ARTIFICIAL INTELLIGENCE, TECH CORPORATE GOVERNANCE AND THE PUBLIC INTEREST REGULATORY RESPONSE.

(*Cambridge Journal of Regions Economy and Society)*

Alan Dignam

Abstract

This article attempts to get to the heart of some of the general misunderstanding and misapplication of Artificial Intelligence (AI) decision-making technology and proposes a regulatory model to place public rather than private interest at the heart of AI regulation. The article proceeds as follows. Section 1 introduces matters. Section 2 proposes a contextualized rather than a deterministic technology lens is needed to cut through the confusion and misunderstanding surrounding AI. Section 3 examines the nature of AI decision-making focusing on its human design and impact, and concludes that technology is not the root of problematic outcomes in the area but rather flawed human design and implementation. Section 4 examines the ownership and control of the major AI developers and concludes that a small group of humans with autocratic tendencies dominate AI development. Section 5 concludes that despite the unfortunate deregulatory instincts of the US and UK governments with regard to technology, AI should be treated in a similar manner to pharmaceutical products by introducing public interest regulation through the medium of a state regulatory body. Similarly a public interest technology regulator would address the autocratic ownership tendencies of the technology sector.

Professor of Corporate Law and Honorary Member 7 King's Bench Walk Chambers,
School of Law,
Queen Mary, University of London
67-69 Lincoln's Inn Fields
London WC2A 3JB
Email: a.dignam@qmul.ac.uk

With thanks to Sarah Morley, David Lawrence and the Welcome Trust for establishing a remarkable network to look at AI within the corporate context. My thanks also to James Maclaurin, Colin Gavaghan, Alistair Knott, John Zerilli, Joy Liddicoat and the New Zealand Law Foundation Centre for Law and Policy in Emerging Technologies as well as the other participants at the Otago/Oxford Roundtable on Artificial Intelligence and Employment in November 2018 for helping to develop the focus of this paper.

"*The future is not google-able*"

William Gibson

INTRODUCTION

Within the past decade Artificial Intelligence (AI) has passed from the realm of Science Fiction to apparent operational and legal reality, with even the European Parliament suggesting a form of legal personhood for AI.[1] Within the academy AI conferences and networks abound as the potential for AI to deeply change our societies has become apparent. However, what is also apparent is that a complex science fiction veneer remains present in that many speakers and writers discuss AI impacts that exist only in novels as if they were real. Phrases like "machine learning", "artificial neural networks" and "training the AI" are used casually and confusingly to imply sentience and superiority, when they merely refer to statistical models using mass computational power to make often deeply flawed decisions/predictions. Crucially, in general there is a wide spread misunderstanding of the limits of using AI statistical modeling to make decisions and its deeper societal and particularly legal implications. In general, AI is misleadingly portrayed and, worryingly, some academic work in the area is supported by the tech industry without many academics seeing, or declaring, a conflict of interest.[2] Even when problems are recognized, the background radiation is that tech and AI are exceptional world changing positive forces determining

[1] European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)

[2] This work is not funded by the tech industry. A. Orlowski, 'Academics 'funded by Google' tend not to mention it in their work.' Two-thirds of the time support is not disclosed, say campaigners, *The Register*, 13 Jul 2017 <https://www.theregister.co.uk/2017/07/13/google_transparency_report_academics/>

our future for good. It is all somewhat reminiscent of the treatment of the tobacco industry in the 1950s.[3]

This article attempts to get to the heart of some of the general misunderstanding and misapplication of AI decision-making technology and proposes a regulatory model to place public rather than private interest at the heart of AI regulation. As such, the article proceeds as follows. The first section, here, introduces matters. The second proposes that a contextualized human rather than a deterministic technology lens is needed to cut through much of the confusion and misunderstanding surrounding AI. The third examines the nature of AI decision-making focusing on its human design and impact, and concludes that technology is not the root of problematic outcomes in the area but rather flawed human design and implementation. The fourth examines the ownership and control of the major AI developers and concludes that a small group of humans with autocratic tendencies dominate AI development. The fourth section concludes that despite the unfortunate deregulatory instincts of the US and UK governments with regard to technology, AI should be treated in a similar manner to pharmaceutical products by introducing public interest regulation through the medium of a state regulatory body. Similarly a public interest technology regulator would address the autocratic ownership tendencies of the technology sector.

AI IN CONTEXT: SCIENCE FICTION V STATISTICAL MODELS

In 2017, the European Parliament passed a resolution suggesting a form of legal personhood for Artificial Intelligence.[4] In legal terms, this was an extraordinary idea given that legal personality is not lightly conferred in any jurisdiction. Legal personality forms the gateway to a legal system and the rights and obligations that confers. Humans and important organizational forms such as companies and unions have historically claimed legal personality and, exceptionally, deities, rivers

[3] S. Elliott. 'When Doctors, and Even Santa, Endorsed Tobacco'
New York Times, 6 October, 2008.
<https://www.nytimes.com/2008/10/07/business/media/07adco.html>
[4] European Parliament, op. cit. n 1.

and certain higher level mammals have been granted legal personality, but not machines or computer programmes.[5] That the European Parliament passed the resolution was even odder given that its own report in 2016 evaluating various problematic issues with AI and robotics concluded:

> In reality, advocates of the legal personality option have a fanciful vision of the robot, inspired by science-fiction novels and cinema. They view the robot — particularly if it is classified as smart and is humanoid — as a genuine thinking artificial creation, humanity's alter ego. We believe it would be inappropriate and out-of-place not only to recognise the existence of an electronic person but to even create any such legal personality. Doing so risks not only assigning rights and obligations to what is just a tool, but also tearing down the boundaries between man and machine, blurring the lines between the living and the inert, the human and the inhuman.[6]

Ultimately the European Commission preferred that 2016 view and did not include legal personality in its 2018 legislative plans.[7]

There is much to draw from the conclusions of the 2016 report in terms of the central danger of AI. Despite the excited sci-fi inspired claims, AI is not human. It is a tool that can be used well or badly. In the end it is human decision-making as to its design and deployment that matters. Although AI is the subject of this article, the lens through which it is examined is human rather than technological agency. It is not always easy to focus on this, as human agency is easily disregarded in the high tech world, where human workers are often deliberately hidden behind technology platforms such as those used by Deliveroo,[8] Google/Facebook,[9] Amazon Mechanical Turk,[10] or IBM's AI Watson for Oncology.[11] Recognizing that technology is not a preordained deterministic

[5] E. O'Donnell, and J. Talbot-Jones. 'Creating legal rights for rivers: lessons from Australia, New Zealand, and India.' (2018) *Ecology and Society* 23(1):7.

[6] European Parliament's Committee on Legal Affairs study on European Civil Law Rules in Robotics, 2016 (PE 571.379) page 15-16.

[7] European Commission
'Artificial intelligence: Commission outlines a European approach to boost investment and set ethical guidelines' Brussels, 2018 <http://europa.eu/rapid/press-release_IP-18-3362_en.htm>

[8] In this advertisement for Deliveroo the human delivery workers are literally invisible <https://www.youtube.com/watch?v=2f4u6HbH5xY>

[9] A. Singh, 'Facebook moderators 'develop PTSD'
The Telegraph, 31 May 2017 <https://www.telegraph.co.uk/news/2017/05/31/facebookmoderators-develop-ptsd-exposed-worst-content-internet/>

[10] https://www.mturk.com/

[11] See discussion below.

force but a social construction produced by human decision making and comprehending who those human are, is key to maximizing the public benefit while minimizing the dangers of AI.[12]

One of the first hurdles to asserting the human focus in the world of AI is the use of language that obscures the reality of AI. AI has two general industry/academic divisions – general or strong AI, meaning sentient human conscious type AI, and narrow or weak AI, meaning basic narrow/weak processes using statistical analysis of mass data sets and computational power. General AI is the ultimate science fiction inspired goal but does not exist.[13] Weak/narrow AI is the reality. Some of the activity that narrow AI carries out mimics aspects of human intelligence such as speech or facial recognition. Some of its computational ability goes beyond human ability and the more complex statistical AI models can deal with some level of subtlety and nuance but is not of a human order of intelligence.[14] A telescope is a complex precision tool designed to enhance human observational ability. It is not intelligent, but it has been intelligently designed by a human. A mechanical watch is a complex instrument that measures time but it does not consciously "decide" what time it is. Harrison's H4 Marine Chronometer was the Sat-Nav of its day and was an extraordinary piece of human life-enhancing technology but no-one then or since has described the H4 as "intelligent" despite its extraordinary complexity.[15] It remains a tool for human use designed by the immense intelligence of John Harrison. Neither the H4 or the telescope are intelligent in the human sense, just as a statistical model described and sold as AI is not intelligent.[16] They are at best precision tools designed by humans for other humans to use. At worst, as will be explored, they can be error-strewn drivers of bias, inequality, death and loss of liberty.

So, what's the problem with using a technically inaccurate phrase to describe this type of statistical computational automation? Misselling statistical tools as AI creates an illusion of and

[12] T. Pinch, and W. Bijker. 'The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other,' (1984) *Social Studies of Science* 14 : 399-441.
[13] R. Yampolskiy, and J. Fox, 'Artificial Intelligence and the Human Mental Model.' in: *The Singularity Hypothesis: a Scientific and Philosophical Assessment.*' A. Eden, J. Moor, J. Soraker, J. and E. Steinhart, (eds.) (2012) 129-145, Springer
[14] G. Ryle, *The Concept of Mind,* (2009) Routledge, S. Russell, P. Norvig, *(2003), Artificial Intelligence: A Modern Approach, (2003)* Prentice Hall, M. Ford, *Architects of Intelligence,* (2018) Packt, and B. Goertzel and C. Pennachin, *Artificial General Intelligence,* (2007) Springer.
[15] D. Sobel, *Longitude,* (1995) Penguin.
[16] M. Broussard, *Artificial Unintelligence,* (2018) MIT Press, 1-39.

reaction to assumed superior intelligence, when in fact it is nothing of the order of human intelligence. Chess computers, AlphaGo and IBM Watson's ability to beat humans within a narrow gaming skill set reinforces the illusion that computers are superior at decision making than humans. That same assumption would not be made about a telescope, even though, applying the tech AI personification logic, a telescope is better at looking at stars than a human is. In reality the telescope simply assists us make better observations and decisions.

The science fiction base for general AI, as alluded to in the 2016 report above, also fuels a dangerous anthropomorphic/personification loop for humans both for those who attempt to build AI statistical models and those end users who assume enhanced intelligence in AI decision making. It seems odd to have to articulate it, but science fiction is not real. The reality is that 1999[17] passed without a semi autonomous main computer supporting a human colony on the moon, 2000AD has come and gone without us discovering Verdus the robot planet,[18] 2001 passed and no HAL 9000[19] emerged, while Blade Runner was set in 2019[20] and there are no androids so far this year. Similarly, although the Marvel Universe is confusingly multi-dimensional, Vision the android is currently not flying around this version of Earth.[21] 2029 has not yet come, so Skynet[22] cannot be ruled out but so far in 2019 robots, like the Daleks, are not great at navigating stairs.

There is no doubt that science fiction can sometimes foreshadow and inspire real world technology by imagining what might be possible but it can also mislead in the AI context. Our best AI decision making programmes are at heart statistical models with the usual human influenced problems of poor design, data and data interpretation that statistical analysis involves.[23] The phrase "Lies, damned lies and statistics" variously attributed to Twain, Disraeli or Balfour has continual historical resonance because it recognizes that human judgment is at the heart of statistical integrity and not pure numerical truth. AI has been enormously oversold in terms of its capabilities, while

[17] Space: 1999 was a popular UK science fiction TV series in the 1970s <https://en.wikipedia.org/wiki/Space:_1999#Other_media>
[18] <https://en.wikipedia.org/wiki/Robo-Hunter>
[19] A. Clarke, *2001: A Space Odyssey* (1968), New American Library.
[20] Blade Runner June 25th 1982 <https://www.warnerbros.com/blade-runner>
[21] https://en.wikipedia.org/wiki/Vision_(Marvel_Comics)
[22] <https://en.wikipedia.org/wiki/The_Terminator>
[23] Broussard, op. cit. n. 16. Chapter 7.

companies also regularly misrepresent products as AI. In 2019, for example, two fifths of AI start up companies in Europe were found to use no recognizable AI technology in their products.[24] That is not to say limited weak machine intelligence is not present or that an AI product doesn't mimic or copy human intelligent behaviour – Apple's Siri and Amazon's Alexa are designed to do just that. As Lipton, one of the worlds leading AI scientists, has described cutting edge AI:

> these are just statistical models, the same as those that Google uses to play board games or that your phone uses to make predictions about what word you're saying in order to transcribe your messages. They are no more sentient than a bowl of noodles, or your shoes.[25]

It is essential to understand this in designing regulatory safeguards for AI use because the humans designing, using or subject to the AI can similarly misunderstand the nature of the intelligence being observed and suffer as a consequence.

For marketing purposes the tech industry often merges general and weak AI relying on a preexisting appetite for science fiction in film, TV and literature to excite the public. That use of si-fi has an interesting duality: utopian visions of a leisure class served by intelligent robots simultaneously creates fear of dystopia, for most si-fi is dystopian, where the humans are enslaved.[26] The utopian vision creates a futuristic positive desire for making the world better,[27] while the dystopian vision reinforces an existential threat, which in the tech industry narrative, technology and technology companies play a part in fighting. Apple's famous Orwell inspired 1984 Superbowl advert for example, has a heroic woman representing Apple fighting off a technological enabled Big Brother figure.[28] In this tech centered world view we must trust the good heroic tech

---

[24] A. Ram, 'Europe's AI start-ups often do not use AI, study finds,' *Financial Times*, March 5 2019, <https://www.ft.com/content/21b19010-3e9f-11e9-b896-fe36ec32aece>

[25] K. Quach, "Facebook pulls plug on language-inventing chatbots?" *The Register* 1 Aug 2017, <https://www.theregister.co.uk/2017/08/01/facebook_chatbots_did_not_invent_new_language/ >

[26] A. Jezard, 'Technophobia is so last Century,' *Financial Times*, March 2 2016, <https://www.ft.com/content/a9ec6360-cf80-11e5-92a1-c5e23ef99c77>

[27] F. Turner, '*From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism'* (2008).

