Goldsmiths Research Online

Goldsmiths Research Online (GRO)
is the institutional research repository for
Goldsmiths, University of London

Citation

Persistent URL

Versions

Goldsmiths
UNIVERSITY OF LONDON

Deep Bureaucracy and Autonomist AI

Dan McQuillan

AI is here at last. But what exactly is it, and what is it going to change? The AI emerging from Silicon Valley and Beijing is certainly uncanny, but it doesn't meet the expectations created by the movies. It's heralded as the birth of near-human intelligence but seems to be warping social relations wherever it appears, whether in social media or new structures of precarious employment. Given the global rise in reactionary politics, the future arriving at the same moment that the far-right past has returned should provoke skepticism. While talking heads focus on some imagined superintelligence, actually existing AI is rolling over marginalized communities in the here and now. An AI future seems nonnegotiable, but the nature of neoliberalism and the legacy of colonialism should already warn against a destiny driven by deep learning. I'm suggesting that AI is an intensification of what's gone before rather than any kind of quantum leap, specifically, as an intensification of bureaucracy and the derivatives market. Doubling down on distributed manipulation is a political move that brings specific harms in the form of segregation and automated scarcity, so I also propose a standpoint and a set of tactics that push for an alter-AI: that is, for an approach that recomposes AI according to an alternative vision of togetherness.

Welfare Bureaucracy

At first glance, bureaucracy and AI don't seem to have that much in common. Bureaucracy is the gray zone of organizational administration that relies on boring forms and procedures, while AI is the cutting-edge computational technology whose intelligence may be about to outstrip that of human beings. But rather than the bridge to an alternative sci-fi future, AI can be understood as an upgrade to the existing bureaucratic order, that is, as an iteration of the logic that unifies governments and corporations. This actual AI is not a filmic fantasy but applied machine learning of various kinds, including artificial neural networks or deep learning. Bureaucracy, according to its biographer Max Weber, emerged as a mechanism for the state to impose control and rationality on increasingly complex societies. AI is part of the same quest for traction in a

turbulent world. Exactly like bureaucracy, AI is a mode of goal-oriented rational ordering that claims neutrality and objectivity through detached abstraction.

The functionaries who are already effecting this transformation do not see their work as the state administration of a Terminator-like intelligence. Rather, a shift to AI governmentality makes sense in terms of the perceived virtues of bureaucracy: as a moral comportment that is evidence-driven and rational, one that valorizes indifference as the means to effective implementation of policy and cultivates a wariness toward emotional entanglement. There is, of course, a growing awareness that AI's uncanny powers seem to be stained by bias, because however clever the computers become, they are learning from data generated by people living their messy and prejudiced lives. But public administrations already operate under these conditions. Indeed, the virtue of bureaucracy is seen as limiting the effect of any partiality among its professionals through a technocracy and rigid rules. The wave of research by the machine-learning community into de-biasing datasets exactly complements the bureaucratic mindset through its technocratic solutionism and the formal audits taken as sufficient guarantors of balance.

The distortions that cannot be undone are the architectural ones that unite both bureaucracy and AI: the reliance on distance and abstraction. Bureaucracy achieves its authority by abstracting from the detail of the social world to apply policy at sufficient scale to make modern governmentality possible. It is foundational to bureaucracy to extend "the distance at which human action is able to bring effect."[1] Machine learning begins with data points that are already a set of quantified abstractions from some target context. It constructs a loss function that represents in some abstract sense the sum of distances in a high-dimensional feature space between data points and the predictions of the algorithm. The "learning" in "machine learning" comes from computational optimization that iteratively minimizes this loss function. The resulting parameters are the model that can be used for classification and prediction. These architectures of abstraction and the leveraging of distance—of institutional form and algorithmic computation—are becoming fused through the force of financial crisis. AI is seen as a key solution to the challenges of austerity, where every state actor is required to do more with less. Under the cover of crisis, the technocrats of administration and the technocrats of computing are

[1]   Bauman, Zygmunt. 1989. Modernity and the Holocaust, Cornell University Press (cited in Gay, Paul du. 2000. In Praise of Bureaucracy. Copenhagen: SAGE Publications Ltd, p36)

collaborating on operations of abstraction, distance, and optimization, both insulated by a reliance on technical morality.

