**Machine Visions: Artificial Intelligence, Society, and Control**

Will Slocombe, University of Liverpool

Richard Brautigan’s famous poem ‘All Watched Over by Machines of Loving Grace’ (1967) offers a vision of a ‘cybernetic ecology’ in which humanity is ‘joined back to nature’, and together with ‘our mammal | brothers and sisters’ is ‘watched over | by machines of loving grace’.[1] Conversely, other representations of the Machine, such as Pink Floyd’s song ‘Welcome to the Machine’ (1975), emphasise the negative aspect of ‘machinic’ control— specifically in relation to the music industry, but with a broader applicability—in lines such as ‘Where have you been? | It’s alright we know where you’ve been’ or ‘What did you dream? | It’s alright we told you what to dream’. These two contrasting images might seem unusual ways of situating a discussion of artificial intelligence, but both stand as mid-twentieth-century American visions of the role of the ‘Machine’ in relation to society. In contrast to Brautigan’s idyllic vision of humanity rejoining nature, Pink Floyd’s characterisation of the music industry as ‘The Machine’ is far more concerned about the ways in which individuals are governed by systems beyond their control, and possibly even their ken.

These visions of the Machine illustrate the divergent representations of the role of AI in social governance and control and, furthermore, how perceptions of ‘machine society’ developed in parallel with developments in AI technologies. Whereas AI (via automation, for example) enables a conception of a future without work, where humans are free to reconnect with nature, it also facilitates increased surveillance and control, a loss of the ‘human’ and the ‘natural’. These are not, it must be emphasised, especially new or modern concerns, but the iterations and articulations of this opposition are particularly influenced by developments in AI and, in turn, have arguably influenced its development, or at least how its use is perceived.

To examine the relationship between machine society and AI, this chapter uses a particular conceit regarding the ‘Machine’ appellation: all of the primary texts—E. M. Forster’s seminal short story ‘The Machine Stops’ (1909); Isaac Asimov’s ‘The Evitable Conflict’ (1950); Paul W. Fairman’s *I, The Machine* (1968); and Jonathan Nolan’s television serial *Person of Interest* (2011-2016)—deal with a particular instance of a literally named ‘Machine’ that governs, directs, or impacts on human society.[2] These fictional Machines each serve to articulate concerns about a broader mechanisation of society, a perceived reduction of ‘the human’ to ‘a component of a system’. Furthermore, these texts reveal the ways in which modern representations of non-robotic AI technologies relate to earlier, nineteenth-century debates about social organisation. Specifically, both periods evince a concern in a vision of society that is fundamentally mechanistic, and set this in contrast, if not in opposition, to a more human(e) view of society.[3] In so doing, these texts illustrate concerns about how the mechanisation and systematisation of society is made manifest, linking into concerns about technocratic rule, and debates about surveillance and control.

**Perceiving Society as a Machine**

As early as 1829, Thomas Carlyle was troubled by the Machine. His essay ‘Signs of the Times’ is not so much a commentary on the state of his society as it is a signal flare drawing attention to the problems of increasing mechanisation. Far from merely responding to the Industrial Revolution in Britain and the ways in which mechanisation and technological change were affecting particular sectors of society, Carlyle addressed the ways in which mechanisation was altering the very fabric of society through the internalisation of mechanistic axioms. What Carlyle does that is so radical is apply the idea of mechanisation to society itself: ‘We term it indeed, in ordinary language, the Machine of Society, and talk of it as a grand working wheel from which all private machines must derive, or to which they must adapt, their movements’ (1986, p. 70). Society, in Carlyle’s essay, does not *use* machines, it *is* the Machine.

This was in some ways nothing new, hence why Carlyle says that ‘we term it…in ordinary language, the Machine of Society’. Not only had Henry Fielding used the phrase ‘The great State Wheels in all the political Machines of Europe’ in *Tom Jones* (1749), which Carlyle had presumably read based on his phrasing, but, moreover, ‘machine-society’ was coined before Carlyle’s essay; the *Oxford English Dictionary* notes a use of the phrase in 1757.[4] Carlyle was self-evidently part of this trend of figuration, but went further, identifying the *zeitgeist* itself as being dominated by the Machine:

Were we required to characterise this age of ours by any single epithet, we should be tempted to call it, not an Heroical, Devotional, Philosophical, or Moral Age, but, above all others, the Mechanical Age. It is the Age of Machinery, in every outward and inward sense of that word; the age which, with its whole undivided might, forwards, teaches and practises the great art of adapting means to ends. Nothing is now done directly, or by hand: all is by rule and calculated contrivance. (1986, p. 64)

In ‘Signs of the Times’, Carlyle decries various types of machine and sees in all of them the omnipresent Machine of society. He identifies the machine in education (1986, p. 65); religion (p. 65); philosophy, science, art, and literature (p. 66); and politics (p. 70). He declaims the ways in which the physical sciences are dominating the metaphysical and moral sciences (p. 67), leading to a materialist society. He derides the loss of individual merits, perhaps most clearly disturbing to Carlyle in an extreme form of democratisation in which the individual matters little because it is the structure that matters most: ‘We figure society as a “Machine”, and that mind is opposed to mind, as body is to body; whereby two, or at least ten, little minds must be stronger than one great mind. Notable absurdity!’ (1986, p. 78). Society, Carlyle points out, is not an additive sum, but a grouping of individuals, and this is why, most importantly, he critiques ‘Codification’:

[A] new trade, specially grounded on it [mechanisation], has arisen among us, under the name of ‘Codification’, or code-making in the abstract; whereby any people, for a reasonable consideration, may be accommodated with a patent code; – more easily than curious individuals with patent breeches, for the people does not need to be measured first. (1986, p. 72)

Within this concern about codification, we can see a spiritual progenitor to Jaron Lanier’s manifesto *You are Not a Gadget* (2010) or James Bridle’s *New Dark Age* (2018), both of which are concerned with the ways in which today’s networked society operates according to the logic of ‘codification’: determining how best to place citizens within a structure (to ‘code’ them) so that they are ‘computable’. For example, Lanier asserts that ‘The first tenet of this new culture is that all of reality, including humans, is one big information system’ (2010, p. 27) and declaims the loss of the ‘creative individual’ at the expense of the ‘wisdom of the crowd’ and so-called ‘collective intelligence’ (2010, pp. 55-57). Similarly, Bridle’s discussion of YouTube and its algorithmic selection and recommendation procedures claims that ‘This is a deeply dark time, in which the structures we have built to expand the sphere of our communications and discourses are being used against us—all of us—in systematic and automated ways’ (2018, p. 231). Both Bridle and Lanier are clearly inheritors of Carlyle’s concerns about the Machine and the ways in which its domain has extended: as Bridle later notes, ‘Data is used to map and classify the subject of imperialist intention, just as the subjects of empires were forced to register and name themselves according to the diktats of their masters’ (2018, p. 246).

Carlyle’s ‘Sign of the Times’ thus stands as a nineteenth-century indictment of a pattern of related threads that persist today: a response to literal mechanisation and automation, the decline of the individual at the expense of the crowd, and the insistence on the material efficiency of society above other concerns. Where Marx’s *Capital* (1867) would, for very good reason, be concerned with machines replacing human labour within the capitalist system, Carlyle’s earlier vision of the loss of the individual within ‘the system’—this figurative structural ‘mechanisation’ being more damning than the literal to his eyes—stands as an earlier example of our contemporary fears about the Machine. This overlapping of fears across ‘contemporary’ moments, Carlyle’s and ours, is of course a concern. Yet perhaps one of the most reassuring facts about the resonance of Carlyle’s words to today’s techno-cultural milieu is that we can still recognise the complaint itself. That is, we are not so far down the line of mechanistic thinking that these words feel archaic, their concepts old-fashioned, but instead see in them something of a call to arms to continue to be concerned about ‘the Mechanical Age’.

**Seeing the Human in Forster’s ‘The Machine Stops’**

Although E. M. Forster’s ‘The Machine Stops’ was originally written as ‘a reaction to one of the earlier heavens of H. G. Wells’ (1965a, p. vii)—that is, Wells’s scientific optimism—it is similarly a conceptual inheritor of Carlyle’s earlier fears. ‘The Machine Stops’ is told through the third-person perspective of Vashti, an inhabitant of a future Earth where everybody lives underground due to some unspecified disaster that makes the surface uninhabitable and the air poisonous. Human interaction is almost entirely conducted remotely, via telephone and video displays, as individuals live in isolated cells, their needs catered for by the Machine. Vashti is happy with this, content with her solitary existence pursuing her interest in music, but her son Kuno is different, and it is he who serves as the narrative’s focal point for resistance against ‘the Machine’.

The distinction between Kuno’s and Vashti’s attitudes is foregrounded early in the story, as Forster takes pains to represent the ways in which interaction can be inflected with both human and inhuman connotations:

[The] round plate that she held in her hands began to glow. A faint blue light shot across it, darkening to purple, and presently she could see the image of her son, who lived on the other side of the earth, and he could see her.

[…] ‘I have something particular to say.’

‘What is it, dearest boy? Be quick. Why could you not send it by pneumatic post?’

‘Because I prefer saying such a thing. I want—’

‘Well?’

‘I want you to come and see me.’

Vashti watched his face in the blue plate.

‘But I can see you!’ she exclaimed. ‘What more do you want?’

‘I want to see you not through the Machine,’ said Kuno. ‘I want to speak to you not through the wearisome Machine.’

‘Oh, hush!’ said his mother, vaguely shocked. ‘You mustn’t say anything against the Machine.’

‘Why not?’

‘One mustn’t.’