[28] Apple's 1984 Superbowl Advert can be viewed here <https://youtu.be/2zfqw8nhUwA>

products to protect us from the bad tech future. More recently AI's ubiquity in general public debate allows additional mainstream association with film and media. For example Google Assistant advertising reworked the Home Alone film series in 2018 around a utopian motif where a grown up Kevin McCallister is never threatened by intruders, because Google Assistant's control of the home keeps him safe.[29] In February 2019, Microsoft AI sponsored The Hollywood Reporter's entire coverage of the Oscars.[30] Indeed, often within the industry and Academy, sci-fi forms the inspiration to try to create general AI and indeed the link goes both ways. Marvin Minsky, one of the key scientific figures in modern AI, both devoured and wrote science fiction, while also being a friend of Isaac Asimov and Arthur C. Clarke. He worked with Clarke and Kubrick on creating the vision of the menacing AI computer HAL 9000 that appears in 2001: A Space Odyssey.[31]

The sci-fi element also has the effect of pushing back at law and regulation by the Orwellian Big Brother state because this would interfere with a utopian tech-immersive future. That sci-fi inspired tech future has been accompanied by a strong belief that new technology and its development is free of law and regulation by its very nature: that new technological frontiers are governed by mathematical calculation and old outdated preexisting laws cannot stand in the way. Technological determinism runs strongly through these claims.[32] In 1996, the publication of "A Declaration of the Independence of Cyberspace" by John Perry Barlow captured what was to become the driving utopian determinist libertarian mindset of nascent internet tech companies – the root of the famous Facebook coder "Move fast and break things" tech philosophy.[33] It begins:

> Governments of the Industrial World, you weary giants of flesh and steel, I come from Cyberspace, the new home of Mind. On behalf of the future, I ask you of the past to leave us alone. You are not welcome among us. You have no sovereignty where we gather.
>
> We have no elected government, nor are we likely to have one, so I address you with no greater authority than that with which liberty itself always speaks. I declare the global social space we are building to be naturally independent of the tyrannies you seek to impose on us.

[29] Google's Advert can be viewed here <https://youtu.be/xKYABI-dGEA>
[30] The Hollywood Reporter coverage can be viewed here <https://youtu.be/RFUFwJMweCo>
[31] Broussard, op. cit. n. 16. pp.71-72.
[32] B. Bimber, 'Karl Marx and the Three Faces of Technological Determinism,' (1990) 20 (2) *Social Studies of Science.* 333–351.
[33] J. Taplin, *Move Fast and Break Things: How Facebook, Google, and Amazon Cornered Culture and Undermined Democracy,* (2017).

> You have no moral right to rule us nor do you possess any methods of enforcement we have true reason to fear.[34]

In this tech utopia Uber is not a taxi company, Deliveroo is not a delivery company, Google and Facebook are not publishers/broadcasters/software or music companies and Tesla is not a car company. They are the utopian future and to restrict their activities through existing laws is to strike at freedom itself. As Vance notes, commenting on tech leaders' futuristic world view, "They were all geeks raised on science fiction and the vision of space we had in the 1960s and 70s. Now they have the money to make this a reality."[35] US tech determinist libertarianism also has a particular business focused flavor. Within the panoply of US libertarian political philosophy this comes closest to the autarchist libertarian school. As its inceptor Robert LeFevre described that distinction in 1965:

> If one believe in freedom, one must believe in economic freedom--full latitude of choice in any and all economic areas, for each person. This can never be accomplished by any procedure, organized or otherwise, which uses violence (even the violence implicit in taxation) to take from an owner anything which is rightfully his.[36]

As such, the drive by tech billionaires such as Tesla's Elon Musk and Amazon's Jeff Benzos to explore space is not just seeking adventure. The ability to operate off-planet carries the autarchic libertarian promise of companies being able to operate in a genuine law free zone.[37]

This background radiation of determinist libertarian push back has an effect as governments, particularly the UK and US governments, are facilitating a deregulated libertarian tech-exceptional view, and industry self-regulation, as will be observed later, is accepted. In the AI context, pushing general AI as if it was real is extremely important to creating these regulatory pushbacks. Its most pervasive and dangerous application is in the narrow/weak AI terminology where "neural

[34] J. Barlow, *A Declaration of the Independence of Cyberspace*, February 8 1996, <https://www.eff.org/cyberspace-independence>
[35] D. Tynan, 'Rocket men: why tech's biggest billionaires want their place in space,' *The Guardian*, December 5 2016, <https://www.theguardian.com/science/2016/dec/05/tech-billionaires-space-exploration-musk-bezos-branson>.
[36] R. LeFevre, 'Autarchy vs Anarchy,' *Rampart Journal of Individualist Thought*, (1965) Vol. 1, No. 4: 30–49, 49.
[37] C. Graham. 'Factories in space: Amazon founder Jeff Bezos unveils vision for the future,' *The Telegraph*, June 2 2016, <https://www.telegraph.co.uk/technology/2016/06/01/factories-in-space-amazon-founder-jeff-bezos-unveils-vision-for/>

networks", "machine learning", and the phrase "artificial intelligence" itself, portray a general AI world of superior intelligence instead of the reality of extremely limited, mostly standard statistical models that are only novel in that they utilize the vast computational power increasingly available to us.

Why does this matter? It matters for the regulatory push back as mentioned, but it matters also because to describe, say, Amazon's now aborted AI hiring project as AI, or California's criminal justice application COMPAS or Watson for Oncology as AI is to mislead those using the AI and those subject to its decision-making. The assumption is it is superior but as will be explored it is not. It can facilitate, as will be observed, unlawful hiring practices, racial and sexual discrimination, unnecessary financial hardship and, in the case of self driving cars or Watson for Oncology, place lives at risk. As Waters notes "Strip away the gee-whizz research that hogs many of the headlines (a computer that can beat humans at Go!) and the technology is at a rudimentary stage."[38]

Hiding that AI is simply comprised of statistical models can also hide that it's not new or techy and so not logically part of our utopian deregulated tech safe harbours. Statistical models were first used in the 17th century and have a long and problematic history when used badly or when their predictive power is misunderstood.[39] Unfortunately, the tech industry seems to have hidden its statistical models so well behind the AI façade that they have produced narrow/weak AI statistical and probability decision models that ignore three centuries of experience working with these models and their dangers. Poorly designed statistical and predictive models will fail; bias and dirty data is a big problem; and lawful human design, interpretation and continual audit of the results is essential. Additionally, software and hardware can be buggy and unreliable and the tech

[38] R. Waters, 'Everything still to play for with AI in its infancy,' *Financial Times,* Feb 14 2019 <https://www.ft.com/content/bf3d708c-3077-11e9-8744-e7016697f225>

[39] W. Willcox, 'The Founder of Statistics' *Review of the International Statistical Institute* (1938) 5(4): 321–328.

industry has a long and painful record of failing to evaluate the risks of its products and being delusional as to its capabilities.[40] As Charette considered:

> software failures tend to resemble the worst conceivable airplane crash, where the pilot was inexperienced but exceedingly rash, flew into an ice storm in an untested aircraft, and worked for an airline that gave lip service to safety while cutting back on training and maintenance.[41]

Doing it properly is a highly-skilled endeavor that is both expensive and time consuming. Mass computational power, while it allows greater statistical scale does not move the needle at all with regard to basic statistical integrity. Highly skilled human design, operation and oversight is still essential even if you've got access to a 93 petaflop computer. Facebook's historical "Move fast and Break Things" philosophy is dangerous for users and those subject to its AI outcomes. Eventually, though, there are significant consequences for the company itself and its clients, when what they break is the law.

However, the statistical models currently called AI, used properly, recognizing their dangers, can be a useful tool for societal progress. That is after all AI's attraction. Freeing humans from dangerous or labour-intensive tasks, improving health and enhancing human ability to analyze the world is a positive goal. A tool like a telescope allows humans to enhance their observational ability and ultimately send a probe outside the solar system, and a clock more accurately allows humans to gauge time, which in turn allows us to run a rail network. Well designed narrow AI can analyse huge data sets a human simply could not do and can produce counterintuitive data analysis outcomes that would be impossible for a human to discover on their own. Used badly, as will be observed, AI has the potential to send innocent people to jail, discriminate against women and minorities, unfairly exclude people from the financial system and public services, injure and kill road users and misdiagnose patients. Maximising the benefit while minimizing the dangers is the challenge.

---

[40] R. Charette, 'Why software fails,' *IEEE Spectrum*, (2005) vol. 42, no. 9, pp. 42-49, Sept.

[41] Charette, id., p.7.

FINDING THE AI DECISION MAKERS.

1.

AI products abound, ranging from domestic appliances, such as smart dishwashers, to smart medical devices, AI phones and computers, smart speakers, AI human resources and AI quasi-judicial decision making models. All are powered by versions of statistical/predictive models within the weak/narrow AI category. This section considers aspects of the design and operation of AI decision-making models focusing particularly on bias, the "black box" proposition, designed novel AI outcomes, complexity/superiority atrophy and low cost AI versus the humans. As before, the focus is on contextualizing the technology by drawing out the humans behind the AI and who are impacted by the AI.

As noted above, weak/narrow AI is not a new technological marvel but rather old mathematical statistical/probability models that have been calibrated and tested against known correct outcomes to make decisions based on those previous outcomes. The process for a basic high quality statistical machine-learning model should work as follows. First, a budget appropriate to the project needs to be set and reviewed as the project progresses. Getting this wrong can strongly undermine the quality of the AI outcomes. Second, detailed decisions need to be made by the designer to work out what the model will do. This first phase is key and should involve humans who have deep and detailed knowledge of the area that will be subject to the AI and the policy objectives to be achieved. Third, the designer decides what data will be used. This should involve an evaluation of the extent and quality of the data available. If the quality is too low then that should end the project until better data is available. High quality data is rarely available. In most cases, the data is not entirely complete or a bit problematic and some version of a median fill function is often used to plug gaps and complete the data set. This is normal but can affect the outcomes, as some of the data is by design incorrect. Fourth, standard statistical machine learning software would then be run, or in more complex situations involving image or language, something like an artificial neural

network.[42] When the process of "training" or calibrating the AI against known decision outcomes reaches at least 90% accuracy, a working AI model exists. To do this properly on a commercial basis requires some very good data scientists and there are not many of those.[43] Humans from a range of backgrounds also need to critically evaluate the outcomes and those outcomes should be lawful and explainable. Where elements of bespoke coding have been incorporated then the programme will likely need to be debugged, again… and again…[44] When the AI goes live it will also need to be audited regularly to recalibrate it as it deals with new situations and to see if its is still performing along the policy lines set by the organization and the law.

The steps described above do not often happen in commercial tech AI development as it is extremely expensive to produce a high quality AI model. In the US state of Idaho for example in 2012 despite widespread knowledge of high level corrupt data and enormously problematic outcomes in testing, a statistical decision making algorithm was put into operation by public officials resulting in Medicaid cuts to 4000 disabled people. The widespread hardship this caused resulted in the American Civil Liberties Union (ACLU) bringing a successful court case to reinstate the payments.[45] As the ACLU noted afterwards:

> the unfortunate part, as we learned in this case, is that it costs a lot of money to actually test these things and make sure they're working right. It cost us probably $50,000, and I don't think that a state Medicaid program is going to be motivated to spend the money that it takes to make sure these things are working right. Or even these private companies that are running credit predictions, housing predictions, recidivism predictions—unless the cost is internalized on them through litigation, and it's understood that "hey, eventually somebody's going to have the money to test this, so it better be working."[46]

It is very important to understand that at the heart of all this is human decision making, not artificial technological decision making. Humans set the budget, design the AI project, decide on the data,

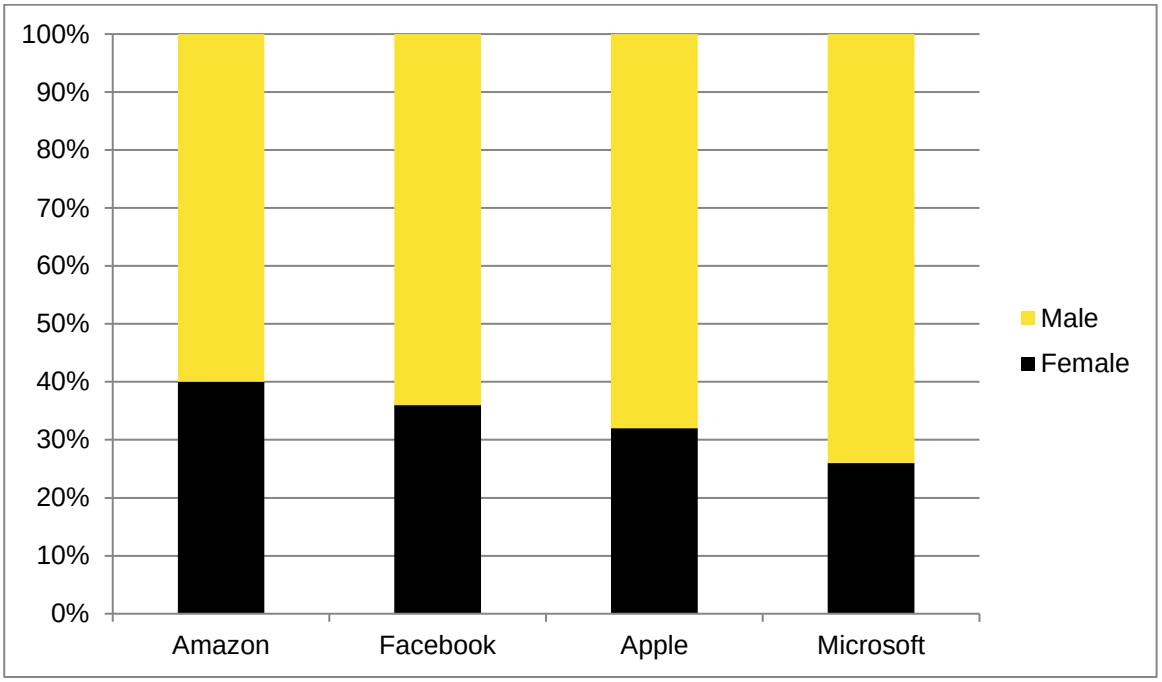
[42] Discussed below.
[43] M. Hamblen, 'IBM and Google Create New Certifications for Data Scientist and Cloud Skills', *SdxCentral*, January 29 2019.
<https://www.sdxcentral.com/articles/news/ibm-and-google-create-new-certifications-for-data-scientist-and-cloud-skills/2019/01/>
[44] Broussard, op. cit. n. 16. Chapter 7.
[45] *Toby Schultz v Richard Armstrong* Case No. 3:12-CV-58-BLW
[46] J. Stanley, 'Pitfalls of Artificial Intelligence Decision Making,' ACLU, June 2 2017.
<https://www.aclu.org/blog/privacy-technology/pitfalls-artificial-intelligence-decisionmaking-highlighted-idaho-aclu-case>

write and debug the code, calibrate it against real human decisions, decide when to put it into operation and use it in the real world. These humans can be problematic.