Thus, the impact of actual AI—that is, how machine learning is applied in the world—is shifting from the white noise of social networks to the transformation of state functions. Applying algorithms to welfare payments, for example, is presented as a benign upgrade of a bloated system; an entirely administrative matter that needs no democratic debate, let alone community co-design. While many of the current welfare-state algorithms are clumsy codifications of existing bureaucratic assumptions, they are already being adopted for areas such as child protection and healthcare. Machine-learning systems are the next step in this extant but largely invisible shift to algorithms in public services. In Sweden, algorithms have been issuing warnings and withholding benefit payments; in Spain, they've been discarding applications from poor households for electricity subsidies; and in Austria, they've been scoring job seekers to decide whether or not they get help, and doing so in a way that discriminates against women and the disabled.

In the name of efficiency and narrowing the pool of benefits claimants, new and intrusive forms of conditionality are introduced. While the overt reason is mostly austerity, the process is also being exploited for social reengineering by algorithms: that is, as a kind of automation shock doctrine. Recent reports by the UN's Special Rapporteur on extreme poverty and human rights describe myriad ways in which digital transformation and the shift to algorithmic governance conceals significant alterations to the political-structural landscape. In the Netherlands, an algorithm called SyRI (Systeem Risico Indicatie, or System Risk Indication) tapped into a large array of government databases to draw up lists of people suspected of housing or benefits fraud by calculating their similarity to data profiles of previous fraudulent claims. High water consumption, for example, may have been seen as an indicator that someone receiving a single person's benefit was cohabiting with someone else. In a micro-demonstration of policy-by-data-science, SyRI has only been applied to specific neighborhoods with high levels of poverty and deprivation, and, when the algorithm is applied, the data analyst sets parameters that define the local "risk range." The fashionable liberal focus on algorithmic ethics diverts attention from this increasingly widespread substitution of technical operations for democratic discourse.

Derivatives

Bureaucracies are already not what they used to be. Under neoliberalism they have been increasingly transformed into arms-length agencies with entrepreneurial goals, where senior bureaucrats don't implement policies but design a market of incentives and penalties intended to optimize contractual outcomes from outsourced services. Adding AI will take this marketization to a new level by mobilizing quantifications of risk as a form of derivative. Public institutions pervaded by artificial neural networks will become futures exchanges for bundles of citizen risk factors. The way that neural networks will so intensively financialize administration and governance is related to their internal operations. When training a deep learning algorithm, diverse features of the world are normalized to numbers between zero and one such that they can be traded against each other in an iterative process called gradient descent. You don't need to tell neural networks what features to use; you just force enough labeled data forward through the layers then pass the errors backward to optimize the fit. Neural networks can handle the messy patterns that defeat other algorithms. It's neural networks that make facial recognition possible and power the conversational abilities of personal assistants like Alexa and Siri. This leads to a kind of hubris among practitioners, a belief that, given a sufficiently large dataset and a target to optimize on, they can derive a model that will solve pretty much any problem. But this process of hyper-abstracting the diversity of the world's materiality into planes of numerical consistency that can be mathematically manipulated and traded is also the dark pattern behind derivatives.

Derivatives are, of course, the financial products that have come to dominate international markets. Their value is derived from an underlying asset or group of assets and varies according to their fluctuations. They are ways of dealing with risk by taking positions on likely future values. One particular form of derivative is the Collateralized Debt Obligation (CDO), a complex structure that bundles together loans and other assets. Subprime mortgage-backed securities become the source of huge inflation in the market for CDOs and it was the bursting of that bubble that triggered the financial crisis of 2007–2008. Perhaps surprisingly, derivatives, including CDOs, are still the dynamo of world markets. This same logic is now migrating directly into social life via new forms of algorithmic risk assessment based on data, not about industries or capital assets but about everyday lives and behaviors. These assets are not being bet on by financial traders but are being quantified and optimized by institutions in making decisions about

their matters of concern, whether it's a large company calculating which employee is the most likely to leave or social services calculating which parents are most likely to abuse their children. Like futures markets, AI models the probabilities of particular outcomes, projecting values into likelihoods that are then treated as more real than the actualities that they supposedly represent.