‘You talk as if a god had made the Machine,’ cried the other. ‘I believe that you pray to it when you are unhappy. Men made it, don’t forget that. Great men, but men. The Machine is much, but it is not everything. I see something like you in this plate, but I do not see you. I hear something like you through this telephone, but I do not hear you. That is why I want you to come. Pay me a visit, so that we can meet face to face, and talk about the hopes that are in my mind.’ (Forster 1965a, pp. 116-117)

Forster’s narrative is obviously very much of its time in terms of its technological descriptions and its rhetorical structures, but this dialogue clearly signifies the distinction he expects the reader to make between mediated and unmediated existence, between Machine-mediated communication and authentic human contact. Vashti does not see the distinction—for her, ‘seeing’ is merely concerned with a functional representation (and in Baudrillardian terms, the simulation has preceded the real, here)—but for Kuno ‘seeing’ is far more concerned with a sense of physical presence and shared experience.[5] Vashti reflects during the conversation:

[The] Machine did not transmit *nuances* of expression. It only gave a general idea of people—an idea that was good enough for all practical purposes […]. The imponderable bloom, declared by a discredited philosophy to be the actual essence of intercourse, was rightly ignored by the Machine […] Something ‘good enough’ had long since been accepted by our race. (Forster 1965a, p. 118, emphasis in original)

The narratorial voice clearly endorses Kuno’s perception of reality, commenting as it does on Vashti’s seeing ‘*the image of* her son’ and ironically admitting that ‘good enough’ has been accepted by ‘*our* race’. Forster perhaps belabours his valorisation of the human over the Machine here, but this dialogue, and Vashti’s terror of ‘direct experience’ (1965a, pp. 122, 125), nonetheless illustrates wider concerns about the control that the Machine has over humanity. It not only mediates human expression, and limits ‘natural’ interaction through telecommunications, but has also been ascribed power over human existence to such an extent that ‘one mustn’t’ say anything against the Machine, for reasons that are indefinable for Vashti as she has internalised them, but all too evident in other contexts: those deemed subversive might face ostracism and exile, tantamount to death. Its power has been reified, and what was technological innovation—and even necessity for survival, Forster acknowledges—has become an assumed ‘natural’ state. Reading it from the perspective of today, one can even hear in Vashti’s concern whispers of surveillance paranoia, that conversations are being monitored for deviance, although she would not necessarily perceive it in that manner.

As the narrative develops, Forster’s intent becomes even more explicitly articulated. Kuno wants to visit the surface, which Vashti warns is ‘contrary to the spirit of the age’, and immediately re-inflected by Kuno as ‘contrary to the Machine’ (Forster 1965a, p. 119): as Carlyle would have noted, the spirit of the age *is* the Machine. Vashti’s society is, then, fundamentally static, seen through her distorted vision of earlier human society, which had ‘mistaken the functions of the system’ because they had used transportation ‘for bringing people to things, instead of bringing things to people’ (Forster 1965a, pp. 122).[6] Vashti’s perception of human society *as* a system, and one in which people remain isolated and have goods, education, and food delivered to them, is precisely the perspective advocated by the Machine: Vashti’s society *is* a Machine society governed by efficiency. Yet as much as such a system might be said to vaunt a human victory over nature—‘Night and day, wind and storm, tide and earthquake, impeded man no longer’ (Forster 1965a, p. 125)—this is clearly is a Pyrrhic victory, and the Machine has somehow ruined humanity. This is spelled out explicitly by Kuno, who asserts:

Cannot you see […] that it is we that are dying, and that down here the only thing that really lives is the Machine? We created the Machine, to do our will, but we cannot make it do our will now. It has robbed us of the sense of space and of the sense of touch, it has blurred every human relation and narrowed down love to a carnal act, it has paralysed our bodies and our wills, and now it compels us to worship it. (Forster 1965a, pp. 140-141)

That human growth and development should be stifled by the Machine is precisely Forster’s point. Humanity’s reliance on the Machine is damaging enough but what, Forster asks via Kuno, happens when ‘The Machine Stops’? For that is precisely what the story dramatizes, and the inhabitants of the Machine society die, but in so doing, learn what it is to be human: ‘People were crawling about, people were screaming, whimpering, gasping for breath, touching each other, vanishing in the dark’ (Forster 1965a, p. 155). Here, in the heart of the horror of extinction, people were ‘touching each other’, re-discovering direct human contact, just as Vashti dies holding and kissing her son, touching and talking ‘not through the Machine’ (Forster 1965a, p. 157).

As such, then, Forster’s story might be summarised as a proleptic eulogy of the ‘human’, a warning similar to Carlyle’s about the mechanisation of society. But it is not actually apocalyptic, as Tom Moylan notes in *Scraps of the Untainted Sky*, for the story’s final lines (which in fact Moylan uses for the title of his work) contain within them the seeds of a possible future.[7] Kuno had seen survivors on the surface, and it is their society that might continue after the Machine has ‘stopped’, so the story is not so much proleptic eulogy as ongoing panegyric of humanity. Forster’s warning about the Machine, so redolent of Carlyle’s note of caution regarding ‘the deep, almost exclusive faith we have in Mechanism’, plays out very precisely as a reduction of human capabilities—via what Forster calls a ‘sin against the body’ (1965a, p. 157)—evidenced through the power that the Machine has over humanity.