1. The problem of AI bias

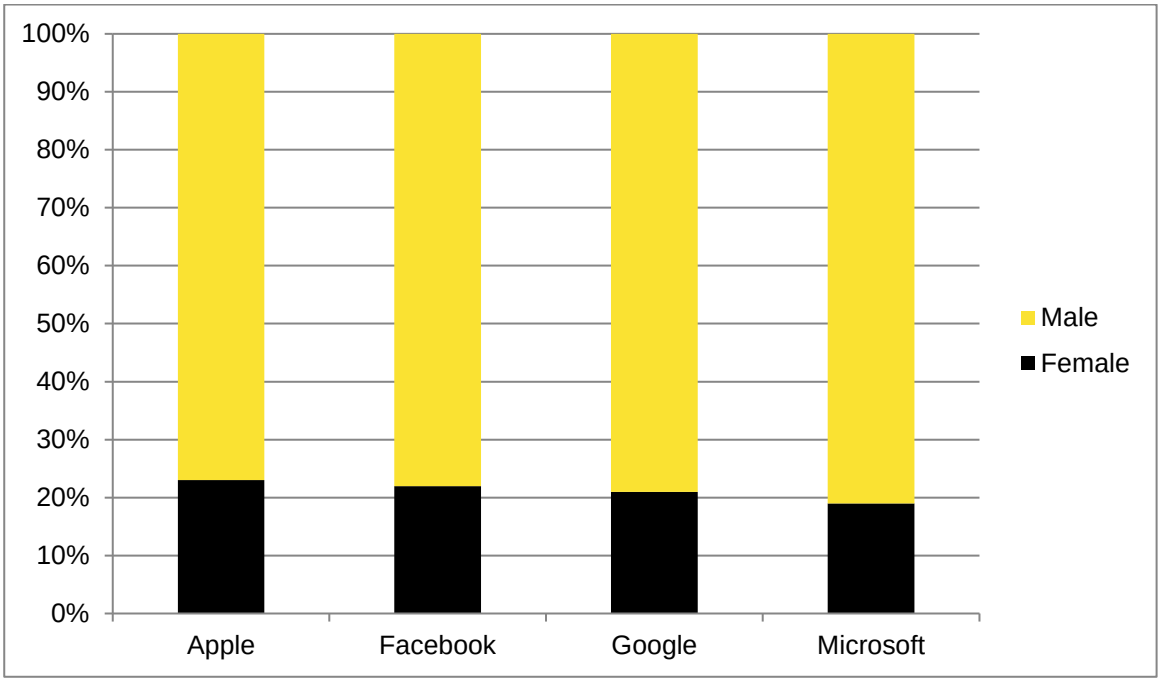
If those humans, particularly those designing the project, are not representative of society, have explicit and/or unconscious worldviews, this can strongly bias the outcomes. So who those human are matters a lot, and as Charts below demonstrate, they are mostly men particularly in the technical AI roles that matter and mostly white men.[47]

Chart 1: Leading AI Tech Companies Gender Breakdown Worldwide 2017[48]



[47] White men predominate with the largest minority being Asian males. M. Garcia, 'Racist in the Machine: The Disturbing Implications of Algorithmic Bias.' *World Policy Journal*, (2016) vol. 33 no. 4, 111-117, 114.
[48] Reuters Data, 2018.

Chart 2: Leading AI Tech Companies Gender Breakdown Worldwide Technical Roles 2017[49]



For example, if the data for testing an AI model is chosen by men who have unconscious bias they may choose samples that are representative of their world experience but unrepresentative generally. This has happened already where systems have been calibrated/trained unthinkingly with dominant images of white men.[50] Similarly with word-embedding AI techniques necessary for understanding written language or the spoken word, traditional gender roles have been attributed by the coders to women and men. In this attribution system women are homemakers and men are scientists.[51] Predominantly white men are responsible for designing AI systems so the AI systems reflect those designers' conscious and unconscious world view. At a practical level the AI systems can, as a result, have operational problems. Facial recognition systems are very good at recognizing white men but poor at recognizing black women.[52] Google's voice recognition system designed and

[49] Reuters Data, 2018.
[50] J. Buolamwini, and T. Gebru, 'Gender Shades.' *Proceedings of Machine Learning Research* (2018) 81:1–15.
[51] T. Bolukbasi, K. Chang, J. Zou, V. Saligrama, and A. Kalai. 'Man is to computer programmer as woman is to homemaker?' in *Advances in Neural Information Processing Systems* (eds) D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, (2016) 4349–4357.
[52] Buolamwini. Op. cit. n.49.

tested on men has significant problems recognizing women's voices.[53] Siri, Apple's virtual assistant, initially refused to provide information on abortion services.[54] California's Criminal Justice COMPAS algorithm, used to assess risk of reoffending, was found to be racially biased.[55] Google's facial recognition system was so poorly calibrated on non-white faces it misidentified black people as gorillas, while Microsoft's Titter bot Tay quickly became a sex crazed racist when programmed to learn from other twitter users' views.[56] These outcomes are all the product of human bias, not determinist technological autonomy.

AI hiring processes, for example, are emerging in the employment sphere and in those cases the calibrating "training" is based around specific definitive known data decision outcomes – those who were hired previously. CVs are analysed to see if they have the characteristics of those who were successfully hired in the past. In the training or calibrating process the AI program runs text analysis on past CVs to find a pattern that matches the known successful outcomes of those hired. If it gets a pattern that matches the known outcomes of those previously hired by humans to about 90% accuracy an operational model exists that could work with the CVs of new applicants.

In August 2018, Amazon abandoned its AI hiring program when it discovered that it was eliminating women from consideration. Based on 10 years previous hiring data the programme excluded women and sometimes hired unqualified workers. On investigation, Amazon was able to identify key aspects of the AI decision-making that led to those outcomes. The AI was "trained" as normal against a decade of CVs of successful candidates. The elements the AI favoured in reaching its decisions turned out to have significant flaws. Words in the CV indicating membership of "women's" clubs or graduates from all women colleges eliminated or downgraded candidates and

[53] R. Tatman, 'Google's speech recognition has a gender bias,' *Making Noise and Hearing Things*, July 12 2016, https://makingnoiseandhearingthings.com/2016/07/12/googles-speech-recognition-has-a-gender-bias/ and H. Devlin and A. Hern, 'Why are there so Few Women in tech,' *The Guardian*, August 8 2016, <https://www.theguardian.com/lifeandstyle/2017/aug/08/why-are-there-so-few-women-in-tech-the-truth-behind-the-google-memo>

[54] D. Rushe, 'Siri's Abortion Bias Embarrasses Apple,' *The Guardian*, December 1 2011, <https://www.theguardian.com/technology/2011/dec/01/siri-abortion-apple-unintenional-omissions>

[55] Broussard, op. cit. n. 16. pp 154-156 and A. Chouldechova, 'Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments,' (2017) *Big Data*. Issue 2: 5

[56] C. Dougherty, 'Google Photos Mistakenly Labels Black People 'Gorillas',' *The New York Times*, July 1 2015, <https://bits.blogs.nytimes.com/2015/07/01/google-photos-mistakenly-labels-black-people-gorillas/> and Garcia, op. cit. n. 47, p. 112.

more subtle eliminations occurred by the AI favouring applicants who used male engineer type verbs such as "executed" or "captured". In other words, the AI had found that the way to get a 90% accurate score in training on past data outcomes was to favour male characteristics and directly discriminate against women.[57] In a tech industry dominated by men it had simply reflected and revealed the bias of past decision-makers to determine future outcomes.[58]

While this made headlines around the world, what perhaps should have made headlines was that this has been happening with employment based computer programmes since the 1970s. The most famous example occurred at St George's Medical School in London where a computer hiring programme excluded women and those with non-European names from consideration. The issue was the same as the Amazon AI example: the program had embedded the previous bias of past human decision makers, which had gone unnoticed by designers and users with similar world views. In 1987, an investigation by the Commission for Racial Equality found that the hospital had engaged in racial and sexual discrimination.[59] Amazon abandoned its hiring AI system shortly after its problems came to light but other major companies such as Unilever and Microsoft's LinkedIn use a similar system or are seeking to implement one.[60] Other Amazon AI decision-making systems with similar problematic outcomes are in use in the public sector within the policing and quasi judicial systems of the UK and US, where the outcomes of their decisions have similarly shown bias; in particular racial bias.[61] A bias-generating AI system, it seems, is unacceptable when operating within Amazon itself but exporting it to other organisations to inflict on the wider public is acceptable. In a broader sense control of AI development within the tech

[57] AI is a fantastic tool for revealing bias.
[58] J. Dastin. Amazon scraps secret AI recruiting tool, *Reuters,* October 10 2018. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>
[59] S. Lowry and G. MacPherson, 'A Blot on the Profession,' (1988) *The British Medical Journal*, Vol 296, No. 6623, pp.657-8 and Garcia, op. cit. n. 47.
[60] R. Booth, 'Concern as Unilever uses AI to do work of recruiters' *The Guardian,* 25 October 2019 and J. Burn-Murdoch, 'The problem with algorithms: magnifying misbehaviour,' *The Guardian*, 14 August 2013, <https://www.theguardian.com/news/datablog/2013/aug/14/problem-with-algorithms-magnifying-misbehaviour>
[61] New Statesman/IBM, 'AI in the public sector,' (2018) *New Statesman* <https://www.newstatesman.com/sites/default/files/ns_ibm_supplement_june_2018_1_.pdf >, J. Angwin, J. Larson, S. Mattu and L. Kirchner, 'Machine Bias,' *ProPublica,* 23 May 2016, <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>and M. Burgess, 'UK police are using AI to inform custodial decisions,' *Wired*, 1 March 2018, <https://www.wired.co.uk/article/police-ai-uk-durham-hart-checkpoint-algorithm-edit>

industry exports gender, racial and technical bias (see the discussion of the Black Box proposition below) outside the tech sphere to companies where those specific bias have not been present or at least not present to the extent they are within a narrow male science dominated tech industry. This tech bias contagion has the potential to do enormous harm.

In a high quality AI model the outcomes should always have human oversight because, as discussed below, it is very difficult to tell a counterintuitive but possibly brilliant outcome from a dangerous error or an unlawful outcome. As Shah noted in 2017, "How to ensure that the algorithm is fair, how to make sure the algorithm is really interpretable and explainable - that's still quite far off".[62] In short, most AI is not an intelligent technological agent but rather one designed, made and implemented by a particularly narrow group of humans. AI can have significant problematic design bias, is unable to entirely replicate previous human decision making, while simultaneously reflecting any bias in past human decisions. This is widely misunderstood even by the AI tech industry, as the Amazon example illustrates.

2. Black Boxes and technical development bias

However, not all AI is basic machine-learning based. Some tasks, such as image recognition and speech recognition/translation, contain nuance and subtleties that normal machine learning will not capture. As such, more complex deep-learning AI models such as artificial neural networks have been developed to operate slightly differently, combining various machine-learning statistical algorithms into a framework originally designed to mimic how the neural networks in the human brain operate.[63] So, from examples of tagged cat images, an artificial neural network can work with new untagged images and eventually identify with hopefully 90%-plus accuracy a cat or maybe

[62] A. Theodorou, R. Wortham and J. Bryson, 'Designing and implementing transparency for real time inspection of autonomous robots,' (2017) *Connection Science,* 29:3, 230-241

[63] W. Sarle, 'Neural networks and statistical models.' (1994) *Proceedings of the Nineteenth Annual SAS Users Group International Conference,* 1538–1550.

even a particular cat. However, their complexity can come with a significant flaw – explaining the basis for the program's decisions can be difficult- this is the root of the black box proposition.[64]

The black box proposition is that the decision making inside the AI system is so complex that it is not possible to know why a particular outcome arose. This does not occur because inside the AI black box is an inscrutable deep thinker, it is still a statistical model, but because the AI has been put into use when only half built: as it is missing a key diagnostic component that would explain the outcomes. For the same reason that bias has historically been a problem with AI and computer program's generally, so too these systems have been deployed because a narrow group of human designers have not been concerned to know the basis of the decision, just that it works within certain technical parameters. In some situations, such as image analysis, where someone is searching for cat pictures on the internet, this technical development bias might be fine as explaining why a cat image was chosen won't be important, but where explainability is important it is problematic.[65]

These systems have been deployed because the designers have not been concerned about the basis of the AI's decision, but with its operational "beauty", even if they have no idea how it works: a form of maths/engineering mindset bias.[66] For computer scientists and engineers the key may be that the AI produces seemingly high percentage outcomes, but for lawyers, medical staff, citizens subject to its outcomes and ultimately for these systems to be lawful or useful in operation, the key is the basis upon which those operational decisions are made. Unfortunately, explainability has not until recently been part of the design remit. [67]

This lack of explainability also raises a significant legal red flag with the application of deep learning AI black box systems in the public sphere. Citizens can challenge state decisions based on flawed decision making or in any situation where recourse to the courts or appeal from the courts

[64] P. Voosen, 'How AI detectives are cracking open the black box of deep learning,' *Science*, 6 July 2017, <https://www.sciencemag.org/news/2017/07/how-ai-detectives-are-cracking-open-black-box-deep-learning>
[65] J. Zerilli, A. Knott, J. Maclaurin, and C. Gavaghan, 'Transparency in Algorithmic and Human Decision-Making: Is There a Double Standard?' (2018) *Philos. Technol.* 1-23 <https://doi.org/10.1007/s13347-018-0330-6>
[66] S. Hossenfelder, *Lost in Math: How Beauty Leads Physics Astray*, (2018).
[67] Zerilli, op. cit. n.64.

depended on knowing the reasons for a decision.[68] Similarly, as observed below, knowing the reason for a medical diagnosis or why a self-driving car crashed is crucial.

In the health sphere, both Google and IBM have been developing AI health applications based around their general AI programmes. IBM's Watson for Oncology offers a salutary lesson in the limits of the black box. As with Google's AlphaGo game-playing AI, IBM developed Watson to play a popular US TV game called Jeopardy where the contestants are given the answer to a question and they must provide the question. Six years and enormous resources went into the project with Watson eventually triumphing over a Jeopardy grand master in 2011. Finding a business outlet for Watson's very specific skill set proved challenging but eventually Watson for Oncology was the outcome. The only problem was, the skill set for Jeopardy was not easily transferable to cancer treatment. The solution was to have a range of outcomes provided by one hospital in the US for Watson to draw on. As Ross and Swetlitz found in 2017 in an investigation into the problems being reported about Watson for Oncology:

> "The system is essentially Memorial Sloan Kettering in a portable box. Its treatment recommendations are based entirely on the training provided by doctors, who determine what information Watson needs to devise its guidance as well as what those recommendations should be."[69]

In some ways IBM became a victim of the general AI misunderstanding and its own marketing claims that Watson would cure cancer, when really Watson is no different from other types of machine learning where tags such as "cat" or "not a cat" are used to identify images to train the AI on. In this case, it was doctors at one US hospital that did the tagging as to symptoms and treatment plans.