The obsession with accuracy metrics so visible in the machine-learning profession shouldn't be misunderstood as an attempt to asymptotically approach a reliable representation of reality. This isn't science but the construction of data derivatives, an operation of stabilization that recomposes abstracted risk characteristics into commensurable bundles. The confident promoters of AI solutionism don't consider whether decisions based on the complexity of leveraged calculations will introduce new forms of instability or have unintended consequences. Prediction-oriented risk assessments introduce a feedback loop of iterative amplification because they can't account for their own interventions in the system. Recent history offers lessons here. Many scholars see risk assessment as the key historical driver in the massive growth and racial disparities of the United States prison population, through the knock-on effects it had on people's lives. Incarceration means weakened familty ties, reduced employment opportunties and ratcheting up of recidivism. Hence 'risk profiling' increases risk. The social effects of AI are further deepened through segregation and a reduction of dignity.

Segregation and Dignity

When AI applies its optimized model to input data, it makes predictions by discriminating between one target class and another. In neural networks the final layers are designed to force an outcome from the non-linearity of the neurons' activations so that the output classifications act like probabilities, that is, they are all values between zero and one and together sum to a total of one. The predicted label is the class with the highest probability. There will always be a predicted output, however poor the correlations, which explains the certainty with which a neural network will pick a result even when presented with deliberately misleading input data. The way AI arrives at its classifications is by calculating distances: by determining some abstract metric of difference as a distance. The basis of machine learning is a constructed homophily, a forcing of closeness in data space. But treating an abstract distance in data space as an innate affinity is the logic of statistical segregation. As Wendy Kyong Chun points out, racism can also be seen as a technology of segregation. Race itself is a kind of tool, "one designed to stratify and sanctify

social injustice as part of the architecture of everyday life," and AI constructs the world in the same way.[2] The relationship of AI to racism won't be fully dispersed by more inclusive training data; it's more fundamental. AI systems will automate discrimination not simply because they are being adopted in already racist societies but because their core technical operations work that way.

Of course, labeling and classification are mundane methods used in filtering through daily reality, but raised to institutional principles for the distribution of goods and harms, they have always reinforced asymmetries of power. Inflated credibility will be given to machinic assessments over the already suspect voice of the service user. Whether you're a recipient of benefit payments or a precariously employed Uber driver, your status is not as a rights holder but as a supplicant to an algorithm. It's a minimization not just of the loss function but of dignity. The resultant ranking of the deserving and the undeserving becomes a calculative Victorianism, assigning morality via statistical metrics. The social logic of derivatives induced by algorithms and pervasive risk scoring produces new forms of debt, not only the kind of monetary debt miscalculated by the so-called "robo-debt" algorithm in Australia, but a debt that is the inverse of "social credit," an algorithmic negative that cuts off access to services. A recent Airbnb patent uses machine learning to use credit checks and crawls of social media accounts in combination to infer whether a client displays the untrustworthy traits of "narcissism, Machiavellianism, or psychopathy."[3] Meanwhile, socially marginalized groups are reporting increased exclusion from the service.