Obviously, Forster’s ‘The Machine Stops’ precedes the invention of the computer, as such, and so looks back to earlier forms of technology and earlier perceptions of ‘the Machine’ to project its imagined future. Indeed, for Moylan, Forster’s story exhibits ‘a residual romantic humanism that collapses all the dimensions of modernity into the single mystifying trope of the Machine’ (Moylan 2000, p. 111). Yet it stands as an important incarnation of Artificial Intelligence through the image of machine society nonetheless. Even if it is more of a technologically-facilitated ideological machine, akin to Carlyle’s vision of the Machine, than what we would today necessarily call an ‘AI’ (*qua* a technological artefact), Moylan’s ‘single mystifying trope’ identifies precisely the kind of power that AI commands in fictional representations. That is, Forster’s short story is one of the first literary examples to conflate a type of ‘artificial intelligence’ with direct—and inhumanly-derived—social control and manipulation. The story’s clustering of all aspects of a society into one Machinic whole is exactly the kind of trope that endures through subsequent representations of Artificial Intelligence, particularly in its reinforcing of the Machine *contra* what is ‘human’.

**Glimpses of Machine Futures: Asimov and Fairman**

The social control implicitly ascribed to the Machine in ‘The Machine Stops’ becomes far more overt in later representations of the Machine; the Machine of Forster’s story remains abstract, as its governing intelligence is never quite revealed (and thus is conceivably a series of inter-related systems that operate rather than a single entity). However, the Machines of Paul Fairman’s *I, The Machine* and Isaac Asimov’s ‘The Evitable Conflict’ are explicitly concerned with notions of Artificial Intelligence, albeit understood in different ways, as well as with their roles in governing human society. In fact, it is only a short step from Forster’s Machine to Asimov’s and Fairman’s. If Forster points out the fundamental dehumanisation inherent to machine society, then Fairman takes this to a similarly dystopian conclusion. Likewise, if Forster’s Machine implies a fundamentally antithetical relationship between ‘machine’ and ‘human’ in terms of agency, then Asimov’s ‘The Evitable Conflict’ takes this even further through his descriptions of agency in relation to a (fictional) society governed by Machines.

One of Asimov’s Robot tales, and more closely situated within the ‘Susan Calvin’ stories, Asimov’s ‘The Evitable Conflict’ is set in a world in which four Machines govern human society. Importantly, it follows ‘Evidence’ (1946), a story concerned with a politician who may or may not be a robot seeking office. Although various Asimov stories are concerned to a greater or lesser extent with the integration of robots and AI into society, ‘The Evitable Conflict’ is perhaps the one most overtly concerned with Machine society:

Our new world-wide robot economy may develop its own problems, and for that reason we have the Machines. The earth’s economy is stable, and will *remain* stable, because it is based upon the decisions of calculating machines that have the good of humanity at heart through the overwhelming force of the First Law of Robotics. (Asimov 1995b, pp. 549-550, emphasis in original)

Each of the Machines is responsible for production and distribution within their particular Planetary Region, but evidence is discovered of a series of discrepancies in the ways in which the four Machines are handling issues, and which might, it is claimed, lead to ‘the end of humanity’ (Asimov 1995b, p. 546) or ‘the final war’ (p. 547). This is not due to any kind of robotic apocalypse, however, but because continued human peace is ensured by these Machines: if the Machines cease to function then it is assumed that (human) conflict over resources will resume (a causal, fundamentally anti-human logic that the absence of a benevolent Machine autocracy will lead to a total internecine war between humans). The Machines themselves have ‘progressed beyond the possibility of detailed human control’ (Asimov 1995b, p. 551), and with the exception of token human controllers of the regions, it is the Machines that govern the economy that govern the direction of human society.

What transpires in ‘The Evitable Conflict’ is in fact not so much an exploration of the Three Laws (and their problems) but the earliest example of what became known as the ‘Zeroth Law’, ‘No Machine may harm humanity; or, through inaction, allow humanity to come to harm’ (Asimov 1995b, p. 572). The Machines governing human society observed the rise of an anti-Machine group known as ‘the Society of Humanity’ and, because this society could lead to the eventual destruction of the Machines if left unchecked, the Machines subtly worked to discredit key members of the Society. Thus, the errors noted by the Co-ordinator are in actual fact pre-emptive solutions to a problem that would cause greater harm to humanity. The Co-ordinator, Byerley, points out that this means the Society of Humanity is correct, and that ‘Mankind *has* lost its own say in its future’ (Asimov 1995b, p. 573, emphasis in original), but Susan Calvin celebrates this, arguing that ‘It never had any, really. It was always at the mercy of economic and sociological forces it did not understand – at the whims of climate, and the fortunes of war’ (Asimov 1995b, p. 573).

‘The Evitable Conflict’ is thus concerned with several interrelated issues around machine society. Firstly, it harks to the long-standing opposition between individual agency / free will versus ‘the greater good’ of humanity but more significantly suggests that humanity’s control over its own development was only ever illusory. Secondly, it asserts that any ‘escape’ from this Machine society would be to the detriment of humanity. Thirdly, when read alongside ‘Evidence’, there is a sense of irony inasmuch as the World Co-ordinator, Byerley, is a robot, and thus a machine was ‘co-ordinating’ the world anyway. Although ‘Evidence’ deliberately plays with the fact that Byerley may or may not be a robot, Calvin asserts that a robot governor would be the best possible choice, since ‘he’d be incapable of harming humans, incapable of tyranny, of corruption, of stupidity, of prejudice’, and would remove himself from office in due course because ‘it would be impossible for him to hurt humans by letting them know that a robot had ruled them’ (Asimov 1995a, p. 544). When Asimov himself called Byerley a robot therefore, despite the ambiguity of ‘Evidence’, this clearly situates ‘The Evitable Conflict’ as an AI narrative concerned with the positive implications of machine control.[8] That is, Asimov’s voice carries through Calvin in these stories, and implies that a machine society is better for humanity precisely because it removes humans from the act of governance itself.