In the context of medical care, Watson for Oncology revealed certain key problems. First, in a complex field such as cancer care a narrow data set of cancer diagnosis based in one hospital in a middle class US setting is unlikely to be easily scalable for treatment all over the world, where very

[68] See for example *Council of Civil Service Unions v Minister for the Civil Service* [1984] UKHL 9
[69] C. Ross and I. Swetlitz, 'IBM Pitched its Watson Supercomputer as a revolution in cancer care.' *STAT*, 5 September 2017, <https://www.statnews.com/2017/09/05/watson-ibm-cancer/>

different medical and nutritional circumstances are present. Second, those using Watson found it difficult to understand why it was recommending a particular course of treatment, when it was counterintuitive or possibly wrong. Watson could provide literature to read that might be relevant but could not explain why it was relevant or the key reasons for the recommendation. The black box again proved problematic. Responding to this demand IBM now sells an AI product that it claims helps understand the Black Box outcomes of its other AI products.[70] Things might have been different had IBM set out to design an AI application to cure cancer, rather than set out to design an AI to play Jeopardy.

In March 2018, an autonomous Uber car crashed into and killed a woman wheeling a bicycle across a road during testing. Problems with the maintenance of the car and the failure of the human safety backup driver in the car combined with its AI decision making to cause the death.[71] Tesla similarly had a fatality linked to the failure of its Autopilot driving system in 2016. While the accidents dented public confidence in autonomous vehicles, the implications for AI more generally was that the investigators found it difficult to determine the decision making of the AI, particularly in the Tesla fatality.[72] Decisions inside the AI black box were so complex it seemed, that they could never discover its role in the fatality.

This marks a significant change to the regulatory challenge, where knowing what happened is essential for a range of regulators from road/air safety to medical and financial. The proposition from the manufacturers is that where AI is involved, it is an autonomous intelligent black box and not a statistical/probability model designed and implemented at too early a stage by the company. If the AI made the decision autonomously then the proposition also seems to be that it was not Uber, Tesla, IBM or Google's fault. With this approach even weak AI would have to be accommodated in an autonomous legal liability process.[73] Given the widespread misunderstanding about it, this is not

[70] IBM Openscale <https://www.ibm.com/cloud/watson-openscale>.
[71] National Transportation Safety Board, (2018) Preliminary Report HWY18MH010, NTSB
[72] J. Stilgoe and A. Winfield, 'Self-driving car companies should not be allowed to investigate their own crashes,' *The Guardian*, 13 April 2018, <https://www.theguardian.com/science/political-science/2018/apr/13/self-driving-car-companies-should-not-be-allowed-to-investigate-their-own-crashes>
[73] C. Reed, E. Kennedy, and S. Silva, 'Responsibility, Autonomy and Accountability: Legal Liability for Machine Learning' (2016). *Queen Mary School of Law Legal Studies Research Paper* No. 243/2016.

an unrealistic scenario as tech companies already push the black box scenario. However, the point of this black box analysis here is not to pursue the legal implications of a possible mistake that has not happened yet but to inform debate so that the mistake is less likely to happen in the first place, as the only situation where the AI black box could be accepted as an autonomous decision maker is where the nature of weak AI and its poor human design is misunderstood. The autonomous black box is an illusion. A black box is a half finished, poorly designed, admittedly complex statistical model, that has been deployed too early in its development. As Stilgoe considered in response to the black box autonomy proposition of self driving cars:

> The terms 'self-driving cars', 'autonomous vehicles' and 'driverless cars' have been used almost interchangeably in public discourse... The differences in nuance implied by these terms should not distract us from a larger concern, which is with the rhetoric of autonomous technology. Technology, however, is never self-driving…Claims that technology has a will of its own… typically disguise a political agenda that is libertarian and deregulatory.[74]

3. Private v Public Interest AI design

Some AI designers, however, understand that because AI is a product of human design the potential exists for AI to shape the world to their own private ends in novel AI circumstances. This presents particular public interest regulatory concerns. Fears arose in the AI literature that in the trolley problem[75] or lose/lose scenarios, the autonomous car AI would be programmed to save the car passengers over those outside it.[76] Indeed, Mercedes executives appeared to confirm that Mercedes AI systems were programmed to save the driver and passengers over those outside the car.[77] Important legal concerns arise here. AI does genuinely present novel possibilities for better decision making that lie outside the realm of human ability but only if the normative public interest

---

[74] J. Stilgoe, J. 'Machine learning, social learning and the governance of self-driving cars.' (2018) *Social Studies of Science*, *48*(1), 25–56 at 35.
[75] In a classic trolley test a human has a choice between switching a trolley (tram/train) car that will kill a group of people to another track where it will kill only one. See P. Foot, *Virtues and Vices* (1978) Chapter 2.
[76] B. Casey, 'Amoral Machines, Or: How Roboticists Can Learn to Stop Worrying and Love the Law' (2017) *Northwestern University Law Review*, Vol. 111, No. 5, pp.231-250.
[77] M. Taylor, 'Self Driving Mercedes will prioritise occupant safety over pedestrians,' *Car and Driver*, 7 October 2016 <https://www.caranddriver.com/news/a15344706/self-driving-mercedes-will-prioritize-occupant-safety-over-pedestrians/>

aspect of the human AI design decisions are recognized and implimented. An autonomous car can make better decisions because its sensors and AI decision making can potentially react quicker than a human can. For example, if something appears suddenly in front of a human driver their reaction is to break suddenly or swerve to avoid the object. The reaction will be the same whether the object is a plastic bag or a human. In each case, the driver chooses to place themselves, the passengers and potentially other road users at risk over an unknown external object because the human cannot determine what the object is in time to make the best decision. In the plastic bag example, breaking or swerving is not likely to be the best choice. The AI can potentially make a better choice if programmed to place the lives of those in the car and other road users in priority to the bag. However, in the trolley test situation where it is a choice between saving the driver and passengers or someone stepping onto the street in front of the car, choices are already being made by the AI industry to favour the driver and the passengers. After all, the industry logic goes, they paid for the product. [78]

Similarly, Facebook's AI algorithms are deliberately used to shape people's perception of events and can facilitate discriminatory behavior.[79] China's development of AI is primarily driven by its recognition of AI's design potential to achieve authoritarian ends. Designed bias was exposed in the financial sector when in 2010 the bank Capital One used its understanding of algorithmic bias to deliberately channel ethnic minority customers to higher interest credit cards.[80] Online retailers, including Amazon, use or have used differential pricing based on algorithm bias.[81] In 2010 and 2015 flash crashes in the financial markets were blamed on algorithmic trading. This was correct but what was little understood was that these were not rogue trading algorithms; these were

[78] J. Bonnefon, A. Shariff, and I. Rahwan. 'Autonomous Vehicles Need Experimental Ethics: Are We Ready for Utilitarian Cars?' (2015) *Science* 352, 1573–1576, P. König, and P. Gordon. 'Using Virtual Reality to Assess Ethical Decisions in Road Traffic Scenarios,' (2017) *Frontiers in Behavioral Neuroscience.* 11, 122.
[79]W. Knight, 'Forget Killer Robots Bias is the Real AI Danger,' (2017) *MIT Technology Review*, 3 October <https://www.technologyreview.com/s/608986/forget-killer-robotsbias-is-the-real-ai-danger/>, A. Madrigal. 'What Facebook did to America,' *The Atlantic*, 12 October 2017, <https://www.theatlantic.com/technology/archive/2017/10/what-facebook-did/542502/>and Department of Housing and Urban Development. "HUD charges Facebook with housing discrimination" 28 March 2019, Press Release no. 19-035.
[80] Garcia, op. cit. n.47. p.114.
[81] V. Heffernan, 'Amazon's Prime Suspect,' *The New York Times*, 6 August 2010, <https://www.nytimes.com/2010/08/08/magazine/08FOB-medium-t.html>

algorithms that were deliberately designed by traders to manipulate the market and cause the crash.[82] Novel "designed" AI decision making such as this, have important public interest regulatory implications. It should not be left to the AI designers to decide financial market prices, interest rates, the price of goods, who to vote for, or who lives and dies.

4. Complexity Impacts

As has been observed AI can have problematic embedded bias, is not quite as good as humans at decision making, its outcomes can lack transparency and can be designed to work against the public interest. An additional exacerbating factor that drives potential AI harm is the general misunderstanding that AI has superior intelligence to humans and therefore that the human user and humans subject to the AI defer to its outcomes even when an outcome is manifestly problematic. In the medical profession there is a long and painful history of deference to computer outcomes in complex diagnosis situations such as radiology dosages, where experienced professionals have simply accepted clearly mistaken dosages produced by a computer. As a New York Times investigation found in 2010 in the context of computerized radiology dosages:

> while this new technology allows doctors to more accurately attack tumors and reduce certain mistakes, its complexity has created new avenues for error — through software flaws, faulty programming, poor safety procedures or inadequate staffing and training. [H]ospitals… are often too trusting of the new computer systems and software, relying on them as if they had been tested over time, when in fact they have not.[83]

The assumed superior ability of the computer to make the right decision even when the decision was manifestly incorrect impacted here even where experts were involved.

[82] S. Brush, T. Schoenberg, S, Ring, 'How a Mystery Trader with an Algorithm May Have Caused the Flash Crash,' *Bloomberg News*, April 22, 2015, <https://www.bloomberg.com/news/articles/2015-04-22/mystery-trader-armed-with-algorithms-rewrites-flash-crash-story> and A. Weinberg, 'Should You Fear the ETF? ETFs are scaring regulators and investors: Here are the dangers—real and perceived.' *Wall Street Journal.* December 6 2015, <https://www.wsj.com/articles/should-you-fear-the-etf-1449457201>

[83]W. Bogdanich., 'Radiation Offers New Cures, and Ways to do Harm,' *The New York Times*, 23 January 2010, <https://www.nytimes.com/2010/01/24/health/24radiation.html>

In 2018 and 2019 two crashes within a few months of each other involving Boeing's new 737 Max 8 quasi-autonomous fly-by-wire system have thrown up issues of complexity atrophy and AI complexity. Pilots cushioned by years of fly-by-wire autonomous flying may have been unable to deal with suddenly having to fly a plane in difficult circumstance where that plane has been aerodynamically designed to be primarily piloted or assisted by AI or may have been fighting to correct a fly-by-wire decision made by the computer system but not understood by the pilot.[84] These combinations may also have greater impact where AI is being used by non-experts such as in an employment or public sector situation. Richard Eppink, the legal director of the ACLU Idaho who has been at the forefront of litigation challenging problematic AI decision making in the public sector considered:

> My hunch is that this kind of thing is happening a lot across the United States and across the world as people move to these computerized systems. Nobody understands them, they think that somebody else does—but in the end we trust them. Even the people in charge of these programs have this trust that these things are working.[85]

In the AI context, complexity/superiority atrophy is compounded by potential bias, the limits of AI decision-making versus humans and the black box problem, making it difficult for users to tell the difference between a counterintuitive and potentially life saving AI outcome and a mistake.

5. Understanding the bargain

Even the highest quality AI decision-making models present challenges for the user and those subject to its decision-making. As already considered, AI decision-making is not 100%

[84] G. Topham, 'Ethiopian flight 302: second new Boeing 737 to crash in four months,' *The Guardian*, 10 March 2019 <https://www.theguardian.com/world/2019/mar/10/ethiopian-flight-302-second-new-boeing-737-max-8-to-crash-in-four-months> and J. Stilgoe, 'Machine learning, social learning and the governance of self-driving cars.' (2018), *Social Studies of Science*, 48(1), 25–56, at 34.

[85] J. Stanley, 'Pitfalls of Artificial Intelligence Decision making Highlighted In Idaho ACLU Case.' <https://www.aclu.org/blog/privacy-technology/pitfalls-artificial-intelligence-decisionmaking-highlighted-idaho-aclu-case>

accurate relative to the human decisions it has been trained against. A really high quality system might get to 97% but many systems are in use at lower accuracy rates. 90% plus accuracy in the calibrating/training process is also not what it seems. As noted before, most large AI data sets are incomplete and have been statistically enhanced to complete them so their accuracy may well be below 90% in application, as some of the data is incorect. Poorly designed software and bugs in the software and problems with hardware can also affect AI accuracy. Indeed, large scale human facial recognition AI systems deployed by the Metropolitan police and South Wales Police have been revealed to have accuracy rates of only 2% and 10% respectively, generating thousands of false positive criminal identifications of innocent citizens.[86] Uncertain environments that do not conform to the training environment can also be problematic. For example currency trading AI used to trade UK currency has been suspended as the unpredictability of fluctuations due to Brexit made it useless as it was trained on historical data.[87] An ahistorical event is impossible to train for. Similarly, AI judicial decision making models become unreliable as new variables appear that depart from the training data.[88] In essence when the future is not like the past AI struggles and if not recognized it locks in a future based on the past.

All this is acceptable and still useful if decisions relate to whether an image is a dog or a cat or predictive text outcomes but 90% accuracy is much less acceptable in other context. In a judicial criminal process as an example, where ostensibly no innocent person should go to jail, then a high quality 97% accurate AI decision system implemented in a judicial context would move to 3 in a hundred innocent citizens, or higher if from a minority group, going to jail who would not have if humans made the decisions. That same system will have much lower accuracy as new variables are encountered that depart from its training data. As mentioned previously, certain types of black box

[86] Sharman, J. 'Metropolitan Police's facial recognition technology 98% inaccurate,' *The Independent,* May 13th 2018, <https://www.independent.co.uk/news/uk/home-news/met-police-facial-recognition-success-south-wales-trial-home-office-false-positive-a8345036.html>

[87] E. Szalay, 'The pound's 'breathless marathon' tests traders nerves,' *Financial Times,* April 12 2019 <https://www.ft.com/content/5ce4ae6a-5d0d-11e9-939a-341f5ada9d40>

[88] S. Levmore, and F. Fagan, 'The Impact of Artificial Intelligence on Rules, Standards, and Judicial Discretion.' (2019) *Southern California Law Review,* 93, <https://ssrn.com/abstract=3362563>

AI would present additional insurmountable challenges in both the criminal context and in judicial review, as the reason for the decision would remain unknown. In 2019 Partnership on AI produced a report stimulated by the widespread use and planned mandatory use of AI decision making in the US judicial criminal system. In recommending that such systems should not be used to solely determine questions of individual liberty it stated:

> [u]sing risk assessment tools to make fair decisions about human liberty would require solving deep ethical, technical, and statistical challenges, including ensuring that the tools are designed and built to mitigate bias at both the model and data layers, and that proper protocols are in place to promote transparency and accountability. The tools currently available and under consideration for widespread use suffer from several of these failures…[89]

In the medical context similar failures might similarly mean 3 in 100 dying who would not have otherwise died, or more if they did not match the training dataset. In the employment context a 90% accurate system would mean 10 in 100 not hired who would have been hired in a human decision making process. Those 10 will likely be statistical outliers who do not fit the norm for the organization, most likely women, ethnic minorities or those with different educational qualifications or backgrounds. Potential organizational innovators would be excluded to the detriment of the organisation.

