: an unattainable dream' (p. 34). Contrary to the logic generated by synthetic data, risk is the unavoidable structural condition of machine-learning algorithms.

Conclusion

In this paper, I have explored some of the ethicopolitical implications of synthetic data for machine learning, defined broadly as data generated *by* algorithms and *for* algorithms. I have focused particularly on two key interlinking ideas: first, how synthetic data promise to emerge as a rich source of exposure to variability for the algorithm. In this view, synthetic data are intimately related to logics of shock, exposure, and resilience. They promise to make algorithmic systems more resilient to volatility in the world by generating the rare, the unusual, and the infinitely variable. Second, this paper has explored how synthetic data enable the emergence of a new logic of risk whereby synthetic data promise to place machine-learning algorithms beyond the realm of risk. In other words, synthetic data embody an emergent logic which conceptualises the possibility of de-risking algorithmic models, to render them risk free. More broadly, these two central ideas emphasise how machine learning is being envisioned differently in the light of synthetic data, that is, how synthetic data constitute an inflection point as they participate in the reconfiguration of the conditions of possibility for machine-learning algorithms in contemporary society. From this, it is evident that synthetic data for machine learning are increasingly becoming an intricate part of our contemporary algorithmic societies. It is therefore crucial to better understand their variegated role and impact, as well as how they promise to variously reconfigure the conditions of possibility for algorithms.

Moreover, there is also a need to ask: as a technology of risk, what new risks are engendered through synthetic data? What new ways of governing risk in society are made possible through synthetic data? One area of future research could be how synthetic data participate in the generation (and amplification) of the anomaly and the anomalous. A generative algorithmic model such as a GAN can be exposed to the patterns, regularities, and features in a data distribution and then iteratively learn to generate either synthetic data that approximate that distribution or synthetic edge cases that are not sufficiently covered in that distribution. The idea is that these edge cases can be used to augment the training dataset of a machine-learning algorithm used for anomaly detection in domains such as immigration, terrorism, finance, and insurance. Put differently, variability or the rare is generated so that it can be amplified in the training data. If there are not enough, say, potential terrorists in the dataset, these can be generated. Synthetic data participate in the emergence of spaces of possibility for detecting the anomalous in increasingly higher resolution and granularity. This raises the question what kinds of anomaly can be generated? (a question which is beyond the remit of this paper). In short, what can this anomaly be generated to look like and used for? These questions attain additional depth when considering the promise of companies such as Datagen to generate any diversity needed such as more synthetic black faces. Therefore, we especially need to ask what new modes and techniques of racialisation and profiling are made possible by synthetic data for machine-learning algorithms (Amaro, 2020; Phan and Wark, 2021).

It is also crucial, however, that studies of synthetic data and machine learning go beyond a framework of exclusion, bias, and marginalisation. Instead, the power and politics of synthetic data for machine learning must be examined in relation to their overarching logic of *generativity*. Examining the ethicopolitics of machine learning, seen through the prism of synthetic data, foregrounds not only how algorithms foreclose possible futures, but also how they participate in opening up new possible futures with new parameters of good and bad, new curves of (ab)normality, new ideas of difference, as well as new kinds of risk or manufactured uncertainties. Yet, these new openings and possibilites are not simply emancipatory. They need to be critically explored in terms of the notion of inclusion, both in terms of what inclusion promises but also the

ethicopolitical complications inherent in the term: that any variance (e.g. minority population) not represented in a training dataset can be algorithmically generated and included. In short, inclusion can be achieved because it can be generated. Instead of taking such claims and promises at face value they must be interrogated for their conceptual and structural implications. As Hoffmann (2021) has argued, inclusion is fundamentally a non-disruptive logic; that is, it represents 'an ethics of social change that does not upset the social order' (p. 3550). A similar point is made by Ahmed (2012: 163) in *On Being Included,* where she powerfully claims that inclusion is not simply a logic but also a technology, one which produces included subjects that nonetheless have to 'consent to the terms of inclusion'. Similarly, as I have sought to show in this paper, the claims that synthetic data are ushering in a new era of generated inclusion and non-risk for machine-learning algorithms is both misguided and dangerous. For it obfuscates how synthetic data are fundamentally a technology of risk, producing the parameters and conditions of what gets to count as risk in a certain context.