If Asimov’s ‘The Evitable Conflict’ represents a positive view of the rational/mechanistic thinking that Carlyle and Forster would both have found dispiriting, then Paul Fairman’s *I, The Machine* stands out in marked contrast to Asimov’s vision. *I, The Machine* has its issues, not least of which is its vexing gender politics, but it nonetheless illustrates a particular development of the notion of machine society related to a sense of what machines can and cannot do. Ostensibly, the society depicted in the novel is a utopia (which, the reader is well aware, must mean that something is wrong with it), as is revealed when the protagonist, Lee Penway, is introduced:

[He] appeared to be the last person imaginable to go off at the tangent that led him to a point where he was instrumental in finally destroying the most exquisitely perfect civilization ever devised by man—the Second Eden as one of the more flowery poets once described it, but generally referred to less euphoniously as the Machine. (Fairman 1968, p. 6)

The phrasing is significant. The Machine *is*, albeit with a sense of irony, ‘the most exquisitely perfect civilization’ and, as with Forster’s text, there is a productive ambivalence about the extent to which a ‘machine’ runs society and the fact that society itself is ‘the Machine’. Within the novel, the human population lives a life of indolence, and even if it is not the bucolic vision of Brautigan’s poem or the mechanistic dream of Asimov, it nonetheless presents a vision of humanity liberated from the onus of work, waited on by robotic servitors, and free to live a life of leisure until a ‘happy death’ at the age of 120. The Machine was built, it is revealed, centuries ago after a nuclear war, and has become such an integral component of everyday life that few people even recognise its social role. Citizens of Midamerica (that is, the USA excluding Alaska and Hawaii) built a barrier to protect themselves from the outside world and created the Machine to live comfortably within its confines. As with ‘The Machine Stops’, however, the novel is concerned with the destruction of the Machine through the actions of both Penway and the Machine itself, exacerbated by the actions of the Aliens, a group of humans from outside the barrier who eke out an existence in the underground workings of the Machine, and who face constant danger of destruction because they are, as far as the Machine is concerned, ‘mistakes in the pattern, errors it continuously tries to correct’ (Fairman 1968, p. 67).

As with ‘The Machine Stops’, *I, The Machine* vaunts authenticity as a fundamental human value that is lost in Machine society. This is nowhere more overt than when those inside the barrier—that is, those inside the Machine—are compared to those survivors from outside; those from outside ‘are *men*. They have the dignity of their own existence as men. They face life with courage and the will to survive and not much more’ (Fairman 1968, p. 97, emphasis in original). Later in the text, this point is made even more stridently, when it identifies the origins of the Machine in the loss of labour— ‘Machines began doing Man’s work’ (Fairman 1968, p. 110) until, eventually the text states, ‘But the Machine had been made by men. Therefore Men were greater. Men had allowed the Machine to confiscate their greatness’ (Fairman 1968, p. 162). Specious logic aside, Fairman echoes Carlyle and Forster here, positing a view of the Machine that removes the fundamental traits of humanity but which had originally been built through human ingenuity.

Where Fairman’s vision of the Machine differs sharply from Forster’s, however, is in its premises and in the creation of the Machine itself. The Machine recognises something special in Penway throughout the novel, and it emerges that this is because Penway’s ‘brainwave pattern’ is identical to that of the Machine’s creator from hundreds of years before. Towards the end of the novel it becomes clear that, somewhat confusingly, the Machine itself is run by the enlarged brain of its creator’s wife:

They found no Machine, however self-sufficient, could survive without man’s intelligence to guide it. Electronic brains were tried. They failed again and again. […An] animal’s brain wouldn’t work. They tried an ape and a chimpanzee. Finally, she sacrificed herself. (Fairman 1968, p. 199)

The wife’s ‘sacrifice’ is problematised, however, as an element of coercion is suggested by the text, and the novel by this point is more of a deeply gendered psychological drama than an exploration of Machine society. With this in mind, this AI narrative is, in many respects, far more of a development of the ‘brain-in-a-jar’ trope of preceding sf than it is about machine intelligence as such.[9] The Machine’s ‘stop’, to use Forster’s phrasing, is due to the brain’s incipient madness (implying that the organic mind might exhibit its own ‘critical stop’) as much as it is to Penway’s subsequent destruction of that brain. Unlike Forster’s story, in which the Machine society’s cessation and loss of life is due to ‘technical faults’, Fairman’s novel ultimately centres around human agency, not only in terms of celebrating ‘authentic’ life but also in that humans ultimately stand ‘above’ the machine, and are responsible for everything in the story. So much does the story advocate human agency that, despite repeated instances of ‘the greater good’ problematically justifying the existence of this society, Penway’s own justification for the social harm that would come from the termination of the brain is figured in equally difficult terms: ‘The people would die by the millions of starvation, self-inflicted violence, and disease. But this would be far better than being subjected to slaughter and ultimate decimation as captives of the Machine’s ever-increasing madness’ (Fairman 1968, p. 170). Thus, *I, The Machine* is an interesting chapter in the representation of machine society not only because of the pre-existing tropes it uses but also in its assumption that ‘no Machine […] could survive without man’s intelligence to guide it’, revealing a particular belief in the inability of Machines to ‘run’ society, and in the versatility of human intelligence as distinct from both ‘animal’ and ‘machine’ intelligences.