Despite the limits of AI decision-making, the potential for bias, the black box issues, the private v public interest clash and the potential for complexity/superiority atrophy, a huge driver of AI decision-making development for the public sector, private sector companies and their clients is cost. Cost savings for AI over human decision-making are enormous.[90] Unilever estimates a saving of 100,000 person hours or aprox £800,000 using its HR AI system instead of interviewing and assessing candidates using humans.[91] Weighed against the long term diversity and innovation of the organization in terms of candidate sex, race ethnicity and educational background, the balance is

[89] Partnership on AI, 'Report on Algorithmic Risk Assessment Tools in the US Criminal Justice System,' (2019) p. 3 https://www.partnershiponai.org/report-on-machine-learning-in-risk-assessment-tools-in-the-u-s-criminal-justice-system/ >
[90] McKinsey, 'Disruptive Technologies', (2013) *McKinsey and Company.* p. 43 <https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Disruptive%20technologies/MGI_Disruptive_technologies_Full_report_May2013.ashx>
[91] Booth, op.cit. n.59.

likely to fall in favour of cost saving. Indeed, even where diversity and innovation is valued by an organization, the AI exclusion of diverse innovators may not be recognized or quantifiable in the cost evaluation of AI implementation. While the flaws are problematic and potentially unlawful they are also complex to understand and shrouded in an exciting sci-fi utopian veneer. If it is good enough and the cost savings are huge then it will get implemented flaws and all.

While in the Mercedes (passengers over pedestrians) and Capital One (charge minorities more) examples the companies recognized the normative possibilities of coding the AI to achieve an outcome that favoured their private interests over the general public, amazingly, Amazon with all its resources in AI development seemingly gave no thought to the basic statistical problem of data bias or that the AI outcomes in an employment context had to be legal. The narrowness of technical design focus, utopian assumptions that the tech is good or the notion that this is a new unregulated frontier has left a troubling absence of basic normative legal design planning at the heart of AI decision-making systems and allows the possibility of both designing against the public interest or freezing historical bias forever, locking out and locking up minorities and outliers from participation in society.

In 2019, MMC Ventures produced a report on the state of AI and identified the potential for AI to facilitate illegal practices and discrimination in housing, employment, education, social welfare, insurance, finance, differential pricing, liberty, privacy and dignity as well as undermining democratic institutions.[92] That is quite a list. The possibility of AI as both a designed and an unthinking driver of inequality is of huge concern. That regulatory space needs to be filled rapidly as poorly designed and potentially strongly narrow private interest serving AI is increasingly rolled out across the public and private sectors, while remembering that it is not the tech that is our central problem but the flawed humans behind its implementation. The mechanism for tackling this will be returned to in the conclusion. The next section considers the unusual human governance behind the tech AI companies, as this forms another important part in understanding the dysfunctional human's shaping the AI field.

[92] MMC, op. cit. n.91 Chapter 8.

AI GOVERNANCE: THE MEN AND ONE WOMAN BEHIND THE GOVERNANCE OF AI

While there are problematic humans on the technical side of designing and implementing AI, if the human contextual lens is set wider, it also reveals problematic humans behind the façade of the AI companies. Five years ago, the major global AI players were UK and US companies, with the UK's DeepMind probably the most advanced AI developer in the world. Since then, the leading UK companies have been bought by the US ones, including DeepMind, which is now owned by Google, and the Chinese government has become a significant force in AI development.[93] Although the UK AI developers are still based in the UK they are now owned and controlled from the US. The global balance of leading AI development has split between the US and China.[94] While China is considered later when discussing dysfunctional public governance, this part considers the private sector governance of the leading private sector AI developers, Uber, Amazon, Facebook, Microsoft, Google, Apple, IBM and Tesla. As will be observed, they almost all have unusual private sector governance structures designed in general to give control to a small group of insiders with tight connections to each other and very similar backgrounds and interests. Governance in these leading AI companies is in general unusually autocratic and unaccountable even in the context of weak private sector accountability. The section proceeds by examining each company and its key ownership and control features.

Uber was until very recently a private company subject to little public accountability. In May 2019 it became a public listed company. It has two key controllers. The first is Travis

[93]S. Gibbs, 'Google Buys two more UK Artificial Intelligence Startups,' *The Guardian*, 23 October 2014 <https://www.theguardian.com/technology/2014/oct/23/google-uk-artificial-intelligence-startups-machine-learning-dark-blue-labs-vision-factory>, M. Murphy 'Facebook buys British artificial intelligence company Bloomsbury,' *The Telegraph,* 2 July 2018, <https://www.telegraph.co.uk/technology/2018/07/02/facebook-buys-british-artificial-intelligence-company-bloomsbury/> and J. Nimmo (2016) Microsoft Buys SwiftKey, The Independent, 4 February 2016 <https://www.independent.co.uk/news/business/analysis-and-features/microsoft-buys-swiftkey-how-artificial-intelligence-became-a-hot-uk-export-a6852266.html>

[94]I. Burrows, 'China's research in artificial intelligence 'far outranks' Huawei threat,' *ABC News,* 2 February 2019, <https://www.abc.net.au/news/2019-02-02/china-research-artificial-intelligence-bigger-threat-than-huawei/10685420 > and A. Slaughter, What Will Really Determine the Winner of the U.S.-China Rivalry Over A.I.?, SLATE, 12 March 2019, <https://slate.com/technology/2019/03/united-states-china-rivalry-artificial-intelligence.html>

Kalanick, who resigned as Chief Executive Officer (CEO) in 2017 because of accusations of perpetuating a culture of sexual harassment and fraud litigation brought by a major shareholder. Kalanick also famously developed and implemented a programme called Greyball to evade law enforcement investigations into Uber's activities globally.[95] The fraud litigation was eventually settled by a broad restructuring of the ownership and control of Uber, whereby Kalanick's power and shareholding was reduced through a major investment by Softbank. Kalanick remains a significant shareholder but is no longer a lone controlling shareholder: he sits on the board and has the right to appoint three board members. As a result, Softbank is now the major shareholder in Uber and has the right to appoint two board members. [96] Together, Softbank and Kalanick hold 22% of Uber's shares. Softbank is a Japanese listed company owned and controlled by Masayosi Son through direct control of 25.5% of its shares. Son is also the president, chairman and CEO of the company.[97] Son is as a result the second significant controller of Uber. Softbank's investment in Uber is part of an explicit attempt to shape the direction of AI to Son's personal futuristic agenda through Softbank's $100 billion vision fund. He has been described both as "crazy' (he once threatened to set him self on fire in a clash with regulators) and as the most powerful person in Silicon Valley.[98] Softbank lists his centrality and unpredictability as one of its key business risks.[99]

The other leading AI companies have been listed for decades so all are now subject to a range of enhanced private sector accountability norms. While criminal and specific market-oriented regulation designed to protect the public operate in the background, overseen by the US Securities and Exchange Commission (SEC), it is important to understand that private sector governance does not aim at any sort of high level accountability process beyond legal duties focused on fidelity and

[95]M. Isaac, 'How Uber Deceives the Authorities Worldwide,' *The New York Times*, 3 March 2017. <https://www.nytimes.com/2017/03/03/technology/uber-greyball-program-evade-authorities.html>
[96] R. Molla and T. Schliefer, 'Here's who Controls Uber,' *Recode*, 8 Jan 2018 <https://www.recode.net/2018/1/8/16865598/uber-softbank-control-board-power-stocks-benchmark-travis-kalanick-dara-khosrowshahi>

[97] Marketscreener, *Masayoshi Biography*, (2019) < https://www.marketscreener.com/business-leaders/Masayoshi-Son-264/biography/ >
[98]K. Brooker, 'The most powerful person in Silicon Valley,' *Fast Company*, 14 January 2019, < https://www.fastcompany.com/90285552/the-most-powerful-person-in-silicon-valley > and K. Inagaki, L. Lewis, and A. Massoudi, 'Masayoshi Son: The unrepentant visionary,' *The Financial Times*, 22 July 2016, < https://www.ft.com/content/7b2da318-4f2d-11e6-8172-e39ecd3b86fc>
[99] <https://group.softbank/en/corp/irinfo/about/risk_factor

competence and certain historically problematic issues such as executive accountability to shareholders. UK and US board legal duties, while nodding at stakeholder governance, are ultimately primarily focused on shareholders' interests and are generally weak.[100] Additionally, the corporate governance codes[101] to which these listed companies adhere aim to underline certain high risk historical agency cost issues such as having an audit, remuneration and appointment committee, not combining the CEO and chair role (dangerous but acceptable in some codes), having independent non-executive directors, pay disclosure and shareholder engagement. The aim overall is to ensure that the executive directors are subject to robust monitoring on behalf of the shareholders. There are three problems with this when applied to these AI tech companies: The legal duties are weak and don't contain any public interest element, standards differ significantly depending on the corporate governance code used and in any case code compliance is optional. Non compliance has no particular consequence other than perhaps shareholder discontent. Third, and probably most significantly, in the tech sector the key board members and executives are generally the shareholders.

There are a number of organisations that monitor governance standards of companies worldwide. Institutional Shareholder Services (ISS) is one of the most significant and compiles a governance risk score for listed companies based on private sector governance compliance norms. A score of 1 is a low risk company in governance terms and a score of 10 is very high risk. Apple and Microsoft have overall scores of 1, IBM and Amazon have scores of 7 and Facebook, Alphabet (Google's parent company) and Tesla have scores of 10.[102] Apple, and Microsoft aside, private sector governance compliance is generally very poor in the tech companies, even within the context of a weak private sector set of accountability norms. However, the majority of these companies are also outliers within listed companies in terms of their ownership and control, which in many ways

[100] J. Armour, B. Black, Cheffins, B. and Nolan, C., 'Private Enforcement of Corporate Law: An Empirical Comparison of the UK and US' (2009), 6 *Journal of Empirical Legal Studies*, 687-722.

[101] In the US there are a number of corporate governance codes. S. Foley, 'The Battle of the US Corporate Governance Codes,' *The Financial Times*, 5 February 2017< https://www.ft.com/content/e52f6f22-e93c-11e6-893c-082c54a7f539> .

[102] ISS Governance Qualityscores can be found here <https://www.issgovernance.com/esg/rankings/governance-qualityscore/>

subverts even the weak private sector accountability processes focused on compliance in terms of accountability to shareholders. As with Uber, when key insiders are also significant shareholders then the accountability numbers become meaningless and the high number governance scores more alarming. Even low numbers can hide an accountability gap. Apple for example has a score of 1 but some of its key executives and board members are its largest individual shareholders.[103] Microsoft has largely migrated away from a founder led and controlled company but still has some significant insider holdings and Bill Gates, its remaining founder, still sits on the Board.[104] IBM is different from the other leading AI tech companies in that it is led by a woman, Virginia M. (Ginni) Rometty. However, as is the norm in the tech sector it also has poor governance and executive accountability generally, as its ISS score indicates, because among other things Rometty simultaneously holds all the key accountability roles on the board as the CEO, president and chair and she and other executives are the largest individual shareholders in the company.[105] Amazon similarly has poor governance practices and is founder-controlled both executively and through Jeff Benzos's 16.1% shareholding, which allows him to hold the chair, president and CEO roles, which waves a big red flag in corporate governance accountability terms as it provides one individual with a dangerous concentration of power and removes any internal board accountability. In 2014 Benzo's famously controlling leadership[106] was rewarded with the International Trade Union Congress "worst boss" award. In conferring the award the Congress noted:

> Amazon operating in Germany treats its workers as if they are robots. The company makes no secret that within just a few years they will replace workers with robots. A rich American corporation operating globally with disdain for dignity, for rights for working people. Jeff

[103] Apples SEC ownership filings can be found here <https://investor.apple.com/investor-relations/sec-filings/default.aspx> and J. Edwards, 'Top Five Apple Shareholders' *Investopedia*, 11 February 2019, <https://www.investopedia.com/articles/markets/120115/top-5-apple-shareholders.asp>

[104] Microsoft's SEC ownership filings can be found here <https://www.microsoft.com/en-us/Investor/sec-filings.aspx> and J. Walton, 'Top Five Microsoft Shareholders,' *Investopedia*, 30 January 2019 <https://www.investopedia.com/articles/investing/122215/top-4-microsoft-shareholders.asp?utm_campaign=rss_stock_analysis&utm_source=rss&utm_medium=referral>

[105] IBM's SEC ownership filings can be found here <https://www.ibm.com/investor/financials/financial-reporting.html > and J. Walton, 'Top Five IBM Shareholders,' *Investopedia*, 17 October 2019 <https://www.investopedia.com/articles/insights/052216/top-5-ibm-shareholders-ibm.asp>

[106] Amazon's SEC ownership filings can be found here <https://ir.aboutamazon.com/sec-filings> and J. Kantor and D. Streitfeld, 'Inside Amazon: Wrestling Big Ideas in a Bruising Workplace,' *The New York Times*, 15 August 2015, <https://www.nytimes.com/2015/08/16/technology/inside-amazon-wrestling-big-ideas-in-a-bruising-workplace.html?_r=0 >

> Bezos represents the inhumanity of employers who are promoting the American corporate model.[107]

Tesla, Facebook and Alphabet have extraordinarily poor governance practices and are founder-controlled through shareholdings or manipulation of shareholder voting rights. Elon Musk holds nearly a third of the shares in Tesla and until recently was its chair and CEO.[108] Although Tesla does not have any dual class voting rights that might give Musk additional voting control, from Tesla's historical SEC filings it is clear that it has supermajority voting clauses in its corporate constitution on issues such as ownership change and board participation, which effectively mean Musk's votes have important blocking rights.[109] His brother, Kimbal Musk, is also a shareholder and sits on the Tesla board even though he has no experience in the industry.[110] Over the course of 2018 and 2019, Elon Musk's lack of accountability and open contempt for US securities law and the SEC, led to a fraud action and settlement with the SEC, whereby the board of Tesla was reformed and Musk gave up his chair role. He also agreed to restrictions on his potentially market-manipulating public statements. He remains as CEO and controlling shareholder. In February 2019, he was alleged to have breached the SEC settlement by tweeting an uncleared misleading claim, after which Tesla's general legal council resigned. The SEC subsequently pursued a contempt charge against Musk.[111] In all, in 2018, over 40 Tesla executives resigned including its general legal council and his subsequent replacement (after two months), as well as the company's chief financial officer.[112]

Facebook is owned and controlled by Mark Zuckerberg through a dual voting structure that gives him direct control over 53.3% of the voting rights and an additional agreement with the second largest shareholder that allows him to control the voting rights of that shareholder as well.