Moreover, the promises of inclusion inherent in the discourses surrounding synthetic are ethicopolitically problematic in three further senses: firstly, in the sense that they seek to amortise the claims that can be made by historically marginalised communities; secondly, they render protected characteristics redundant in AI systems; and lastly, they reduce the 'ethics' of AI to elements such as entries to the training data that do not disrupt the fundamental power and politics of algorithms. As such, synthetic data for machine learning may be participating in generating a new kind of world, but this will likely be a world where the overall logics of AI remain intact and our current grounds for critique of AI will be deeply unsettled and maybe even uprooted. The politics of synthetic data for machine learning is therefore in dire need of further critical attention. Moreover, the emergence of synthetic data should urge us to rethink and reimagine the relationship between machine learning, power, and ethics more broadly.

Acknowledgements

This paper has greatly benefitted from helpful advice and discussions with Louise Amoore, Ludovico Rella, and Alexander Campolo. My thanks also go to the two anonymous reviewers for their engagement and comment. The research has received funding from the European Research Council (ERC) under Horizon 2020, Advanced Investigator Grant ERC-2019-ADG-883107-ALGOSOC.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Horizon 2020 Framework Programme, (grant number ERC-2019-AdG-883107-ALGOSOC)

ORCID iD

Benjamin N Jacobsen https://orcid.org/0000-0002-6656-8892

Notes

1. The studies that currently come closest to the domain of synthetic data are those exploring the sociopolitical ramifications of deepfakes (see for instance Diakopoulos and Johnson, 2020; de Vries, 2020).
2. Other examples of the actual use of synthetic datasets for the training of facial recognition algorithms include governments that incorporate synthetic faces in the training of biometric passport recognition algorithms.
3. Backpropagation refers to the way in which neural network algorithms learn to recognise patterns such as people and objects within images. Backpropagation sends information backwards through the layers of the neural net, indicating how the algorithm should incrementally tweak its internal parameters or weights to achieve the optimal result. In short, backpropagation iteratively minimises the gap between actual output and desired output (LeCun et al., 2015).
4. The attempt to capture and generate the variability of the world as probability distributions necessarily raises much-discussed questions concerning the nature of reality (which unfortunately is beyond the remit of this paper). However, whilst we need to remain highly suspect of claims that reality *can* be fully captured by algorithmic systems and generated as synthetic training data, I argue it is crucial to examine what precisely such claims *do* and *make possible.* That is, viewing the world in terms of probability distributions and variability make possible a particular approach the world. Such claims operate as what Boltanski and Thevenot (2006) call 'regimes of justification', providing the justificatory scaffolding for algorithmic systems to continue capturing and generating the world as data distributions that function as representations of the world. The main issue therefore is not if probability distributions accurately capture the world, but rather that they participate in the construction of a world that only comes to matter as an array of probability distributions.
5. Bansal and Ogale currently work as research scientists at the AV company Waymo. Kirzhevsky, who won the ImageNet competition in 2012 with his groundbreaking convolutional neural network algorithm AlexNet, worked at Google Brain when this paper was published. He currently works at the deepfake research company Dessa.
6. By using real liver lesion data only, the accuracy rate of the convolutional neural net algorithm correctly classifying liver lesions in CT scans was 78.6%. Adding synthetic liver lesions, however, the accuracy rate increased to 85.7% (Frid-Adar et al., 2018: 326)
7. Technological risks are commonly understood as 'those that arise specifically from the use and operation of human-made instruments or systems, as opposed to risks from natural disasters that we do not presume to control, such as earthquakes, storms, or epidemic disease' (Jasanoff, 2016: 34). Yet, Sheila

Jasanoff acknowledges that such a distinction 'is hard to sustain in an interdependent world' (p. 34).
8. See https://hazy.com/company
9. See https://www.syndata.co/
10. See https://www.synthesized.io/. One of their recent case studies is titled 'Test Rigorously and Risk-Free with High-Quality Synthesized Data'.
11. Instead of viewing risks as having an external reality independent of the tools, apparatuses, and technologies that have been developed, we should see risk as 'manufactured uncertainties', that is, as a set of practices of 'manufacturing particular uncertainties that may have harmful consequences to 'life' in the broadest sense of the term' (Adam and van Loon, 2000: p. 2).
12. See https://datagen.tech/.
13. For more on the embodied ramifications of biometrics, see Btijah Ajana's (2013) book *Governing Through Biometrics: The Biopolitics of Identity.*

References

Adam B and van Loon J (2000) Introduction: Repositioning risk; the challenge for social theory. In: Adam B, Beck U and van Loon J (eds) *The Risk Society and Beyond: Critical Issues for Social Theory*. London: Sage, pp. 1–33.

