



Strange affair of man with the machine

Karamjit S. Gill¹

Accepted: 3 October 2020 / Published online: 13 October 2020
© Springer-Verlag London Ltd., part of Springer Nature 2020

We create devices and then they create us

Narcissus-like, we gaze into a pool of technology and
see ourselves

We acquiesce in our own demise, setting out as
participants

and metamorphosing into victims

(Cooley 2013)

In paying our memorial tribute to Mike Cooley, the founding chairman of AI&Society, we are reminded of his poem “INSULTING MACHINES” above. It makes us reflect on the strange affair of man with the machine. Whilst we may feel mesmerised by the computing power of our creation, the AI machine, we may also feel that we are being seduced to participate in our own demise, as helpless victims. Professor Bell (2018) notes that although the idea of AI was ‘codified’ as a computational paradigm at the historical Dartford conference in 1956, its focus on abstraction and computation was an attempt to ‘make machines use language, form abstractions and concepts, solve the kinds of problems now reserved for humans, and improve themselves.’ Bell notes that while Wiener’s cybernetics cultivated the interface between biological and technical systems, McCarthy, Minsky and Shannon appeared in 1956 to create an intellectual agenda to announce computational AI. What was lost in this computational enterprise was the social piece, the human piece and the biological component inherent in cybernetics, so this earlier AI became just a technical artefact. Bell reflects that the birth of computational AI coincided with cybernetics going underground in Europe and North America, and any notion that you could use the new power of computation to drive social science looked more like socialism—in a period when it was not a good thing to be. Although the initial computational formulation was a product of a particular time and place, it has influenced the framing of research agendas ever since. But what

exactly is AI in 2020, and why does it loom very large in our conversations about the future? Bell says that the 1956 blue print of AI coincided with the interest of US Defense Department in pursuing the simultaneous machine translation from Russian (not just logic but context); 1960–70s was a period of reckoning that the translation culture and its contexts were not realisable; 1990 saw the pursuit of Intelligent AI; twenty-first century AIs appropriated an abundance of data to train algorithms. AI is now about ‘Can I know your desires for goods and services, an ultimate manifestation of capitalism.’ Bell further says that the question now is ‘Not what AI but Whose AI? What work is it doing and why?’ Would AI look differently in different countries—different data sets and different logics, what do AIs know, and what might they do? These are questions of consciousness and not of intelligence: can we imagine non-human objects to have consciousness, have intelligence! We have now gone way past the era of human–machine collaboration and heuristics of problem solving of the earlier AIs, we now live in the era of the prediction AIs. Whilst the academic community may be overjoyed with their work on prediction and affective computing to solve societal problems, the same prediction paradigm is being appropriated by high-tech companies and security agencies for automating mass surveillance of people and communities. We learn from Zuboff (2019) that high tech, not content with automation of human experiences into behavioral surplus, has misappropriated affective computing architecture with the aim of automation of human emotion, the creation of an emotion chip, the creation of emotion AI. The implications of automating ‘us’ is to instill an awe of ‘inevitability of technology’ and the culture of ‘economic and market dependency’ and a sense of helplessness in the face of when the computer says “NO”. We wonder whether the creators of the computational paradigm in 1956 would have imagined that one day their dream of functional rationality would be misappropriated by high-tech companies in the 2020s to automate not just problem solving processes but to venture into automating human behaviour in the pursuit of automating the human itself. On reading Zuboff (ibid), we should perhaps not be too surprised that prediction paradigm rooted in the computation paradigm of 1956 would

✉ Karamjit S. Gill
editoraisoc@yahoo.co.uk

¹ University of Brighton, Brighton, UK

continue to follow the path of data mining to build big data commons for behavioural mining, construct living laboratories for reality mining of human experiences, and launch an unprecedented instrumentation for prediction and tuning of societal policies. It is instructive to note that the very prediction paradigm that is elevated to reality mining has now become a catalyst for the creation of uncontrollable echo chambers that thrive by copying ideas, copying feedback, listening to the same sermon, same voice, many gurus but same sermon, no change, no new vision, no transformation—a paradox of algorithmic exploration.

