

ALMOST HUMAN: LAW AND HUMAN AGENCY  
IN THE TIME OF ARTIFICIAL INTELLIGENCE

ANDREW MURRAY

Sixth Annual  
T.M.C. Asser Lecture



ASSER PRESS



ALMOST HUMAN: LAW AND HUMAN AGENCY  
IN THE TIME OF ARTIFICIAL INTELLIGENCE



ALMOST HUMAN:  
LAW AND HUMAN AGENCY  
IN THE TIME OF ARTIFICIAL  
INTELLIGENCE

*by*

ANDREW MURRAY



ASSER PRESS

Published by T.M.C. Asser Press  
P.O. Box 30461, 2500 GL The Hague, The Netherlands  
www.asserpress.nl

*The Annual T.M.C. Asser lecture has been established in honour of the Dutch jurist and Nobel Peace Prize Laureate, Tobias Michael Carel Asser (Amsterdam, 28 April 1838 – The Hague, 29 July 1913), and his significant contributions to the development of public and private international law. The Annual Lecture builds on his vision and mission, it invites distinguished international scholars to take inspiration from Asser's idea of cultivating trust and respect through law and legal institutions, and to examine what it could mean in their area of expertise today. It is the T.M.C. Asser Instituut's flagship lecture and its date commemorates the foundation of the Institute in December 1965.*

*The T.M.C. Asser Instituut is grateful for the support of the Municipality of The Hague.*

Series editor: Janne E. Nijman

ISBN 978 90 6704 366 3

All rights reserved.

© 2021, Andrew Murray and T.M.C. Asser Press, The Hague, The Netherlands

No part of the material protected by the copyright notices may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without written permission from the copyright owner.

PRINTED IN THE NETHERLANDS

## FOREWORD

### *Intelligence without a Conscience? A Plea for Regulation of the Digital World*

The Sixth Annual Asser Lecture was unlike any other. Part of the tradition of the Asser lecture is to meet for a late afternoon of critical reflection on the state of international or European law, at the premises of the Peace Palace, built in the days of Tobias Asser himself.

In the midst of the Covid-19 pandemic, on November 26<sup>th</sup>, 2020, our speaker was in London, our guests were at home or in their offices, and we, the Asser team, were in a studio in The Hague. No drinks and buzzing in-person conversations about the lecture afterwards. On the other hand, we were delighted so many people could join. Thanks to technology, hundreds of people from around the world attended the lecture virtually.

While the 2020 Asser Lecture was different, some things were the same. This year we again organised the lecture in the context of our research programme, ‘International and European law as a source of trust in a hyperconnected world’ (2016-2020). It seemed only natural to examine the implications of hyperconnectivity, datafication and algorithms for law and governance at the end of 2020 – a year severely marked by the Covid-19 pandemic increased human dependency on digital connectivity and on new technologies such as Artificial Intelligence (AI). These technologies do however not merely enable us to forge human connections around the world, they have also come to feed distrust among citizens.

Only a week prior to the lecture, we saw a lawyer of the then US President Donald Trump create suspicion about voting machines being hacked by ‘a secret algorithm’ and stealing millions of votes. This adds up to social media in which citizens roam around algorithm-

governed echo chambers, with rising levels of distrust towards each other and towards public institutions. Increasingly, algorithms are shaping or even taking over human decision-making. Trained with human data, artificial intelligence exacerbates human biases often with harmful consequences. Urgent questions emerge: when are algorithms trustworthy? And, when is trust in artificial intelligence misplaced? To put it in the words of Cambridge philosopher Onora O’Neill, who delivered the second Asser Lecture.<sup>1</sup>