Ahmed S (2012) *On Being Included: Racism and Diversity in Institutional Life*. Durham and London: Duke University Press.

Ajana B (2013) *Governing Through Biometrics: The Biopolitics of Identity*. Basingstoke: Palgrave Macmillan.

Amaro R (2020) Threshold Value. E-Flux Architecture. Available at: https://www.e-flux.com/architecture/education/322664/threshold-value/.

Amoore L (2009) Lines of sight: On the visualization of unknown futures. *Citizenship Studies* 13(1): 17–30.

Amoore L (2011) Data derivatives: On the emergence of a security risk calculus for our times. *Theory, Culture & Society* 28(6): 24–43.

Amoore L (2020) *Cloud Ethics: Machine Learning and the Attributes of Ourselves and Others*. Durham and London: Duke University Press.

Andrejevic M and Selwyn N (2019) Facial recognition technology in schools: Critical questions and concerns. *Learning, Media and Technology* 45(2): 115–128.

Angwin J, Larson J, Mattu S et al. (2016) Machine Bias. ProPublica. Available at: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Bansal M, Krizhevsky A and Ogale A (2019) Chaffeurnet: Learning to drive by imitating the best and synthesizing the worst. In Robotics: Science and Systems, Freiburg im Breisgau, June 22–26, 2019, pp. 1–20.

Beck (1992) *Risk Society: Towards a New Modernity*. London: Sage Publications.

Beer D (2016) How should we do the history of big data? *Big Data & Society* 3(1): 1–10.

Beer D (2019) *The Data Gaze: Capitalism, Power and Perception*. London: Sage.

Berlant L (2007) On the case. *Critical Inquiry* 33(4): 663–672.

Boltanksi L and Thevenot L (2006) *On Justification: Economies of Worth*. Princeton and Oxford: Princeton University Press.

Boyd D and Crawford K (2012) Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society* 15(5): 662–679.

Bruder J and Halpern O (2021) Optimal brain damage: Theorizing our nervous present. *Culture Machine* 20: 1–25. Available at: https://culturemachine.net/wp-content/uploads/2021/09/Bruder-Halpern.pdf.

Bucher T (2018) *If…Then: Algorithmic Power and Politics*. Oxford: Oxford University Press.

Buolamwini J and Gebru T (2018) Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research* 81: 1–15.

Chen RJ, Lu MY, Chen TY, et al. (2021) Synthetic data in machine learning for medicine and healthcare. *Nature* 5: 493–397.

Courville A, Goodfellow I and Bengio Y (2016) *Deep Learning*. Cambridge, MA and London: MIT Press. Available at: https://www.deeplearningbook.org/.

Crary J (1990) *Techniques of the Observer: On Vision and Modernity in the Nineteenth Century*. Cambridge, MA and London: MIT Press.

Crawford K (2021) *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven and London: Yale University Press.

Crawford K and Paglen T (2019) Excavating AI: The politics of images in machine learning training sets. September 19. Available at: https://excavating.ai/.

Denton E, Hanna A, Amironesei R, et al. (2021) On the genealogy of machine learning datasets: A critical history of ImageNet. *Big Data & Society* 8(2): 1–14.

De Vries K (2020) You never fake alone: Creative AI in action. *Information, Communication & Society* 23(14): 2110–2127.

Diakopoulos N and Johnson D (2020) Anticipating and addressing the ethical implications of deepfakes in the context of elections. *New Media & Society* 23(7): 1–27.

Dourish P (2016) Algorithms and their others: Algorithmic culture in context. *Big Data & Society* 3(2): 1–11.