Harari (2020) argues that the epidemic of surveillance technologies that track, monitor and manipulate people, marks an important watershed in the history of surveillance. The danger lies in not just the normalisation of the use/misuse of mass surveillance tools, but also the implication of a dramatic transition from “over the skin” to “under the skin” surveillance that has arrived with coronavirus. For example, what is now demanded of us is not just what is outside our skin but also what is inside—not just the blood pressure and temperature of our fingers, but also the blood pressure under the skin. He asks us to imagine a future scenario in which every citizen would be required to wear ‘a biometric bracelet that monitors body temperature and heart-rate 24 h a day’. We should, remember that technology can also use biometric data to predict human behaviour and manipulate our feelings and sell us anything they want—be it a product or a politician’. We learn from Weizberg (2020) that whilst international agencies such as the UNHC may express concern about sharing sensitive biometric data of refugees with the security agencies, they together with the World Bank and the World Food Programme seek technological solutions to the elusive problem of identity and citizenship status. Under the techno-centric umbrella of improved accountability, increased efficiency, and greater objectivity, biometrics is being used as a blunt instrument for digital surveillance. The surveillance technology may offer ‘seductively easy solutions’ to complex population problems by tying ‘legal status directly to the body’, which also leaves people, their messy lives, life choices, hopes and their survival strategies at the mercy of the digital scanner of ‘dispassionate bureaucracies’. It should be alarming to note that whilst biometrics is increasingly seen as a panacea for a range of problems being addressed by the global development agenda of international agencies such as UN’s Sustainable Development Goals, and World Bank, and the World Food Programme, it has also become a surveillance tool for the externalisation of ‘European asylum policy’ to the ‘Global South’. We also learn from Weizberg (ibid.) that across the ‘Global South, biometric identifiers are increasingly linked to voting, aid distribution, refugee management and financial services’ and the most vulnerable populations are now being used as ‘laboratories for experimental tech.’ It should thus come as

no surprise that human rights advocates worry about international agencies such as UNHCR sharing sensitive biometric data of refugees with the security agencies, thereby making it ‘accessible to other actors beyond the UNHCR’s own biometric identity management system.’ We learn from Zhouboff (op.cit: 215) that it is becoming clear that this ‘surveillance gamification’ is not concerned with ethical, trust and moral implications of misappropriating body as data and behavioural surplus, it is more concerned with taking punitive measures against humans who may breach, even unwittingly, the performance of algorithms. A surveillance algorithm could activate consequences of this breach in the form of ‘a violation algorithm’, ‘a curfew algorithm’ ‘a monitoring algorithm’, ‘an adherence algorithm’, ‘a credit algorithms’. So we have reached a digital future in which reality mining obliterates the past, mortgages the future and speaks in present tense: a future dominated by the arrogance of the prediction paradigm, and in which humans are penalised if they ‘insult the machine’, but the machine is protected from violating not just the human body but also violating their privacy, ethics and moral being. As Cooley (2013), in his poem on INSULTING MACHINES puts it:

Potential and reality are torn apart as change is confused with progress
with slender knowledge of deep subjects
- you proceed with present tense technology,
obliterating the past and with the future already mortgaged
The court of history may find you intoxicated with species arrogance
recklessly proceeding without a Hippocratic Oath.
(Cooley 2013)

Whilst the medical and health professionals and data science researchers see COVID-19 data as a guide to predict scenarios of infection, fatality, and develop guidelines for safety, the same data are being appropriated by surveillance proponents to promote machine learning algorithms and apps as instrumental tools for locating, facial recognition, monitoring and tracing people under the cloak of cloak of public safety, national security, fraud detection, and even disease control and diagnosis. As was noted in Gill (2020) there are, for example, offers of facial recognition systems for predicting the behaviour of citizens, offers of surveillance drones for ‘biometric readings’, predictive Policing is offered as an effective tool to predict, contact trace and reduce crime rates (e.g. Australia’s CovidSafe), and offers of prediction algorithms (e.g. Zegami) to assess the outcome of patient X-rays and diagnose COVID-19 virus. As machine learning and data analytics are offered to ‘accelerate solutions and minimize the impacts of the virus, help expedite the drug development process, and forecast infection rates, these automation tools also raises ethical issues of data protection,

privacy, potential bias in the data, lack of transparency, explainability and accountability. Furthermore, this raises questions of potential negative implications for the therapeutic alliance in patient–clinician relationships.