In short, we chose the topic for the 2020 Asser lecture before the Covid-19 crisis had really started. Over the past few years, questions of law and tech have become increasingly important in our research. In September 2020, for example, we were able to launch a new interdisciplinary research project: Designing International Law and Ethics into Military Artificial Intelligence (DILEMA). Led by Asser Senior Researcher Dr Berenice Boutin, and supported by the NWO [Nederlandse Organisatie voor Wetenschappelijk Onderzoek - Dutch Research Council], the project explores how to safeguard human agency in military applications of artificial intelligence and it seeks to ensure compliance with international law and accountability by design. If we allow for these military applications to be used – I deliberately use *if* since I am personally much persuaded by arguments made for example by leading international lawyer Mary Ellen O’Connell that autonomous decision-making, when non-human agency decides over human life, is inherently in violation of the humanity principle<sup>2</sup> – then

<sup>1</sup> Onora O’Neill, *Accountable Institutions, Trustworthy Cultures*, the Second Annual T.M.C. Asser Lecture (The Hague: Asser Press, 2017).

<sup>2</sup> Cf Mary Ellen O’Connell, *Banning Autonomous Killing*, in *The American Way of Bombing. Changing Ethical and Legal Norms from Flying Fortresses to Drones* 224 (M. Evangelista and H. Shue eds. Cornell University Press, 2014) 236 ‘Resort to weapons has always been accompanied with some legal and moral restraint, including the complete ban on certain types of weapons. In the near future, robotic weapons are expected to be available with programs able to select and destroy targets without a human operator in the loop. Such a development would conflict with the historical, legal, and moral understanding that killing should be based on a good-faith understanding of real necessity and carried out by someone who may be held accountable for a wrong decision. Even if a computer could be so programmed, it is imperative that human beings not give up sovereignty over these vital aspects of what it is to be human: to have a conscience and to be subject to accountability. Too much of our current system of community and personhood are based on these two factors

these applications should at a minimum be in full conformity with international humanitarian law and international human rights law. The concept of “meaningful human control” has caught on in discussions on autonomous weapons systems to uphold human agency in warfare, and therewith responsibility, answerability, and enforceability – in brief it demands accountability that is meaningful and goes beyond mere transparency of unsorted information and data.<sup>3</sup>

Artificial intelligence challenges law in a fundamental way. Law is a social construct, it presumes and is a product of human thought, human experience, and human interaction; in short: law is a product of human agency. Humans interpret the law, they apply the law. Humans are able and free to respect the law when acting or making a decision. What gets lost when human agency (human autonomy) disappears and is supported or even replaced by autonomous decision-making? What are the implications of algorithmic decision-making for the rule of law, for democracy and human rights? In short, how to approach the governance of AI technologies?

With these poignant questions, we turned to a most eminent scholar in the field who has not shied away from grappling with these fundamental questions ever since he started to work on information technology law and the regulation of cyberspace. Andrew Murray, Professor of Law at the London School of Economics and Political Sciences, has been one of the first global experts who argued for the protection and promotion of human rights within the digital environment. As early as 15 years ago, in *Human Rights in the Digital Age*, a book co-edited with Mathias Klang, he examined how digital technologies impact the enjoyment of human rights. He has written the leading textbook *Information Technology Law: The Law and Society* and has recently co-written the book *Rethinking the Jurisprudence of Cyberspace* with Chris Reed. In *Regulating AI and Machine Learning: Setting the Regulatory Agenda*, co-written with Julia Black and published

to risk their elimination. This point is all the stronger when we realize the risk is being promoted for the sake of creating new means of killing.’

<sup>3</sup> Cf Onora O’Neill on transparency, Onora O’Neill, *A Question of Trust* (Cambridge: Cambridge University Press, 2002) 63-79.

in the European Journal of Law and Technology in 2019. Andrew is Director of the LSE Law, Technology and Society group and also leads the AI, Ethics and Governance subgroup. In 2018/19 Professor Murray was the specialist advisor to the House of Lords Communications Committee inquiry “Regulating in a Digital World”. In other words, a true forerunner in this area of law and policy development.

The written version of Andrew Murray’s lecture in front of you, is an in-depth examination of the implications of the structural changes caused by datafication for human agency and for the Rule of Law. It is highly relevant for lawyers and policy makers who are confronted with the technological revolution and the enormous responsibility that the need to regulate evokes.