Ewald F (1991) Insurance and risk. In: Burchell G, Gordon C and Miller P (eds) *The Foucault Effect: Studies in Governmentality*. Chicago: The University of Chicago Press, pp. 197–211.

Foucault M (2007) *Security, Territory, Population: Lectures at the College de France 1977–1978*. Basingstoke: Palgrave Macmillan.

Frid-Adar M, Diamant I, Amitai M, et al. (2018) GAN-based Synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing* 321(10): 321–331.

Gebru T, Morgenstern J, Vecchione B, et al. (2020) Datasheets for datasets. ArXiv: 1–18.

Gentric S (2021) Deep learning, a key technology behind IDEMIA's algorithms. Idemia. Available at: https://www.idemia.com/news/deep-learning-key-technology-behind-idemias-algorithms-2021-07-26.

Goodfellow I, Pouget-Abadie J, Mirza M, et al. (2014) Generative adversarial nets. Proceedings of the International Conference on Neural Information Processing Systems (NIPS), pp. 1–9.

Halpern O (2021) Planetary intelligence. In: Roberge J and Castelle M (eds) *The Cultural Life of Machine Learning: An Incursion Into Critical AI Studies*. London: Palgrave Macmillan, pp. 227–256.

Halpern O (2022) The future will not be calculated: Neural nets, neoliberalism, and reactionary politics. *Critical Inquiry* 48(2): 334–359.

Halpern O, Jagoda P, Kirkwood JW, et al. (2022) Surplus data: An Introduction. *Critical Inquiry* 48(2): 197–210.

Halpern O, Mitchell R and Geoghegan BD (2017) The smartness mandate: Notes toward a critique. *Grey Room* 68: 106–129.

Heaven WD (2020) Our weird behavior during the pandemic is messing with AI models. MIT Technology Review. Available at: https://www.technologyreview.com/2020/05/11/1001563/covid-pandemic-broken-ai-machine-learning-amazon-retail-fraud-humans-in-the-loop/.

Heaven WD (2021) Synthetic Data for AI. MIT Technology Review. Available at: https://www.technologyreview.com/2022/02/23/1044965/ai-synthetic-data-2/.

Hoffmann AL (2021). Terms of inclusion: Data, discourse, violence. *New Media & Society* 23(12): 3539–3556.

Hooker S (2021) Moving beyond 'algorithmic bias is a data problem'. *Patterns* 2: 1–4.

Hutchinson B, Smart A, Hanna A, et al. (2021) Towards accountability for machine learning datasets. Proceedings of the 2021 conference on fairness, accountability, and transparency, pp. 560–575.

Jasanoff S (2016) *The Ethics of Invention: Technology and the Human Future*. New York and London: W. W. Norton.

Kitchin R (2014) Big data, new epistemologies and paradigm shifts. *Big Data & Society* 1(1): 1–12.

Kowalski M, Garbin SJ, Estellers V, et al (2020) CONFIG: Controllable neural face image generation. In: *2020 european conference on computer vision*, 23–28 August, pp. 299–315.

LeCun Y, Bengio Y and Hinton G (2015) Deep learning. *Nature* 521: 436–444.

Newman P (2021) Hard Miles without Hard Miles. *NVIDIA GTC Conference 2021*. November Monday 8th to Thursday 11th (Observed 10 November 2021).

Nikolenko SI (2019) Synthetic data for deep learning. ArXiv: 1–156.

Phan T and Wark S (2021) What personalisation can do for *you!* or: How to do racial discrimination without 'race'? *Culture Machine* 20: 1–29.

Tremblay J, Prakash A, Acuna D, et al. (2018) Training deep networks with synthetic data: Bridging the reality gap by domain randomization. *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1082–1090.

White A (2021) By 2024, 60% of the data used for the development of AI and analytics projects will be synthetically generated. Gartner. Available at: https://blogs.gartner.com/andrew_white/2021/07/24/by-2024-60-of-the-data-used-for-the-development-of-ai-and-analytics-projects-will-be-synthetically-generated/?_ga=2.103596488.916571214.1647348850-633920548.1645012714.

Zeilinger M (2021) Generative adversarial copy machines. *Culture Machine* 20: 1–23.

Zuboff S (2019) *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. London: Profile Books.