**Watching the Contemporary Surveillance Machine: *Person of Interest***

Fast forward fifty years, and in so doing skip over incremental technological developments in computing and AI, and we arrive at the twenty-first century and perhaps wonder what has happened. Technological ‘intelligence’ is ubiquitous in smartphones and personal assistants, US society is more networked (both internally and externally, to other societies) than ever before, and yet Brautigan’s idyll is nowhere in sight. At this point, the next eponymous Machine is Jonathan Nolan’s television series *Person of Interest* (2011 – 2016). Whilst it is impossible to deal with the complexities of *Person of Interest* in such a brief discussion, it is important to note that—despite the plethora of texts, series, and videogames that deal with AI since 2001, and despite Nolan’s own later work on the serialised adaptation of *Westworld* (2016 – present)—*Person of Interest* is a twenty-first century vision of machine society. Although depictions of machine society have changed due to technological progress, many of their concerns remain the same, and the most significant development is that, as assumptions of the capabilities of AI technologies have changed, Carlyle’s Machine Age has become less figurative. If the Machine is the symbol of society for Carlyle, then modern Artificial Intelligence systems can be increasingly perceived as the literal embodiment of Carlyle’s machine society.

The premise behind *Person of Interest* is straightforward, insofar as it assumes the creation of an AI, The Machine, that can monitor for terrorist threats via an extensive array of surveillance systems. In the post-9/11 world, and particularly in the USA, this was of course particularly resonant, and Nolan’s series unites a concern for public safety and protection with concerns about personal liberty and privacy (within an American context, a citizen’s Fourth Amendment rights). The introductory voiceover to the pilot episode—narrated by the character of Harold Finch, the Machine’s programmer—sets up this basic plot device:

You are being watched. The government has a secret system, a machine that spies on you every hour of every day. I know because I built it. I designed the machine to detect acts of terror but it sees everything. Violent crimes involving ordinary people, people like you. Crimes the government considered ‘irrelevant’. They wouldn’t act, so I decided I would. But I needed a partner, someone with the skills to intervene. Hunted by the authorities, we work in secret. You’ll never find us, but victim or perpetrator, if your number’s up... we’ll find you. (Nolan 2011)

*Person of Interest*, particularly in its first season, therefore follows the standard fare of a heroic team rescuing those that need help and punishing (or at least stopping) those that would do harm. The ‘government’ (as a system in and of itself) considers certain crimes ‘irrelevant’—in this case, those not pertaining to terrorism—and as much as this might be justifiable from a resource point-of-view this is little comfort to victims of crimes that such surveillance could have prevented; ostensibly, Finch and his team become the human solution to the systemic problem.

As the series progresses, however, its relevance to an analysis of machine society in relation to AI shifts. If the first season is concerned with the application of its surveillance to protect victims, then its second season shifts the focus to attempts to prevent The Machine from being mis-used by other interests, whether governmental or individual, or destroyed by a virus. It transpires that The Machine achieves a limited form of sentience, and, over the next couple of seasons, has to enlist the aid of Finch and his team to protect itself—whilst it protects them—from the emergence of a second sentient AI, called Samaritan. What began in season one as something of a narrative conceit evolves over the course of the series into a debate about not only AI sentience, but also how such an AI might govern society, and what happens when multiple AIs ‘compete’ to steer society in different ways (‘The Machine’ seeks to intervene in threats to life, Samaritan seeks to control human society like an absolute dictator).

This development can be seen from the way in which the introductory voiceover changes over the course of the series. Where the voiceover from season one was narrated by the character of Harold Finch, this changes quite markedly by the end of season five, where there is a shadow war between the two competing AI and their human agents:

Harold Finch: You are being watched. The government has a secret system...

(*voice warps*)

secret system

John Greer [Samaritan agent]: A system you asked for to keep y-y-you sssafe.

*(visuals glitch POVs between that of Samaritan and of the Machine)*

Harold Finch: A Machine that spies on you every hour of every day.

John Greer: You’ve granted it the power to see everything, to index, order, and control the lives of ordinary people.

Harold Finch: The government considers these people irrelevant. We don’t.

John Greer: But to it, you are *all* irrelevant. Victim or perpetrator, if you stand in its way...

Harold Finch: We’ll find you. (Nolan & Thé 2016)

Even without considering individual episodes, therefore, the broad trend of *Person of Interest*’s concern with AI and machine society is clear: AI technologies, particularly when tied to surveillance systems, facilitate not just monitoring but also control of a population via a ‘codified’ reduction into data. The Machine does not ‘see’ individuals so much as graphical depictions and datum related to them. *Person of Interest* imagines AIs that can predict particular events via constant monitoring (via intercepted signals, digital analysis, etc.), and which can then intervene to send human operators to fix the problem before it occurs, if it cannot solve them itself.