Murray asks whether we are ‘developing the right approach to the regulation of AI and Machine Learning?’ These technologies ‘have the capacity to undermine our systems of communication, politics, media, and culture, but above all else our system of law and the rule of law [...] because at a fundamental level the use of AI and Machine Learning to supplement, assist, or in time replace, human decision-making is a change to human autonomy, and through that to thought, society and ultimately law as a product of these.’<sup>4</sup> Mind you, these technologies change human autonomy, which together with human conscience, defines for many what makes us human.<sup>5</sup> ‘Law presupposes its application by the human conscience, and without the latter the very existence of the former becomes unimaginable.’<sup>6</sup>

First, Murray turns against the current practice to develop field specific and overly ethical frameworks and he argues for *legal* frameworks, which regulate AI at a global level and demand compliance with in-

<sup>4</sup> Andrew Murray, *Almost Human: Law and Human Agency in the Time of Artificial Intelligence*, the Sixth Annual T.M.C. Asser Lecture (The Hague: Asser Press, 2021) 35.

<sup>5</sup> Nijman, J.E. (2021). *Ius gentium et naturae: The Human Conscience and Early Modern International Law*. In P. Slotte & J. Haskell (Eds.), *Christianity and International Law: An Introduction (Law and Christianity Series)* (Cambridge: Cambridge University Press, 2021) 153-176. doi:10.1017/9781108565646.008

<sup>6</sup> *Ibidem*.

ternational and human rights law. He is very explicit: we need *new* laws and regulation at a European and global level that is tailored specifically to AI and move quickly beyond the over-reliance on soft governance of AI. Secondly, he warns against taking a risk-based approach to regulation for it does not take the fundamental problem of the input effect on human autonomy seriously – Laws no and Siri yes? The input effect on human autonomy is a very high risk for humans and their societies as it redefines, undermines and destroys what defines humans. The currently popular risk-approach does not capture or avert this danger. And so Murray calls for global leadership and a global institutional response, this is the time to define how corporations and governments will use AI and Machine Learning:

‘The challenge of AI Regulation and Governance is a global one – just as we have the International Telecommunication Union and the International Civil Aviation Organization, we must have an International Office for AI, tasked with the development of a positive, international, legal framework for the development and deployment of AI. This is urgent. For in twenty years the technology will be ambient, and we will have missed our chance at meaningful, modern, positive, regulation of AI.’<sup>7</sup>

Andrew’s wonderful piece of scholarship on how law has to rule AI-driven action and decision-making makes us realise what is at stake today: our democracies, human autonomy, the Rule of Law, human rights, our basic public values and ultimately also law itself, which is not inherently good but may help us order our world and to make it more safe, equal and just. Similarly, AI and Machine Learning can bring both great benefits and great harm to humanity. AI technologies can help us in the face of climate crisis and may lead us to what James Lovelock has called the *Novacene: The Coming Age of Hyper-Intelligence* (2019).<sup>8</sup>

<sup>7</sup> Andrew Murray, *Almost Human: Law and Human Agency in the Time of Artificial Intelligence*, the Sixth Annual T.M.C. Asser Lecture (The Hague: Asser Press, 2021) 42.

<sup>8</sup> James Lovelock, *Novacene: The Coming Age of Hyperintelligence* (MIT Press 2019).

## FOREWORD

Law, as Murray shows, is crucial to the course we take. An international and human rights law approach to the design, development and deployment of AI systems may assist us in assuring these systems fulfil their potential for good.

In the spirit of the Annual T.M.C. Asser Lecture tradition, Andrew Murray has put ‘his finger on the pulse of his time’ and has pointed to the need for the development of international law to guide and constrain the design, development, and deployment of AI and Machine Learning systems. *Au travail!*

PROF DR JANNE E. NIJMAN

*Chair of the Executive Board and Academic Director  
of the T.M.C. Asser Instituut, The Hague*