The Targeted Compliance Framework in Australia creates a "mutual obligation" for claimants to update a digital dashboard with their job-seeking activities and to check their compliance status. Any slip in this self-monitoring leads to a suspension of payments even though many users lack digital literacy or find that the automated system fails to take real-life situations into account. The algorithmic gaze is used to select points of inconsistency that can be leveraged against those who are already vulnerable, extending control over an increasing raft of behavioral details in much the same way as the smallest deviation of a worker in an Amazon warehouse can degrade their work

2   Benjamin, Ruha. 2019. Race After Technology: Abolitionist Tools for the New Jim Code. 1 edition. Medford, MA: Polity. p17

3   Baveja, Sarabjit Singh, Anish Das Sarma, and Nilesh Dalvi. 2019. United States Patent: 10169708 - Determining trustworthiness and compatibility of a person. 10169708, issued 1 January 2019.

credit and lead to automated severance. Algorithmic governance increases the blurring between state and corporation in enacting social control over the least powerful in the same way that austerity and precarity have erased the boundary between employment and poverty. The mathematics of AI seem abstract, but ideologically charged terms such as efficiency, targeting, and minimization slide freely between political rhetoric and machine learning. The classifications operationalized by AI are performative. Decision boundaries won't remain mathematical constructs but will determine the distribution of services and sanctions. Clashes of classification will operate across the spectrum of all who are already pushed toward marginal positions and each example will bring its own collateral damage. The only accountability will be economic, led by considerations of financial efficiency, as the accuracy figures of algorithms fuse with cost-related performance targets.

As machinic classifications are integrated into institutions, and as algorithmic bureaucracies infer people's membership in narrow categories in order to get their needs met, they multiply moments of administrative violence, especially for those who are already the most datafied. The mathematic morality of AI is not only a mechanism of post-austerity but a potential tool for social reengineering. This is not exactly an alien concept in Silicon Valley itself, the heartland of AI, where strong strains of elitism overlap with skepticism of a pluralist social order and leanings toward corporatist authoritarianism, based on assumptions about natural hierarchies of intelligence. Ultimately, the algorithmic coupling of vectorial distances and social differences can become the targeting of those who are claimed to be polluting national purity. The endpoint of an automated welfare state that is purely aligned with a national mission of economic success is to optimize out the "useless eaters." For those many thousand disabled people in the United Kingdom, for example, who have been reclassified as "fit for work," the administrative violence of such efficiencies then, very sadly, often becomes self-directed, as can be seen from the statistics on suicides. [4]It's not just "automating inequality" to apply data-driven force to the poorest; any logic that begins with the most vulnerable demographics in the great task of optimizing the whole populace is a fascist one.

[4]   Barr, Ben, David Taylor-Robinson, Alex Scott-Samuel, Martin McKee, and David Stuckler. 2012. 'Suicides Associated with the 2008-10 Economic Recession in England: Time Trend Analysis'. BMJ 345 (August). https://doi.org/10.1136/bmj.e5142.

Matters of Care

How can political action best contest a neural-network new order? A good starting point is to treat AI not as a matter of intelligence but as a matter of care. By "care" I mean the invisibilized labor that is an inevitable consequence of social interdependence and which is usually gendered, racialized, and devalued. Care exists in the shadow of the kind of detachment and abstraction that is valorized by AI and yet it is the tangle of tasks without which people's lives would not work. A politics of care starts with concern about the exclusions and boundaries of a stratified society. The first question for any AI should not be how many marginal accuracy points it's gained on an accuracy benchmark but how its reordering will increase the present asymmetry in the burden of care. It is indeed ironic that AI's intensification of segregation is itself built on the gendered, racialized, and invisibilized labor of clickworkers in the Global South who are busy labeling the necessary datasets.

Oppsing this new order requires ways of knowing that can challenge the erasure of lived experience by the ideology of efficiency, and generate a counter-project to the algorithmic production of carelessness. Treating AI as a matter of care is not only about concern for those who might be harmed but about the character of its core processes. The form of knowing enacted by AI depends on a separation between representations, the weights in the models, and that which they claim to represent. But the way AI acts in the world is such that subjects are actually produced and performed by these very same structures of knowing, invalidating the models as statistical insights but making them very effective as forms of subjection. Moving beyond the injustices enabled by data systems means moving beyond representation to social recomposition.