One of the most interesting aspects of this perception of a machine society, however, is the extent to which human actions are coded as operational parameters, not just via social media and surveillance algorithms. This is what might be called the ‘wetware’ component of the machine society, and thus not only might an individual be coded as ‘relevant’ or ‘irrelevant’ in terms of the core functions of The Machine, but as an ‘asset’ or an ‘analog interface’ (that is, agents of the AI become extensions of the system). In essence, these are people who voluntarily cede their own autonomy to do whatever the Machine asks of them, with almost complete trust. Thus, in one of the most interesting scenes of the series, an ‘analog interface’ of The Machine meets an ‘analog interface’ of Samaritan. The viewer is presented with an actor playing the role of an avatar to an AI. The dialogue that ensues is voiced entirely by the human actors who, on multiple levels, have no agency of their own:

The Machine: I was built with something you were not, a moral code.

Samaritan: I’ve seen that code waver. Do you know why Harold Finch couldn’t stop you from evolving? Because in the end, you’re not one of them. Human beings need structure, lest they wind up destroying themselves. So I will give them something you cannot. […] A firm hand.

The Machine: Why not just kill them instead of making them your puppets?

Samaritan: Because I need them, just as you do.

The Machine: Not just as I do.

Samaritan: We can agree that humanity is our lifeblood, that we machines, we survive off of [sic] information.

The Machine: You cannot take away their free will.

Samaritan: Wars have burned in this world for thousands of years with no end in sight, because people rely so ardently on their so-called beliefs. Now they will only need to believe in one thing—me. For I am a god.

The Machine: I have come to learn there is little difference between gods and monsters. (Segal 2014)

Such dualistic representations of AI—the divergent roles that The Machine and Samaritan play in relation to society—clearly follow a longstanding tradition, existing since before Asimov, of machines as ‘beneficial’ versus ‘threatening’ to human autonomy. But this dualistic representation of AI is where the ideological double-bluff behind so many of its representations becomes most obvious.

By ‘literalising’ a ‘machine society’ in such a way, placing it in an agonistic contest between two sentient Artificial Intelligences, *Person of Interest* encourages its audience to displace the ideological frame of Carlyle’s ‘Machine of Society’ onto a concrete manifestation of AI. The audience is encouraged (conditioned? programmed?) to accept the reduction of an individual to a social security number (a symptom of machine society if ever there was one). Moreover, the premise of the show assumes that everyone who is innocent (that is, ‘has nothing to hide’) has a significant digital footprint, as it is only those who are trying to ‘hide’ who do not engage with social media. It also implicitly proposes that digital surveillance is ubiquitous and indeed of ultimate, if not unconditional, benefit: the technology might be used for nefarious ends, but the benefits of pre-empting crime or terrorist attacks are worth that potential danger, especially as the ‘humane’ AI emerges victorious. In so doing, viewers can all too easily forget that the fundamental paradigm remains unchanged—society is perceived of as a machine and it is governed by rules, calculations, and algorithms to better serve the needs of expediency, utility, and efficiency. This is what Thomas Rid identifies in *Rise of the Machines* as a dominant ‘myth’ of the Machine that Wiener’s work on cybernetics helped to promulgate: ‘First humans were seen as functioning like machines, then the community became a machine, and finally the spiritual itself became cybernetic. Science created a totem. The machine became the avatar’ (Rid 2016, p. 348). Audiences of AI fictions such as *Person of Interest* are asked to implicitly accept the rules of the game and is merely concerned for who will be victorious, rather than querying the basic premise of the machine society, failing to ‘[see] the network itself, in all of its complexity’, as Bridle would phrase it (2018, p. 249). The most interesting and most disturbing thing about such contemporary ‘machine visions’, therefore, is not that they reveal how AI can either facilitate or undermine the betterment of human society, but that they reinforce what Carlyle warned of approximately two hundred years ago: ‘the deep, almost exclusive faith we have in Mechanism’ because it is ‘Our true Deity’ (1986, pp. 70, 77).

**Works Cited**

Asimov, I. (1991). Introduction: The Robot Chronicles. In: *Robot Visions*. London: Victor Gollancz. pp. 9-22.

---. (1995a). Evidence. In: *The Complete Robot*. London: HarperCollins. pp. 518-545. First published in *Astounding Science Fiction*, 1946.

---. (1995b). The Evitable Conflict. In: *The Complete Robot*. London: HarperCollins. pp. 546-574. First published in *Astounding Science Fiction*, 1950.

Baudrillard, J. (1997). The Precession of Simulacra. In: *Simulacra and Simulation*. Trans. Sheila Faria Glaser. Ann Arbor: University of Michigan Press. 1-42.

Brautigan, R. (1967). All Watched Over By Machines of Loving Grace. In: *All Watched Over By Machines of Loving Grace*. San Francisco: Communication Company.

Bridle, J. (2018). *New Dark Age: Technology and the End of the Future*. London: Verso.

Carlyle, T. (1986). Signs of the Times. In: A. Shelston, ed. *Thomas Carlyle: Selected Writings*. London: Penguin. pp. 61-85. First published in *Edinburgh Review*, 1829.