# ALMOST HUMAN: LAW AND HUMAN AGENCY IN THE TIME OF ARTIFICIAL INTELLIGENCE

ANDREW MURRAY

## I. INTRODUCTION

The contribution of Tobias Asser to contemporary jurisprudence is of particular significance. We are shaped by our social environment. As jurists, as researchers, and as wider members of a community of educators, scholars, and practicing lawyers our worldview is both expanded and constrained by what we see, read, and understand. All social sciences, and law is a social science, must be understood through a societal lens. In short law demands context. As Philip Selznick observes, ‘we cannot separate positive law from debatable principles of fairness, truth-finding, and morality. The more attention we give to basic legal principles such as due process of law, contractual obligation, or fiduciary duty, the harder it is to draw a bright line between distinctively legal norms and other social, intellectual, or even theological standards, such as moral equality, sexual mores, or parental authority and responsibility.’<sup>1</sup>

Asser would have recognised this assertion. As Anne Orford set out in the Fifth Annual T.M.C. Asser Lecture, Tobias Asser was immersed in debates about the relation of law and political economy.<sup>2</sup> Fuelled by the ‘unrest and revolution caused by market liberalization, industrialization and urbanization’ of the 19<sup>th</sup> Century ‘workers and middle-class radicals [took] to the barricades in European capitals, [and launched] campaigns for universal suffrage and transnational repub-

<sup>1</sup> Philip Selznick, ‘Law in Context’ Revisited’ (2003) 30 *Journal of Law and Society* 177, 179.

<sup>2</sup> Anne Orford, *International Law and the Social Question* (T.M.C. Asser Press, 2020), 1.

lican movements.”<sup>3</sup> Asser ‘saw it as his vocation to study the objectives of the law of commerce within its wider social context.’<sup>4</sup> Today Asser, trained as he was in ‘the technical intricacies of political economy’,<sup>5</sup> would recognise the equally disruptive effects of technology on our social and societal institutions.

Recent years have seen an increasing sense of both personal and societal isolation. Lately this has been most obviously driven by the covid-19 pandemic which began in Spring 2020 and which at the time of writing is ongoing. The literal requirement that we be socially distant makes our social isolationism apparent. However, the now clearly apparent current changes in our world were embedded long before covid-19 emerged. The recent changes to our lives may seem to be the result of the global health emergency, but the underlying source is not to be found in a public health crisis, rather it is symptomatic of a deeper societal change driven by developing technological capacities and the tantalising opportunities offered by digitisation.

When covid emerged and developed into a pandemic we were able to continue our lives, empowered by this technological capacity-building which had been going on for many years in an almost invisible fashion. Technologies that had previously been passive influences on our lives – video conferencing which previously had been reserved for the business traveller who just couldn’t physically get to the venue for a meeting or conference, or for everyday social interactions with friends and family, or file sharing which had mostly been used internally to share documents among teams or for the convenience of ‘backing up’ to the cloud now took on a whole new meaning. We began to live more in the virtual moment and less in the physical moment. Universities re-oriented their teaching offering around digital videoconference facilities; gyms became virtual with Zoom yoga and spin classes emerging, and facilities like Google Classroom became central to primary and secondary education. Virtual Learning Environments, or VLEs, for so long relegated to a support for in-person learning

<sup>3</sup> Orford, *International Law and the Social Question*, 1.

<sup>4</sup> Orford, *International Law and the Social Question*, 2.

<sup>5</sup> Orford, *International Law and the Social Question*, 2.

became independent learning environments and libraries adapted to become resource managers for digital learning tools rather than physical repositories of knowledge. The same happened outside academia. Tools like Microsoft Teams and Slack became work hubs connecting remote teams of employees on common projects collaborating in real time not in the office but online. Not only our professional lives moved online. Covid fuelled an explosive uptake in online shopping, including online groceries, and even entertainment moved online with online theatre and comedy gigs. The fact that all the technology existed to allow us to switch so many aspects of our lives online including the Sixth Annual T.M.C. Asser Lecture, employing broadcast quality video and audio connections to livestream events, and to take questions and answers, all at no cost to the attendee and limited cost to the organiser is, if you stop to think about it for a moment, quite astonishing.