---

[107]ITUC, 'Amazon's Jeff Bezos Wins ITUC's World's Worst Boss Poll,' *ITUC,* 22 May 2014 <https://www.ituc-csi.org/amazon-s-jeff-bezos-wins-ituc-s >

[108] <http://ir.tesla.com/corporate-governance/board-of-directors> < https://www.nasdaq.com/symbol/tsla/ownership-summary >and <https://www.nasdaq.com/quotes/insiders/musk-elon-831665 >

[109] Tesla's SEC ownership filings can be found here <http://ir.teslamotors.com/sec-filings>

[110] M. Brown, 'Tesla Investors Are Moving to Boot Elon Musk's Brother From Board,' *Inverse,* 9 May 2018, <https://www.inverse.com/article/44661-tesla-kimbal-musk>

[111] A. Sage, 'SEC seeks contempt charge against Tesla's Musk,' *Reuters*, 25 February 2019, <https://www.reuters.com/article/us-tesla-musk-sec-idUSKCN1QE2OR>

[112] S. Cao, 'Tesla Has Lost Over 40 Execs in a Year, Including 2 Top Lawyers in 2 Months,' The Observer, 20 February 2019, <https://observer.com/2019/02/elon-musk-dane-butswinkas-tesla-general-counsel-resigns/>

He is the controlling shareholder, chair and CEO.[113] This size of voting control, executive position and board role, gives Zuckerburg control over every accountability mechanism in the company and removes any form of private sector accountability. The extent of Zuckerberg's control is highly unusual even within companies with dual class rights. In 2012, on Facebook's listing on the NASDAQ stock exchange, ISS described the structure as "an autocratic model of governance".[114] In a decision of dubious legality in 2017, Zuckerberg sought to sell some of his shares but retain voting control in a move that was ended only after shareholder litigation ensued.[115] The autocrat tag proved prescient in the Cambridge Analytica scandal (Facebook users private data was harvested and used for political purposes without their consent) that engulfed Facebook and Zuckerburg in 2018, and which followed on from an earlier little-noticed similar scandal in 2009.[116] What became clear in the investigations that followed was that Facebook, despite its scale, influence and power, was a one person company. Zuckerberg's autocratic tendencies have been well documented.[117] Indeed, he even recognises his vast power himself.[118]

Google is similarly structured through a dual-class voting structure. Alphabet is the listed parent company of Google, the posterchild for tech avoidance of private sector accountability. Google was the first Tech company to list in 2004 with dual class shares designed to ensure founder

[113] Facebook's SEC ownership filings can be found here <https://investor.fb.com/financials/default.aspx> and F. Manjoo, 'How Mark Zuckerberg Became Too Big to Fail,' *The New York Times*, Nov 1 2018, <https://www.nytimes.com/2018/11/01/technology/mark-zuckerberg-facebook.html>

[114] M. De La Merced (2012) 'I.S.S. Adds to Criticism of Facebook's Governance,' *The New York Times,* 13 February 2012, <https://dealbook.nytimes.com/2012/02/13/i-s-s-adds-to-criticism-of-facebooks-governance/ >

[115] M. Isaac, 'Facebook Drops Stock Move That Would Have Solidified Zuckerberg's Control,' *The New York Times*, 22 September 2017,< https://www.nytimes.com/2017/09/22/technology/facebook-stock-reclassification-zuckerberg-control.html>

[116] N. Confessore, 'Cambridge Analytica and Facebook,' *The New York Times,* 4 April 2018, < https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html >and K. Shubber, 'Facebook Data Scandal: The Legal Questions,' The Financial Times, 21 March 2018, <https://www.ft.com/content/257d4598-2cb9-11e8-a34a-7e7563b0b0f4>

[117] B. Mezrich, *The Accidental Billionaires: The Founding of Facebook, A Tale of Sex, Money, Genius, and Betrayal* (2009) and S. Frenkel, N. Confessore, C. Kang, M. Rosenberg and J. Nicas, 'Delay, Deny and Deflect: How Facebook's Leaders Fought Through Crisis,' *The New York Times,* 14 November 2018 < https://www.nytimes.com/2018/11/14/technology/facebook-data-russia-election-racism.html?action=click&module=Top%20Stories&pgtype=Homepage > and K. McCarthy, 'Why does it cost 20 times as much to protect Mark Zuckerberg as Tim Cook?,' *The Register*, 6 February 2017 <https://www.theregister.co.uk/2017/02/06/why_does_it_cost_20_times_as_much_to_protect_mark_zuckerberg_as_tim_cook/>

[118]J. Wong, 'Zuckerberg's control of Facebook is near absolute – who will hold him accountable?' *The Guardian*, 21 November 2018, < https://www.theguardian.com/technology/2018/nov/21/zuckerbergs-control-of-facebook-is-near-absolute-who-will-hold-him-accountable >

control on listing. Up until that point such structures only existed in family media companies such as the New York Times and the Wall Street Journal and was justified on the basis of protecting editorial independence. Such structures were rare until 2004 and considered generally unsuitable and dangerous for shareholders, as they have been shown to exacerbate accountability problems and encourage value-destroying decision making.[119] Google justified the dual-share structure as protecting against takeovers and to allow focus on the long term, while making clear that Google was in reality comprised of its founders, and an investment in the company was an investment in them.[120] On listing, Larry Page and Sergey Brin controlled over 50% of the company's voting rights.[121] Page is now also the CEO and Brin the President of Alphabet.[122]

In 2014, having sold off enough shares to put their 50% majority control in danger they issued new C class shares with no votes so that they and other executives could continue to cash in their shares without losing control.[123] The dubious legality of doing this and potential litigation, that is, acting in their own interest instead of in the interests of the company, was avoided by the company underwriting any shareholder losses (estimated to be $500 million) caused by the new class structure.[124] As discussed above, attempts by Zuckerberg to do the same as the Cambridge Analytica scandal unfolded, failed. As with Facebook, Google has similarly continually been involved in controversy globally about data gathering and misuse, while more recently its employees have walked out in protest at its engagement with autocratic governments,[125] military

[119] R. Masulis, C. Wang, and F. Xie, 'Agency Problems at Dual-Class Companies.' *Journal of Finance*, ECGI - Finance Working Paper No. 209/2008.
[120] Google, 'Letter From the Founders,' *The New York Times*, 29 April 2004, <https://www.nytimes.com/2004/04/29/business/letter-from-the-founders.html >
[121] Alphabet's SEC filings can be found here https://abc.xyz/investor/other/additional-financial-information/ and Google id.
[122] <https://www.marketscreener.com/ALPHABET-24203373/?type_recherche=rapide&mots=alphabet>
[123] N. Summers (2014) Why Google Is Issuing a New Kind of Toothless Stock, Bloomberg, 3 April, https://www.bloomberg.com/news/articles/2014-04-03/why-google-is-issuing-c-shares-a-new-kind-of-powerless-stock accessed 17 March 2019.
[124] M. Liedtke (2015) Stock split could cost Google over $500 million, Psy Org, 2 April https://phys.org/news/2015-04-stock-google-million.html accessed 17 March 2019.
[125] K. Conger and D. Wakabayashi (2018) Google Employees Protest Secret Work on Censored Search Engine for China, The New York Times, 16 August https://www.nytimes.com/2018/08/16/technology/google-employees-protest-search-censored-china.html accessed 17 March 2019.

use AI[126] and its handling of sexual discrimination and harassment claims.[127] In 2018, one of Google's senior AI researchers resigned because of Google's breach of its own AI principles in assisting the Chinese government to develop project Dragonfly, a censorship and data gathering tool. In a letter to the US Senate Commerce Committee asking for more oversight of Google's activities he stated:

> Dragonfly is part of a broad pattern of unaccountable decision making across the tech industry. It has been made clear, both by word and by action, that the leadership at Google will be clamping down on the types of internal investigation that were necessary to bring Project Dragonfly to light.[128]

In 2018 Google removed its "Don't be Evil" motto as an ethical imperative at the centre of its code of conduct.[129]

To put those tech founder shareholding percentages in context, a big stake for our largest global institutional investors (pension insurance and investment funds) in a public listed company would be ownership of 7-8% of its shares. Aggressive hedge fund investors such as Elliot Management would buy such a 7-8% stake if they wanted to get board seats and influence a company's strategic direction.[130] In other words, a stake of that magnitude provides significant input into the control of a company. Benzos's 16.1% shareholding, while small by comparison to Zuckerberg and the others, is by normal standards a huge chunk of voting rights and gives him de facto control of the board. Similarly Son's 25.5% shareholding in Softbank confers de facto control. Softbank and Kalanick's 22% combined share of Uber, as well as their control of five board seats gives them similar levels of de facto control. A stake approaching 30% such as Elon Musks gives

[126] S. Shane and D. Wakabayashi (2018) 'The Business of War': Google Employees Protest Work for the Pentagon, The New York Times, 4 April, https://www.nytimes.com/2018/04/04/technology/google-letter-ceo-pentagon-project.html accessed 17 March 2019.
[127] D. Wakabayashi and K. Benner (2018) How Google Protected Andy Rubin, the 'Father of Android', The New York Times, 25 October, https://www.nytimes.com/2018/10/25/technology/google-sexual-harassment-andy-rubin.html accessed 17 March 2019.
[128] Dr. Jack Poulson (2018) Letter to Senate Commerce Committee, 24 September https://int.nyt.com/data/documenthelper/328-jack-poulson-dragonfly/87933ffa89dfa78d9007/optimized/full.pdf#page=1 accessed 17 March 2019.
[129] Its still in the code but reduced from an central instruction to a suggested reporting mechanism see K. Conger (2018) Google Removes 'Don't Be Evil' Clause From Its Code of Conduct, Gizmodo, 18 May https://gizmodo.com/google-removes-nearly-all-mentions-of-dont-be-evil-from-1826153393 accessed 17 March 2019.
[130] J. Wieczner (2017) Inside Elliott Management: How Paul Singer's Hedge Fund Always Wins, Fortune, 15 December, http://fortune.com/2017/12/07/elliott-management-hedge-fund-paul-singer/ accessed 17 March 2019.

him, along with his constitutional blocking rights, both de facto and legal control over who are the members of the board of Tesla and with it total strategic control of the company. The voting rights of Zuckerberg, Brin and Page are extraordinary and well beyond the voting control necessary to de facto control the board of their companies. In both Facebook and Google (Alphabet) they have both de facto and total legal control of all the organs of the company. The dubious attempts, successful in Google's case, to maintain more than 50% voting control (the trigger for full de jure voting control), while taking the economic benefits from sale of their shares, exhibits a bizarre level of unnecessary controlling paranoia and greed. In short, our AI tech leadership are particularly autocratic and dysfunctional.

Another significant governance factor is how interlinked the tech owners and the key AI companies are. Apart from Softbank, they are US Companies. Apart from IBM and Softbank they are located in Silicon Valley and Seattle. Musk and Klanick are connected to each other as they were co-founders of another company, Paypal. Page is a shareholder in Tesla and a close friend of Musk.[131] Zuckerberg and Kalanick are also friends.[132] A major tech investor, Peter Thiel, also links Amazon, Facebook, Tesla, Uber and Softbank. Thiel was the most significant early funder of Facebook and remains on its board. Thiel has shared business ventures with Softbank[133] and Jeff Benzos[134] and links Zuckerberg, Musk and Kalanick as he was also a Paypal founder. Son and Apple are linked through Apple's investment in Son's Vision fund and through Softbank's

[131]J. Koebler (2015) Elon Musk and Larry Page Have the World's Weirdest Friendship, Motherboard, 25 May, https://motherboard.vice.com/en_us/article/bmj4zz/elon-musk-and-larry-page-have-the-worlds-weirdest-friendship accessed 17 March 2019. There are other links between paypal and Google through Google's purchase of YouTube which was founded by former paypal executives. See https://en.wikipedia.org/wiki/Steve_Chen accessed 17 March 2019.

[132] P. Marinova, (2017) Uber CEO Travis Kalanick and Mark Zuckerberg Celebrate Friend's Birthday at 'Babes and Balls' Party, Fortune, 10 March, http://fortune.com/2017/03/10/uber-ceo-travis-kalanick-babes-balls/ accessed 17 March 2019.

[133] D. Bosa (2019) SoftBank is leading a $1 billion investment in logistics start-up Flexport, CNBC, 21 February, https://www.cnbc.com/2019/02/21/softbank-leads-1-billion-investment-in-logistics-start-up-flexport.html accessed 17 March 2019.

[134] A. Balakrisnan (2017) Silicon Valley bigwigs like Larry Ellison, Peter Thiel, and Sergey Brin are trying to end aging, CNBC, 27 March, https://www.cnbc.com/2017/03/27/anti-aging-technology-larry-ellison-peter-thiel-sergey-brin.html accessed 17 March 2019.

ownership of ARM, the chip designer for iphones.[135] Facebook and IBM have a marketing partnership.[136] IBM also has a crucial infrastructure partnership with Amazon.[137] Gates the founder and board member of Microsoft, and Benzos are business partners, Seattle neighbours and friends.[138] Microsoft became a major investor in Facebook in 2007 and the two companies also have joint advertising and infrastructure partnerships.[139] Google part owns Uber.[140] Benzos was also an early investor in Google and owns nearly 4% of Uber.[141] More explicitly, Microsoft, IBM, Amazon, Facebook and Google have an AI partnership aimed at shaping industry technical standards and public perception of AI Technology.[142] To an extent, geography seems to play a key part in those connections, with Silicon Valley having one group of connections and the Seattle pair another with Rometty alone in New York, but that said, clear business interest overlaps exist.