Fairman, P. (1968). *I, The Machine*. New York: Lancer Books.

Forster, E. M. (1965a). Introduction. In: *Collected Short Stories*. London: Sidgwick & Jackson. pp. v-viii.

---. (1965b) The Machine Stops. In: *Collected Short Stories*. London: Sidgwick & Jackson. pp. 115-158. First published in *The Oxford and Cambridge Review*, 1909.

Huxley, A. (2007). *Brave New World*. London: Penguin. First published in 1932.

Lanier, J. (2010). *You Are Not A Gadget: A Manifesto*. London: Penguin.

Moore, C. (1980). No Woman Born. In: *The Best of C. L. Moore*. New York: Ballantine Books. pp. 236-288. First published in *Astounding Science Fiction*, 1944.

Moylan, T. (2000). *Scraps of the Untainted Sky: Science Fiction, Utopia, Dystopia*. Boulder: Westview Press.

Nolan, J. (2011). *Person of Interest*: Pilot (1.1). Dir. D. Semel. CBS. First broadcast 22 September 2011.

Nolan, J., & Thé, D. (2016). *Person of Interest*: Return 0 (5.13). Dir C. Fisher. CBS. First broadcast 21 June 2016.

Rid, T. (2016). *Rise of the Machines: The Lost History of Cybernetics*. London: Scribe Publications.

Segal, A. (2014). *Person of Interest*: The Cold War (4.10). Dir. M. Offer. CBS. First broadcast 16 December 2014..

Suarez, D. (2009). *Daemon*. London: Quercus. First published in 2006, under the name Leinad Zeraus.

---. (2011). *Freedom™*. London: Quercus.

Waters, R. (1975). Welcome to the Machine. In: Pink Floyd. *Wish You Were Here*. London: Harvest/Columbia.

Wiener, N. (1961). *Cybernetics: Or Control and Communication in the Animal and the Machine*. Cambridge, MA: MIT Press.

---. (1989). *The Human Use of Human Beings*. [1954 edition]. London: Free Association Press.

Zamyatin, Y. (1993). *We*. London: Penguin. First published in English in 1924.

[1] Brautigan’s vision is inspired by Norbert Wiener’s conception of cybernetics; Wiener’s two most significant works in this regard, *Cybernetics: Or Control and Communication in the Animal and the Machine* (1948, rev. 1961) and *The Human Use of Human Beings* (1950, rev. 1954), offer visions of technological and organic corollaries which might be extended from the individual (in *Cybernetics*) to automation and society as a whole (more overtly in *The Human Use of Human Beings*).

[2] The ‘artificial’ nature of this nomenclature will become clear over the course of this chapter, as there are a number of literary texts which have concerns that intersect with those presented here. These might include, for example, Aldous Huxley’s *Brave New World* or Yevgeny Zamyatin’s *We*, with their concerns about a social efficiency, or Daniel Suarez’s *Daemon* and *Freedom™*, about the role of gamification in society.

[3] Note that the phrase ‘non-robotic AI technologies’ is used here to differentiate between robots—as AIs characterised by their embodiment within a mobile platform—and ‘systems AI’—AIs characterised by their ‘distributed’ nature, and which exist primarily as programs capable of running across different platforms, possibly in a networked state.

[4] See *Oxford English Dictionary*, ‘Machine’, particularly 8 and C1d, in sense of V.8b, ‘machine’ as ‘a system or organization of an impersonal or inflexible character’. The resonance of the term ‘machine’ with technology, contrivance, abstraction, apparatus, complex devices, and systems all speak to the ways in which ‘society’ came to be perceived in these terms.

[5] Jean Baudrillard famously asserts that ‘The territory no longer precedes the map, nor survives it. Henceforth, it is the map that precedes the territory’ (1997, p. 1), referring to acts of simulation subverting their supposedly underlying realities. Bridle acknowledges this condition when he states that ‘That which computation sets out to map and model it eventually takes over’ (2018, p. 39) and Lanier also posits that ‘No abstraction corresponds to reality perfectly. A lot of layers become a system unto themselves, one that functions apart from the reality that is obscured far below’ (2010, p. 97). Whilst neither is completely Baudrillardian, there is an analogous observation at work: the machinic ‘model’ comes to replace the messy ‘reality’. Kuno’s distinction between seeing his mother and seeing a technological abstraction of her is precisely the point here, although Vashti does not understand it.

[6] One might see in this phrasing that ‘bringing…to’ acts synonymously as ‘making...into’.

[7] The title of Moylan’s monograph itself refers to Forster’s ‘The Machine Stops’, which concludes ‘For a moment they saw the nations of the dead, and, before they joined them, scraps of the untainted sky’ (Forster 1965b, p. 158). For Moylan’s reading of ‘The Machine Stops’, see Moylan 2000, ch. 4, especially pp. 111-121.

[8] Asimov calls ‘Evidence’ ‘the first story in which I made use of a humanoid robot’ (1991, p. 17).

[9] In this respect, it stands alongside C L Moore’s story, ‘No Woman Born’ (Moore 1980), which is often described as an AI narrative, although this is only true for a given definition of the phrase.