Gill (ibid) argues that the concern is not just with the automation of behaviour and emotion but also the automation of behavioural interventions and modifications. This automation thereby leads to the exclusion of human engagement to formulate and institute ethical constraints, thereby leads to the exclusion of human interventions from the misappropriation of predictive and affective architectures for high-tech and its market forces for profit. Whilst in the 1980s, we faced the challenge of turning ‘judgment’ to ‘calculation’, now in 2020s we face the challenge of turning human to data. It is no longer about the exclusion of the social but the exclusion of the human itself. At this stage, we wonder whether we would be able to extricate ourselves from the straight jacket of the Fustian Exchange of the prediction paradigm in which we feel helpless and abandoned when the computer says ‘NO’. Whilst prediction technologies of behavioural mining, tuning and surveillance are playing havoc with the identities and lives of people, the high-tech companies are building big data commons and behavioural and experience mining laboratories to first create and then satisfy a culture of instant gratification, ranging from credit cards to fast food, that offers immediate pleasure even in the knowledge that it brings long-term pain. In Carol Ann Duffy (Ramm 2017), we learn of the Faustian seduction: “I grew to love the lifestyle, / not the life”. Goethe’s story of Margareta (Gretchen), provides the most poignant episode of the Faustian exchange. Faust pursues Gretchen, seduces her, and then—unwittingly—destroys her and her family. Mephistopheles guides his hand but Faust’s actions are unbearably his own (the demon goads him: “Who was it who ruined her? I, or you?”). The Gretchen story has become a powerful cultural motif, inspiring elegies such as Byron’s:

Her faults were mine – her virtues were her own –
I loved her, and destroy’d her!...
If I had never lived, that which I love
Had still been living; had I never loved,
That which I love would still be beautiful.

(<https://www.bbc.com/culture/article/20170907-what-the-myth-of-faust-can-teach-us>)

Faust tells Gretchen (Ramm ibid.): “My sweet, believe me, what’s called intellect/Is often shallowness and vanity”, and almost every iteration of the legend underscores this disenchantment: it is Byron’s Manfred who discerns “the fatal truth, /The Tree of Knowledge is not that of Life”. Intellectual pursuits have isolated Faust, and failed to provide him with wisdom: “The very thing one needs one does not know/And what one knows is needless information”. Even

when the quest for knowledge is successful, it conjures up dark forces, as in *Frankenstein*. Like Goethe’s Faust, the proponents and disciples of the prediction paradigm seem to have cast aside their love for scholarship in order to become ‘men of action’, to tame the civilisation and its social and cultural forces of nature, whose tacit and subsidiary contextual dimensions unsettle their deterministic visions, and fill them with anxiety. Their prediction project is beyond human. As Cooley (2013) puts it:

The diagnosis is serious: a rapidly spreading species’
loss of nerve;
Tacit knowledge is demeaned whilst propositional
knowledge is revered.
Who needs imagination when there are facts ?
(Cooley 2013)

From Gill (2020), we learn that the prediction paradigm could neither predict the COVID-19 Tsunami, nor it could provide any relief or diagnosis to people who suffer from COVID-19. Just as the tsunami of the virus cannot be controlled without human engagement and intervention (e.g. medical intervention, social distancing), the virus of the prediction paradigm cannot be controlled without social, ethical and moral constraints and interventions. It is worth repeating the argument that within the academic zones of MIT and Stanford, the prediction paradigm may have been constrained by ethical limits. But once it found its way to Silicon Valley, it was unconstrained by any ethical limits. Further, the COVID-19 pandemic has shown us that just as economy is a very narrow way of organising life and deciding who is important and who is not, so is making the digital future as our home a narrow technological way of thinking about what can be, what should be and what ought to be done for the benefit of society. What we have also learnt from COVID-19 is that the spread of the virus crosses social, cultural, religious, ethnic and geographical boundaries, and thus can neither be controlled by these boundaries, nor can be abstracted away by quantification or wished or washed away through the technological narrative. So any attempt to externalise the spread of the virus to others or outside sources is not only shirking our social and ethical responsibility to mitigate its impact, but also harms others. For Pereira (2019), the problem of the prediction paradigm is not that we have lived with prediction, it is that we are in awe of the power of the machine, and are giving it too much power to automate human behaviour, without social, cultural and legal ethical and moral constraints. This machine learning agency is based on the idea that systems can mine and learn from the huge volume of data, and thereby identify patterns of similarity to make decisions with minimal, if any, human intervention. If this bounded algorithmic agency lacks ethical constraints, then what makes us assured that the prediction paradigm can be tamed by ethical and moral constraints,