Here again, there is a specific group of humans behind the tech companies, shaping the agenda. All apart from Rometty are male. All apart from Son and Rometty are white males and all including Son were educated in the US and almost all operate autocratic governance models in their companies. They have little accountability unless they break criminal laws. Billionaire tech

[135] A. Massoudi, J. Fontanella-Khan and R. Waters (2016) SoftBank to acquire UK's Arm Holdings for £24.3bn, The Financial Times, 18 July, https://www.ft.com/content/235b1af4-4c7f-11e6-8172-e39ecd3b86fc and R. Mickle and L. Hoffman (2017) Apple Confirms $1 Billion Investment in SoftBank Vision Fund, The Wall Street Journal, 5 January, https://www.wsj.com/articles/apple-confirms-1-billion-investment-in-softbank-vision-fund-1483552846 accessed 17 March 2019.

[136] S. Lohr (2015) IBM and Facebook in Marketing Partnership, 6 May, https://bits.blogs.nytimes.com/2015/05/06/ibm-and-facebook-in-marketing-partnership/ accessed 17 March 2019.

[137] https://mainframedebate.com/2018/07/26/amazon-and-ibm-announce-linuxone-server-partnership/

[138]Mainframedabate (2018) Amazon and IBM Announce LinuxONE Server Partnership, Mainframedebate, 26 July, S. Best (2016) Bill Gates teams up with Richard Branson and Jeff Bezos to lead a $1 BILLION venture to tackle climate change, The Daily Mail, 12 December, https://www.dailymail.co.uk/sciencetech/article-4025044/Bill-Gates-teams-Richard-Branson-Jeff-Bezos-lead-1-BILLION-venture-tackle-climate-change.html and R. Williams (2017) The weirdly intertwined lives of Bill Gates and Jeff Bezos, inews, 28th July, https://inews.co.uk/news/technology/the-weirdly-intertwined-lives-of-bill-gates-and-jeff-bezos/ accessed 17 March 2019.

[139] B. Stone (2007) Microsoft to Pay $240 Million for Stake in Facebook, The New York Times, 25 October, https://www.nytimes.com/2007/10/25/technology/24cnd-facebook.html and D. Bach (2017) Microsoft, Facebook and Telxius complete the highest-capacity subsea cable to cross the Atlantic, Microsoft, 21 September, https://news.microsoft.com/features/microsoft-facebook-telxius-complete-highest-capacity-subsea-cable-cross-atlantic/ accessed 17 March 2019.

[140] Bullock, N. and Bradshaw, T (2019) Uber seeks $91.5bn valuation in this year's largest IPO, Financial Times, April 26th https://on.ft.com/2W5xwa4 accessed 29 April 2019.

[141] B. Stone. (2013) The Everything Store: Jeff Bezos and the Age of Amazon, Random House and French, S. (2018) All the companies in Jeff Bezos's empire, in one (large) chart, Marketwatch Jan 30, https://www.marketwatch.com/story/its-not-just-amazon-and-whole-foods-heres-jeff-bezos-enormous-empire-in-one-chart-2017-06-21 accessed 17 March 2019.

[142]A. Hern (2016) 'Partnership on AI' formed by Google, Facebook, Amazon, IBM and Microsoft, The Guardian, https://www.theguardian.com/technology/2016/sep/28/google-facebook-amazon-ibm-microsoft-partnership-on-ai-tech-firms accessed 17 March 2019.

autocrats who live, work and hang out together in small geographic locations brings with it an obvious and urgent concern of disproportionate influence and unrepresentative group think.

That would be the case even if the tech community did not have a particularly US determinist autarchist libertarian[143] deregulatory world view, as noted earlier. Thiel, for example, is the arch libertarian of the tech community and a supporter of President Donald Trump.[144] Similarly Musk and Klanick share those libertarian leanings and both accepted short-lived advisory roles in Trump's administration.[145] Son has similarly promoted Trump and his deregulatory agenda.[146] Tim Cook and Rometty both currently sit on Trump's Business Advisory Board.[147] Zuckerberg seems to share or at least value Thiel's views.[148] Page regularly laments government interference in Google's attempts to create a techno libertarian utopia.[149] Benzos has funded the Reason Foundation, a libertarian thinktank.[150] Tech billionaires generally also appear to have a very specific admiration of Ayn Rand's individualist libertarian work.[151]

That utopian techno autarchist libertarianism also matters beyond the important technological power of these tech leaders to shape agendas through their companies or through their

[143] Building on Emmerson and reaching its peak of individual and market freedom in R. LeFevre (1965) "Autarchy vs Anarchy" Rampart Journal of Individualist Thought Vol. 1, No. 4 (Winter) 30–49. See also P. Boorsook, (2000) Cyberselfish. a critical romp through the terribly libertarian culture of high tech, PublicAffairs.

[144] M. Broussard (2018) p 83 and D. Streitfeld and J. Williams (2017) Peter Thiel, Trump Adviser, Has a Backup Country: New Zealand, The New York Times, 25 January, https://www.nytimes.com/2017/01/25/technology/peter-thiel-new-zealand-citizenship.html accessed 17 March 2019.

[145] R. Reader (2017) Elon Musk and Bob Iger leave Trump's advisory council over Paris agreement exit, Fast Company, 6 January, https://www.fastcompany.com/4039550/elon-musk-leaves-trumps-advisory-council-over-paris-agreement-exit accessed 17 March 2019.

[146] M. Yamazaki (2017) SoftBank CEO Masayoshi Son: Trump's promised deregulation is going to be good for business, Business Insider, 8 February https://www.businessinsider.com/r-softbanks-son-says-trumps-promised-deregulation-will-help-his-business-2017-2?r=US&IR=T accessed 17 March 2019.

[147] A Newcomb, 2019) Tim Cook, Ginni Rometty Join Trump's Panel on A.I. and Workforce Automation, Fortune, 14 February, http://fortune.com/2019/02/13/tim-cook-ginni-rometty-join-president-trumps-panel-on-workforce-automation/ accessed 17 March 2019.

[148] M. Hiltzik (2016) Facebook won't dump billionaire Peter Thiel from its board because it cares about 'diversity, Los Angles Times, 21 October, https://www.latimes.com/business/hiltzik/la-fi-hiltzik-thiel-facebook-20161021-snap-story.html accessed 17 March 2019.

[149] A. Sankin (2013) tech Burning Man Pitched By Google CEO Larry Page, Huffington Post, 20 May, https://www.huffingtonpost.co.uk/2013/05/20/tech-burning-man_n_3308018.html and S. Gallagher (2013) Larry Page wants you to stop worrying and let him fix the world, Ars Technica, 20 May, https://arstechnica.com/information-technology/2013/05/larry-page-wants-you-to-stop-worrying-and-let-him-fix-the-world/ accessed 17 March 2019.

[150] M. Welch (2013) Amazon's Jeff Bezos Buys the Washington Post for $250 Million, 5 August, https://reason.com/blog/2013/08/05/amazons-jeff-bezos-buys-the-washington-p accessed 17 March 2019.

[151] J. Freedland (2017) The new age of Ayn Rand: how she won over Trump and Silicon Valley, The Guardian, 10 April, https://www.theguardian.com/books/2017/apr/10/new-age-ayn-rand-conquered-trump-white-house-silicon-valley accessed 17 March 2019, and Streeter, Thomas. 2003 "The Deep Romantic Chasm: Libertarianism, Neoliberalism, and the Computer Culture." In Critical Cultural Policy Studies: A Reader. Justin Lewis and TobyMiller (eds.), 161-171. Malden: Basil Blackwell.

group think. As already noted, reflecting a tech deterministic libertarian ethos, tech companies have an unusual relationship with the law. A key element of that is the emergence of the internet, which has allowed traditional and new businesses to deliver their products and services in radically new ways and to shelter behind the tech libertarian utopian view that this is all new and therefore existing law does not apply. That is the point of the Declaration of the Independence of Cyberspace in 1996 considered in Section 2. Tesla insists it is not a normal car company subject to normal safety considerations.[152] Facebook, Apple and Google are not publishers or broadcasters so a range of laws to protect children or pay artists should not apply.[153] Amazon and Tesla push the concept of employer to the limit in terms of working conditions.[154] Uber claims it is not a taxi company subject to employment law or taxi regulation.[155] As further discussed in Section 5, the UK and US governments have, until recently, been deregulating the sector and allowing it to self-regulate.

While these companies can work together to shape technical AI standards and public perception of AI, they also engage in significant lobbying to shape or remove the law. Indeed, while all have significant lobbying engagement with government, overall Google is the second most active company across all industry sectors lobbying EU institutions, with Microsoft number eight, Facebook number twelve and IBM number sixteen.[156] In the US, Google, Amazon, Facebook,

[152] T. Simonite (2016) Tesla's Dubious Claims About Autopilot's Safety Record, Technology Review, 6 July, https://www.technologyreview.com/s/601849/teslas-dubious-claims-about-autopilots-safety-record/ accessed 17 March 2019.

[153] *Metropolitan International Schools Ltd v (1) Designtechnica Corporation (2) Google UK Ltd & (3) Google Inc* [2009] EWHC 1765 (QB), S. Levin (2018) Is Facebook a publisher? In public it says no, but in court it says yes, The Guardian, 3 July, https://www.theguardian.com/technology/2018/jul/02/facebook-mark-zuckerberg-platform-publisher-lawsuit, The Telegraph, (2010) What is the Google Books Settlement?, The Daily Telegraph, 5 February, https://www.telegraph.co.uk/technology/google/7164237/What-is-the-Google-Books-Settlement.html and R. Williams (2019) NSPCC: Make tech firm directors legally and personally responsible for child safety, 12 February, https://inews.co.uk/news/technology/nspcc-make-tech-firm-directors-legally-and-personally-responsible-for-child-safety/ accessed 17 March 2019.

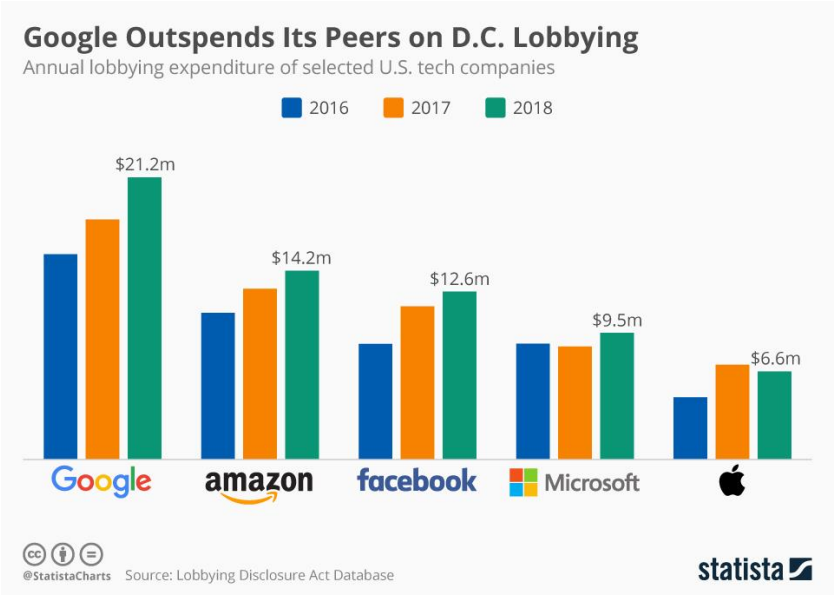
[154]G. Hall, (2017) Elon Musk disputes union claims, promises rollercoaster, 'crazy good' fun at Tesla's Fremont factory, Silicon Valley Business Journal, 27 Febuary, https://www.bizjournals.com/sanjose/news/2017/02/27/elon-musk-tesla-fremont-factory-uaw-tsla.html and I. Johnson (2015) Amazon: Devastating expose accuses internet retailer of oppressive and callous attitude to staff, The Independent, 17 August, https://www.independent.co.uk/news/business/news/amazon-devastating-expose-accuses-internet-retailer-of-oppressive-and-callous-attitude-to-staff-10458159.html accessed 17 March 2019.

[155] S. O'Connor, (2017) Uber loses appeal in UK employment case, The Financial Times, 10 November, https://www.ft.com/content/84de88bc-c5ee-11e7-a1d2-6786f39ef675 and S. Butler (2017) Uber stripped of London licence due to lack of corporate responsibility, The Guardian, 23 September, https://www.theguardian.com/technology/2017/sep/22/uber-licence-transport-for-london-tfl accessed 17 March 2019.

[156] Data available at https://www.integritywatch.eu/lobbyist.html accessed 17 March 2019.

Apple and Microsoft were five of the biggest spenders on lobbying in 2018 with Google, as Table 1 illustrates, significantly outspending its peers.[157]

Table 1. Leading US Technology Company Lobbying Expenditure 2016-18[158]



As well as core technology issues, the tech companies' lobbying also focused on tax law reform, which has become a feature of their autarchic activities globally, not just in terms of using aggressive multi-jurisdictional corporate law strategies to reduce tax paid, but also to fight changes to the law that might impact on tax or change the law to reduce their taxation.[159] The EU estimated in 2017 that international tech businesses paid less than half the tax of traditional companies.[160]

This small group of companies dominate global private sector AI development.[161] Their products are not neutral or benign but are developed by mostly white male AI technical staff with

[157] Data available here https://www.senate.gov/legislative/Public_Disclosure/LDA_reports.htm
[158] Source: Statista https://www.statista.com/chart/10393/lobbying-expenditure-of-tech-companies/ accessed 2 May 2019.
[159] E. Yeomans (2018) tech giants build army of lobbyists to fight off tougher regulations, The Sunday Times, 23 October, https://www.thetimes.co.uk/edition/news/tech-giants-build-army-of-lobbyists-to-fight-off-tougher-regulations-2qq39v0vj accessed 17 March 2019.
[160]European Commission (2017) A Fair and Efficient Tax System in the European Union for the Digital Single Market, COM(2017) 547, https://ec.europa.eu/taxation_customs/sites/taxation/files/1_en_act_part1_v10_en.pdf and A. Hill, M. Khan, R. Waters (2018) The global hunt to tax Big tech, The Financial Times, 2 November, https://www.ft.com/content/79b56392-dde5-11e8-8f50-cbae5495d92b accessed 17 March 2019.
[161]T. Ball (2017) Top 10 players in Artificial Intelligence, CBR, 14 August, https://www.cbronline.com/internet-of-things/cognitive-computing/top-10-players-artificial-intelligence-ai/ and C. Forrest (2016) The five upstarts that are leading the AI and machine learning revolution, ZDNET, 1 December,

world views that can consciously or unconsciously exclude, kill, lock up and impoverish millions of humans given their global technological reach. Their technical staff work for mostly white male autocrats with similar educational backgrounds and world view, ideally seeking an autarchist libertarian utopia free from law and taxation. This is not a group that will provide a solution to AI dysfunctionality and urgent public interest considerations. Indeed, as this article has argued, a contextualised understanding of this technology is necessary to see that these dysfunctional humans are the reason for the existence of dysfunctional AI and that AI is not a superior deterministic technological force. A solution will not lie with the industry itself.