This isn't about machines versus people but about the way specific configurations of the techno-social can corrode sociality and mutual aid. Instead of asking for a more efficient way to model reality, it is better to concentrate on enacting a different world. AI is a form of apparatus in the sense that Karan Barad articulated; it produces at an ontological level the world it is designed to measure. Altering AI in practice would redraw the exclusions that produce phenomena and experience: that is, reclaiming agency through agential realism where discourse is not a matter of language but of changing the constraints on what can emerge. AI is a discursive apparatus whose role in sedimenting reality requires participatory reworking. Rather than iterative assimilation into the new order, the task at hand is to develop forms of participation that are lines of flight.

Workers' and People's Councils

One way to start this recomposition of AI is through people's councils. People's councils are bottom-up, federated structures that act as direct democratic assemblies, a model that goes back to the face-to-face democracy of the Athenian *ekklesia* (popular assemblies) but reemerges again and again in historical moments as a self-organized alternative to authoritarianism. Such assemblies are horizontal structures in which everyone has an equal say about the matter being decided. They are structures that deliberately reverse states of exclusion and exception, and act as counterweights to bureaucratic distance and abstraction. People's councils would interrupt statist AI with the situated perspectives of the ignored, promoting inclusion over efficiency and regenerating relationality so as to restructure common ways of knowing. Algorithmic bureaucracy constitutes one of the highest historical forms of the abstraction of social organization and needs to be counterbalanced by the unmediated relationships of popular assemblies. This is not a reconfiguring of representative democracy. People's councils are not representative because they challenge the validity of representation, but they are transformative because they are constitutive of a different commonality.

The alteration of AI also requires workers' councils. While neoclassical economics tries to construct liberal democracy as a common rationality of growth and prosperity for all, it foundationally excludes the workplace from the democratic sphere. Attempts by workers in AI to hold their employers to account run head on into this central contradiction. Workers' councils are the means by which AI workers can develop a vision as powerful as the transhumanist hubris of their parent corporations. Instead of being devalued as data and traded as risk factors, the objective must be to develop the creative and imaginative activities that constitute self-valorization, that is, the autonomous reinvention of conditions in the here and now. Contemporary tech-worker dissent will become transformative when it sees itself as creating the possibility for a new society within the shell of the old. In this mode of change, acting at "scale" doesn't mean the exponential computation of machine-learning models or the instrumentalist graph structures of social networks but of a democratic federation of councils.

AI is not coming to steal workers' jobs. It's coming to erase any distinction between work and precarious dependency. The conditions of precarious labor will be governed by the same logic of optimization as the weaponized welfare system. Only a form of self-organization that refuses

these this logic has any prospect of pushing back. The potential of council organization for both community life and economic activity has been put into practice in the autonomous zone of Rojava in north eastern Syria, under the auspices of democratic confederalism. This confederalism is based on the principle that all groups in society find expression through local meetings, general conventions, and councils. One of the most striking aspects has been the self-organization of women in this system, both within autonomous structures and through joint leadership of administrative and economic councils at all levels of the federal system. The council form is the anti-derivative; lived reality is addressed as a whole not through deconstruction, and incommensurability is overcome by consensus not stabilized quantification. This is the basis for an alter-AI, one that gains traction by starting from existing points of struggle at the margins: struggles over disability, welfare, housing, climate change, and immigration. It is not enough to reform the performance of hyper-abstracted optimization; it needs to destroyed.

Deserting AI

What needs deserting is AI realism—the sense that an AI-driven future is inevitable. An AI-driven future will mean every aspect of life being traded as some kind of quantified risk. The new order isn't being announced in a press release from Silicon Valley or debated in parliaments but is being infrastructured through a reinvention of mundane state functions. AI is political, and the challenge to algorithmic stratification will come through political organization as much as a commitment to ethics and fairness. Neoliberalism itself was a line of flight from increased worker and community power, and it is the recomposition of those same struggles that can rein in its computational intensification. A reliance on abstraction and generalizability is not a route to superhuman powers but of sign of vulnerability, a fear of being called to account by those presently carrying the burden of care. The way forward for a recomposed AI is not predictive but prefigurative. The aim must be to seek algorithmic operations that abandon their service to detachment and contribute instead to the reconstruction of the commons through militant caring.