when it comes to the automation of human behaviour and emotion?

The prediction paradigm has highlighted our daily Faustian choices. Amore (2020) argues that the injustices of predictive models have been with us for some time. The effects of modelling people's future potential are present in almost all spheres of our lives ranging from predictive policing, visa application, immigration control, children abuse and risk, welfare claims, to student exams, university admission, what we watch, loan applications, and staff recruitment. Our life chances—if we get a visa, whether our welfare claims are flagged as fraudulent, or whether we're designated at risk of reoffending—are becoming tightly bound up with algorithmic outputs. For example, she notes that the predictive algorithm developed by the qualifications regulator Ofqual for predicting student exam results exemplifies the injustice of the prediction paradigm. It not only disregarded the hard work of many young people in a process that ascribed weight to the past performance of schools and colleges, it downgraded the experience of students and teachers. Amore further points out that 'Resistance to algorithms has often focused on issues such as data protection and privacy', but the embedding of predicting algorithms in societal domains such as the above, focuses not just on how the data might be used in the future, but how data are being actively used to change our futures. These discriminatory and opaque predictions not only reduce the potential pathways open to people, but limit their life chances. These algorithms illustrate the technical embodiment of a deeply political idea: that a person is only as good as their circumstances dictate. 'In the future, algorithmic injustices could mean people's choices in education, health, criminal justice and immigration are all diminished by a calculation that pays no attention to our individual personhood.' Amore (ibid.) alerts us that it is time to bring to focus the effects of algorithmic injustice for all to see. The danger is that predictive algorithms offer political and policy-makers the allure of definitive solutions and the promise of reducing intractable decisions to simplified outputs. This logic runs counter to democratic politics, which express the contingency of the world and the deliberative nature of collective decision-making. Algorithmic solutions translate this contingency into clear parameters that can be tweaked and weights that can be adjusted, such that even major errors and inaccuracies can be fine-tuned or adjusted. This algorithmic worldview is one of defending the "robustness", "validity" and "optimisation" of opaque systems and their outputs, closing off spaces for public challenges that are vital to democracy.

The story so far is that the computational model of AI of 1950s, found its rebirth in the prediction paradigm, and consequently we now face challenges of surveillance capitalism. So what went wrong on the way to seeking human-like intelligence? Bell (op.cit.) notes that by separating biology

from cognition, what was lost was the 'social piece, human piece and biological component'. The earlier AI was just a technical artefact, a logical system based on symbols, engaged in an imitation game that created virtual models of the environment, which could then be projected back onto the world itself. Zouboff (2019) says that on the way, we lost the body. The prediction paradigm first appropriated the outer body by mining human behaviour and human experience, and then hollowed the inner being by mining human emotion in search of an emotion chip. Ben Medlock (2017) sheds further light on the limitation of the prediction paradigm and inadequacy of its algorithms to seek human-like AI, without embracing embodiment. He argues that the earlier AI models of symbolic logic such as 'SHRDLU', proved 'hopelessly inadequate when faced with real-world problems, where fine-tuned symbols broke down in the face of ambiguous definitions and myriad shades of interpretation.' Although the recent shift of AI to machine learning has produced many practical applications that, for example, surpass us at speech recognition, image processing, beat us at chess, *Jeopardy!* and Go, and compose pop music, machine learning 'algorithms are a long way from being able to think like us.' We are bodily beings of evolved biology. The human cell, as a biological information processor, is a remarkable piece of networked machinery that has, over centuries, evolved, 'intelligently' adapting and working together to mould us into robust, self-sustaining agents. In asking us to make a 'leap to go from smart, self-organising cells to the brainy sort of intelligence', Medlock (ibid.) quotes Antonio Damasio "we think with our whole body, not just with the brain." He further says that it is the bodily survival in an uncertain world that is the basis of the flexibility and power of human intelligence. This argument suggests that it is questionable whether the prediction paradigm and machine learning approaches will be 'able to capture anything like the richness and diversity of embodied imperative, rooted in the symbiotic relationships of body and the environment. As Cooley (2013) puts it:

A human enhancing symbiosis ignored
whilst a dangerous convergence proceeds apace
as human beings confer life on machines and in so
doing diminish themselves.
Your calculus may be greater than his calculus
but will it pass the Sullenberger Hudson river test ?
Meantime, the virtual is confused with the real
-as parents lavish attention on the virtual child
whilst their real child dies of neglect and starvation.
(Cooley 2013)

The omission of biology in the creation of the computational paradigm, the separation of the body from the brain, the prediction paradigm's hollowing of the body, and ignoring the symbiosis of body and its environment,

has continued to promote and entrench the single story of the universality of the instrumental calculus. Weizenbaum (1976), as early as 1970s, was concerned that instrumental reason is so penetrated in the culture of computation that questions and challenges of human purpose are either ignored or misrepresented, as if every aspect of the real world can be formalised and represented in term of logical calculus. This gave Weizenbaum an insight into a fundamental problem; human beings are liable to attribute to the machine, in this case a diagnostic programme in the field of medical care, more intelligence than it possesses. In doing so, we lose our distance, we fail to realize what the limitations are. Weizenbaum points out that those who aspire to equating machine intelligence to human intelligence keep convincing themselves that by outplaying human Go players, composing music, or creating human-like social robots, machines have either already or are soon going to outsmart human beings. This belief in machine intelligence sees no distinction between the functional machine and the imaginative human being. It seems that in this pursuit of machine intelligence, the validation of human intelligence has been reduced to the display of technological wonders, just as scientific knowledge has been reduced to wonders of data science.

In this age of the fascination with big data and prediction, what is at stake is not just the hollowing and loss of the body but also the loss of wisdom in action. Whilst the data–information–knowledge–wisdom–action loop allowed for human engagement, the algorithmic jump from data to action has not only eliminated practical wisdom, it has also eliminated human from intelligence as if there were no ‘human’ in ‘intelligence’. As Huffington (2018) notes, it is as if humans were simply intelligent machines that could be seamlessly blended with the most intelligent of artificial intelligence with nothing essential lost. What this elimination fails to grasp is that human engagement in action connects self with others, and it is this connectedness that ‘gives meaning to life’. Further that it is this engagement that ‘ultimately determines why technological progress decoupled from wisdom is so dangerous to our humanity.’ Huffington alerts us to the danger of ‘disentangling wisdom from intelligence’ and being drowned ‘in data and starved for wisdom’, and asks us ‘to take steps to protect our humanity from the onslaught of technology in every aspect of our lives as we’re becoming increasingly addicted to our smartphones and all our ubiquitous screens.’ In the pursuit of protecting ‘innately human qualities like wisdom and wonder’, she quotes Harari that up until now, “we humans have built our identity on being *Homo sapiens*, the smartest entities around.” But “as we prepare to be humbled by ever smarter machines,” Harari urges us to “rebrand ourselves as *Homo sentiens*.”

We should, however, be mindful that just ‘branding’ ourselves may be intellectually stimulating, and we need to

take serious note of the seductive control exerted by the big tech machine, such as Facebook, in turning their users into helpless observers of their own and other lived worlds. The machine first disconnects individuals from their social and cultural contexts, creates a society of individuals dependent upon the machine feedback, in the process seducing individuals dispossessed of their social and cultural skills, and ultimately becoming the only social sanctuary without exit. The machine then plays a similar game with society, expressed in this poem written in the spirit of Zhuboff’s argument:

Hail the machine! Thou society of individuals- magical gifts
Of dreams of certainty, hollowed of
Social and cultural roots- a future of
Inevitability, dispossession, helplessness.
Cometh the prediction paradise’s seduction of
Digital behavioural bread crumbs—allure of alignment,
Checkmate thou! Alternative reality-
Game of Surveillance gamification.
Hail the machine! To the digital future’s
Rendition of humanity.