CONCLUSION

In many ways, the biggest challenge lies not in identifying the problematic issues with AI but rather the state of the public governance response. When it comes to technological change, governments have historically been wary. Famously, Elizabeth the First turned down a grant of a patent for a knitting machine to William Lee, remarking:

> Thou aimest high, Master Lee. Consider thou what the invention could do to my poor subjects. It would assuredly bring to them ruin by depriving them of employment, thus making them beggars.[162]

The Hapsburg, Francis the First, banned railways for fear of its potential to fuel threatening social change. Indeed, so deep was this fear that when he did allow a railway to be built the train had to be pulled by horses and its lines were deliberately irregular so that it could never be used by a steam engine. Tzar Nicholas the First banned cotton mills and iron foundries and the Sultans of the Ottoman Empire banned printing.[163] More recently, China at first similarly feared and resisted technological change because of its potential to trigger social change. But as it became clear that the state could use technology to Xi Jinping's authoritarian ends, China has become a major player in

https://www.zdnet.com/article/the-fiveupstarts-that-are-leading-the-ai-and-machine-learning-revolution/ accessed 17 March 2019.

[162] D. Acemoglu, and J. Robinson, *Why nations fail,* (2012). pp.182 and 212-219.

[163] Acemoglu, id.

technology generally and AI particularly.[164] The lesson from China's AI engagement for democracies is that AI in the wrong hands can be very dangerous. Unfortunately, until recently, the UK and US governments have not engaged in shaping or controlling AI for the democratic public good, instead leaving it to be shaped and controlled by a small group of tech companies, in their own interests.

The moment has come to recognise AI as a problematic social construction produced by human decisions in the tech industry and to regulate in the public not private interest. However, in both the UK and US, governance dysfunction abounds as deregulation and self-regulation have formed key parts of nationalist AI strategies. For the past few years, as part of a post-Brexit nationalist tech-friendly strategy, the UK has been quietly deregulating to facilitate AI development.[165] The under secretary of state Claire Perry stated in 2015, "I believe we have one of the most welcoming regulatory environments for development of this technology anywhere in the world."[166] The UK has specifically allowed tech companies to take the lead on its independent post-Brexit Industrial Digitisation Strategy. Published in 2017, and led by the CEO of Siemens, this strategy advocates speeding up and increasing the adoption of automation and AI and warns:

> other countries are stealing a march on the UK. There are coherent government strategies in place in most developed countries, for example in Germany (Industrie 4.0), China (Made in China 2025), and the USA (America Makes). So, the UK needs to act quickly if it is to harness the potential of this agenda.[167]

A second UK AI-specific report co-authored by Jérôme Pesenti, Vice President of Artificial Intelligence at Facebook, continues the urgent-boosterish competitive theme.[168] There is little

[164] R.Veugelers, 'The challenge of China's rise as a science and technology powerhouse,' (2017) *Bruegel Policy Contribution*, 19, < http://bruegel.org/wp-content/uploads/2017/07/PC-19-2017.pdf > and I. Burrows, 'China's research in artificial intelligence 'far outranks' Huawei threat,' *ABC News*, 2 February 2019, <https://www.abc.net.au/news/2019-02-02/china-research-artificial-intelligence-bigger-threat-than-huawei/10685420>

[165] A. Perkins, 'Government to review law before self-driving cars arrive on UK roads,' *The Guardian*, 6 March 2018, <https://www.theguardian.com/technology/2018/mar/06/self-driving-cars-in-uk-riding-on-legal-review>

[166] Department of Transport, *The Pathway to Driverless Cars*, (2015) 5 <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/401562/pathway-driverless-cars-summary.pdf >

[167]Department for Business, Energy & Industrial Strategy (DBEIS), *Made Smarter Review*, (2017) 9. <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/655570/20171027_MadeSmarter_FINAL_DIGITAL.pdf>

[168] Department for Digital, Culture, Media & Sport and DBEIS

concern in either report for the issues raised in this article. As a result, the UK is now the centre of Europe's deregulated AI startup boom.[169]

In the US, a similar welcoming legislative approach has been taken to tech companies since the advent of the commercial internet. Section 230 of the Communications Decency Act of 1996 was introduced to protect the nascent US internet companies by giving them broad legal immunity for content posted on their sites. The overall effect has been more problematic than originally foreseen in disadvantaging traditional publishing, facilitating extreme content and removing legal and moral responsibility for content from powerful no longer nascent US companies such as Facebook and Google.[170] More recently, President Trump has, within the technology sector, through a combination of his appointees to the Federal Communications Commission and executive orders,[171] pursued a very successful deregulatory agenda ending net neutrality and generally allowing industry self-regulation.[172]

However, the Facebook Cambridge Analytica scandal and its connection to Russian interference in the 2016 US Presidential election, as well as general privacy and monopoly concerns have begun changing the public interest response.[173] Across the EU, the General Data Protection Regulation (EU) 2016/679 (GDPR) is in place to deal with the tech companies' misuse of information and the European Commission has produced ethical guidelines for AI as well as putting in place a policy framework to encourage investment in "trustworthy" AI.[174] In 2017 the UK

Growing The Artificial Intelligence Industry in the UK, (2017) <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/652097/Growing_the_artificial_intelligence_industry_in_the_UK.pdf>

[169] MMC, op. cit. n.91, Part 3

[170] A. Selyukh, 'Section 230: A Key Legal Shield For Facebook, Google Is About To Change,' *NPR*, 21 March 2018, <https://www.npr.org/sections/alltechconsidered/2018/03/21/591622450/section-230-a-key-legal-shield-for-facebook-google-is-about-to-change?t=1549713961283>

[171] W. Knight, 'Trump has a plan to keep America first in artificial intelligence,' (2019) *MIT Technology Review,* 10 <https://www.technologyreview.com/s/612926/trump-will-sign-an-executive-order-to-put-america-first-in-artificial-intelligence/>

[172] B. Davis-Noll and A. Dawson, 'Deregulation Run Amok,' (2018) *Institute for Policy Integrity*, <https://policyintegrity.org/files/publications/Deregulation_Run_Amok_Report.pdf>

[173] C.Cath, S.Wachter, B.Mittelstadt, M.Taddeo, and L.Floridi, 'Artificial Intelligence and the 'Good Society'' (2017) *Sci. Eng. Ethics*, 1–24.

[174] European Commission, *Ethics Guidelines for Trustworthy AI.* (2019) <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> and European Commission High-Level Expert Group on Artificial Intelligence, 'Policy and investment recommendations for trustworthy Artificial Intelligence' (2019) <https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence>

Government set up the Centre for Data Ethics and Innovation to advise on data and AI regulatory issues.[175] The report of the UK Digital Competition Expert Panel in March 2019 recommended a regulatory unit to examine the anti-competitive behaviour of the large tech companies.[176]

Individual states have begun introducing digital tech company taxes, while a global tech tax compact is back on the agenda.[177] In the US, individual states such as California have passed privacy laws aimed at protecting people from tech company misuse.[178] The US Congress has begun oversight hearings on the behaviour of the tech companies with regard to data misuse, while the US Department of Justice is investigating the big tech companies on a wide range of issues from data privacy to anti-competitive behaviour.[179] Breaking up the big US tech companies has also become a campaign issue in the Democrat party presidential primary elections.[180]

The moment has also come to specifically regulate AI. As has been argued here, AI creates significant private versus public interest flashpoints. Solutions that are in the public interest and that are human focused are possible. Treating the tech industry as the pharmaceutical industry is treated would be one way to solve certain key issues. In such a regulatory model, AI products would be submitted for testing and licencing to an agency (say, the "Artificial Intelligence Licencing Agency" (AILA)) in a model that would be a mix of the UK National Institute for Health and Care Excellence and the UK's Human Fertilisation and Embryology Authority, requiring a mixture of technical, ethical, legal and economic expertise. In overall terms, the aim of AILA would be to address the flashpoints identified within this article. First, ensuring the public interest is served in situations of genuinely novel AI decision-making rather than solely the private interest of the developer or their client. Second, ensuring the law is inserted within the AI design phase at an early

[175] <https://www.gov.uk/government/groups/centre-for-data-ethics-and-innovation-cdei>
[176] Digital Competition Expert Panel, 'Unlocking digital competition,' (2019) *HM Treasury*, <https://www.gov.uk/government/collections/digital-competition-expert-panel>
[177] V. Dendrinou and A. Weber, 'EU Gives Up on tech Tax Plans,' *Bloomberg*, 12 March 2019, <https://www.bloomberg.com/news/articles/2019-03-12/eu-gives-up-on-tech-tax-plans-as-countries-vow-to-go-it-alone>
[178] A. Adler, 'California Passes Strict Internet Privacy Law,' *NPR*, 29 June 2018, <https://www.npr.org/2018/06/29/624336039/california-passes-strict-internet-privacy-law-with-implications-for-the-country>
[179] A. Selyukh, 'Attorneys General Zoom In On tech Privacy And Power,' *NPR*, 25 September 2018, <https://www.npr.org/2018/09/25/651472693/attorneys-general-zoom-in-on-tech-privacy-and-power>
[180] A. Shahani, 'Democratic Candidates Target tech Giants,' *NPR*, 11 March 2019. <https://www.npr.org/2019/03/11/702102576/democratic-candidates-target-tech-giants-who-are-major-party-donors>

stage and is evaluated throughout. Third, ensuring that high levels of statistical integrity are maintained within the design process. Four, AILA would ensure bias abating design and testing was implemented throughout the AI development process, ensuring that those designing and/or externally evaluating the AI were a diverse group in terms of background, gender and skills. Five, the sphere within which the AI was being deployed would be considered. High risk sectors such as criminal, employment, medical and the public sector generally might be inappropriate/unlawful for AI applications, except in exception circumstances. In evaluating black box AI, AILA would work within a general principle that if it is not explainable, it is not deployable. Exceptions would be allowed where there was a significant public interest in its deployment such as medical scanning. Six, the agency would also develop public interest AI policy and recommendations on dealing with complexity/superiority atrophy situations where there might be significant public harm.

Within the scope of this article, the possibility of large scale displacement of human employees by combinations of AI and automation, as it moves into cognative employment roles, has not been considered.[181] However, AILA could also consider the social and employment displacement effect of the AI implementation and cost the licence accordingly. Where there was no displacement effect the lowest cost licence would be given. Where there was a high risk of human displacement due to its implementation (and that could be extraterritorial displacement as well) the costing would be set around or above the 15% cost saving threshold that is estimated to trigger mass automation.[182] The cost equation would then move back to employing humans. In the most extreme displacement situations such as replacing workers in a low skill, high unemployment area the licence could be refused.

Our current version of UK and US financial capitalism does not deal well with urgent public interest issues such as AI when the dominant economic and legal paradigm focuses almost

[181] D. Autor, 'Why Are There Still So Many Jobs? The History and Future of Workplace Automation.' (2015) 29 *Journal of Economic Perspectives*, 3-30.
[182] P. Davidson, 'More robots coming to U.S. factories,' *USA Today*, 10 February 2015, <https://eu.usatoday.com/story/money/2015/02/09/bcg-report-on-factory-robots/23143259/>

exclusively on shareholder' private interests.[183] And while a critique of the totality of financial market capitalism is beyond this work[184] we can observe that tech autocratic ownership is uniquely antagonistic to both stakeholder and efficient market hypotheses conceptions of the corporations. Its justification rests solely on arguments that it is necessary to support innovation. However, while evidence exists to show this has some effect, it quickly dissipates and becomes negative.[185] Remarkably autocratic tech ownership serves neither stakeholder communities or efficiency justifications but rather the interests of tech owners and the finance industry: a sort of techno-financial coup has occurred. Allowing tech founder-companies to list as dual class entities has allowed a small unrepresentative group access to vast public market funding, which in turn has allowed them to wield disproportionate power over the shape of technology and in turn our democracies. Dealing with it requires creating a public interest technology regulator (PITR): placing a public interest requirement within board fiduciary obligations; removing dual class rights and other enhanced voting mechanisms from tech companies; providing PITR with rights to sit on the board, a veto over tech listings, continued listing and board personnel; nationalising and/or breaking them up, as Senator Elizabeth Warren has suggested, would move us towards a solution to that techno-financial coup.[186]

In the determinist autarchic libertarian technological context, law is subject to a fair amount of false news. Regulation is treated with horror and law as an enabling force is rarely mentioned. However, law that prohibits discrimination is enabling of minorities and women. Quite apart from the moral dimension of inclusion and fairness, without their participation in society creativity and economic progress is stifled. Employment law protects the economically vulnerable from the economically powerful. The ability to challenge in court the decisions of the state and those in

[183] Greenfield, K. (2007) Saving the World With Corporate Law? Boston College Law School Research Paper No. 130. SSRN: https://ssrn.com/abstract=978242.10 and Dignam, A. (2013) The Future of Shareholder Democracy in the Shadow of the Financial Crisis, Seattle University Law Review, No. 36:639.
[184] See for example Stout, L. (2012) The Shareholder Value Myth, Berrett-Koehler and Blair, M. and Stout, L, (1999) A Team Production Theory of Corporate Law. Virginia Law Review, Vol. 85, No. 2, pp. 248-328.
[185] Baran, L. Forst, A. and Via, M. (2018) Dual Class Share Structure and Innovation, SSRN: https://ssrn.com/abstract=3183517

[186] BBC News (2019) Elizabeth Warren vows to break up tech giants if elected in 2020, BBC, 9 March, https://www.bbc.co.uk/news/world-us-canada-47509945 accessed 17 March 2019.

positions of private power, forms a key part of the state's democratic legitimacy and the power of the citizen. Privacy laws protect key aspects of human dignity upon which human freedom and creativity is based. Progressive taxation is wealth redistributing. Laws and regulatory actions based on principles of democratic sovereignty, justice and fairness are not negative aspects of our society. Technology is not by its nature placed outside that lawful sphere and those who implement technological tools in a democracy, that have or can cause public harm, are subject to public interest regulation. These are not negative actions of the state but rather a positive protective force for creative, inclusive, human development in the public interest, not the narrow private interests of the very, very few.