In paying our memorial tribute to Mike Cooley (March 1934–Sept 2020), we remember him as the founding chairman of *AI&Society*, the author of the seminal book, *Architect or Bee*, the recipient of the alternative Nobel Prize, and the architect of the human-centred movement. We should be mindful of his warning (Cooley 2018), when he says that the final act of metamorphosis of the artificial is becoming so complete, so technologically elegant, so powerful of calculation, and so intelligent, that, in some respects, we may no longer be able to tell them apart from humans. In spite of this unsettling prediction, Cooley (Gill 2020), as ever an optimist, asserts that as architects of our own history, we should have the foresight to rewrite the final script, and circumvent the final act of automation, and avoid the slippery slope of calculation to judgment. Cooley urges AI scientists, engineers and practitioners to be architects of the technological future and NOT let the Silicon Valleys discover this future for us, when he says that:

‘the future is not “out there” in the sense that a coastline is out there before somebody goes to discover it’.
It has yet to be built by humans.’

References

- Amoore L (2020) Why ‘Ditch the algorithm’ is the future of political protest. The Guardian, Wed 19 Aug 2020. <https://www.theguardian.com>

- [n.com/commentisfree/2020/aug/19/ditch-the-algorithm-generation-students-a-levels-politics](https://www.youtube.com/watch?v=1KfcWMJ_u0w). Accessed 19 Aug 2020
- Bell G (2018) Decolonising Artificial Intelligence? https://www.youtube.com/watch?v=1KfcWMJ_u0w. Accessed 20 Sep 2020
- Cooley MJ (2013) Ways of knowing. *AI&Society* 28:373–375. <https://doi.org/10.1007/s00146-013-0518-3>
- Cooley MJ (2018) *DELINQUENT GENIUS: the strange affair of man and his technology*. Spokesman, Nottingham (ISBN: 978-0-851-24878-3)
- Gill KS (2020) Prediction paradigm: the human price of instrumentalism. *AI&Society* 35(3):509–517
- Harari YN (2020) The world after coronavirus, *Financial Times*, 20 March 2020. https://www.ft.com/content/19d90308-6858-11ea-a3c9-1fe6fedcca75?fbclid=IwAR1dkuffhX1gIJFL9ERKvyXRaOSkK7qE_9NOIXZJWvQQw3aHyrXIRRn1vU. Accessed 20 June 2020
- Huffington A (2018) We're drowning in data but starved for wisdom: we're more than just intelligent machines. *THRIVE GLOBAL*. March 27, 2018. <https://medium.com/thrive-global/were-drowning-in-data-but-starved-for-wisdom-bd2375baca5>. Accessed 22 Sep 2020
- Medlock B (2017) This post originally appeared on Aeon. March 14, 2017. https://getpocket.com/explore/item/the-body-is-the-missing-link-for-truly-intelligent-machines?utm_source=pocket-newtab. Accessed 22 Sep 2020
- Pereira LM (2019) Should I kill or rather not? *AI&Society* 34:939–943. <https://doi.org/10.1007/s00146-018-0850-8>
- Ramm B (2017) What the myth of faust can teach us. *BBC Culture*. 26th September 2017. <https://www.bbc.com/culture/article/20170907-what-the-myth-of-faust-can-teach-us>. Accessed 21 Sep 2020
- Weizberg K (2020) Machine-readable refugees, *london review of books*—14 September 2020. <https://www.lrb.co.uk/blog/2020/september/machine-readable-refugees>. Accessed 16 Sep 2020
- Weizenbaum J (1976) *Computer power and human reason: from judgment to calculation*. W. H Freeman, Francisco
- Zuboff S (2019) *The age of surveillance capitalism*. Profile Books, London

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.