This is the true revolution. The restrictions of covid, we hope, are temporary. Vaccinations are now available and are being deployed and soon we hope to return to normal life. Soon the phrase social distancing will take on a socially historical meaning, and will no longer be socially present. Hopefully that will remain the case indefinitely. The technology which allowed us to move such large parts of our lives online so easily is though a permanent change: it is part of our new world. When covid recedes will we all jump aboard planes and travel the world for conferences and workshops as regularly as we once did? It is likely we will not. These enabling technologies which have become central to our lives in the last eighteen months did not arrive newly minted in 2020. They have been developed over the previous twenty years or more. They were built to enable a human-digital hybrid world – one where our online and offline lives merge, or to use the phrase employed by Mireille Hildebrandt and Bert-Jaap Koops, they become ‘ambient’.<sup>6</sup>

<sup>6</sup> Mireille Hildebrandt and Bert-Jaap Koops, ‘The Challenges of Ambient Law and Legal Protection in the Profiling Era’ (2010) 73 *Modern Law Review* 428. <https://doi.org/10.1111/j.1468-2230.2010.00806.x>

However, these digital processes and technologies are disruptive to human agency and community. They are much more disruptive than any passing health crisis could ever be. Technology breaks down the traditional boundaries between individuals and communities, but it replaces them with new more limited horizons. Today there are second generation digital natives who communicate almost exclusively through digital intermediaries: instant messaging systems such as WhatsApp, photo messaging such as SnapChat and Instagram, audio messaging such as Soundcloud, and video messaging such as TikTok.<sup>7</sup> For these youngsters the idea of a phone as a device to make and receive voice calls is quite alien. In just one generation our social behaviour has changed. The multimedia opportunities offered by the smartphone, which remember first launched in 2007 with the original iPhone, have been fully integrated into the lives of these second-generation digital natives. In less than 15 years the concept of the phone in your pocket had been fully replaced by the idea of a multimedia entertainment and asymmetrical communications device. The smartphone has had a greater impact on society in a shorter time than any other product in history.<sup>8</sup>

Today many of us find ourselves doing business through digital intermediaries. Through online meeting systems such as Zoom, Teams, or Facetime, or through productivity tools such as Slack, and SharePoint. Our social connections are similarly intermediated through platforms such as Facebook, Instagram, or Twitter. What has changed is the unfiltered nature of communication between nodes. Our propensity

<sup>7</sup> A digital native is someone who has grown up with digital technology. The phrase was probably first used in 1995 by John Perry Barlow in an interview with *Australian Personal Computer*. Barlow adopted the phrase in his famous *Declaration of the Independence of Cyberspace* (<https://www.eff.org/cyberspace-independence>). It is most associated with Mark Prensky's 2001 paper, 'Digital Natives, Digital Immigrants Part 1' (2001) 9 *On the Horizon* 1. <https://doi.org/10.1108/10748120110424816>.

<sup>8</sup> This has been extensively written upon. See e.g. Muhammad Sarwar and Tariq Rahim Soomro, 'Impact of Smartphone's [sic] on Society' (2013) 98 *European Journal of Scientific Research* 216; Jane Vincent and Leslie Haddon (eds) *Smartphone Cultures* (Routledge, 2019); Yudhijit Bhattacharjee, 'Smartphones revolutionize our lives—but at what cost?', *National Geographic* 25 January 2019: <https://www.nationalgeographic.com/science/article/smartphones-revolutionize-our-lives-but-at-what-cost>.

to communicate directly, without intermediary, is being continually eroded. How often now do you book a restaurant using an online reservation service or ‘hire’ a taxi via an app? These are all intermediated services and as a result are quite different to the traditional person to person service offered via a telephone call where you speak directly to a person. Our ability today to do even these simple tasks relies upon digital technology which selects and presents information to us.

Today much of the smart money is moving into the so-called AI revolution and the creation of smart environments. Smart homes, smart offices, smart neighbourhoods and smart cities.<sup>9</sup> These take many forms. At one end are citywide initiatives such as the Amsterdam Smart City initiative launched in 2009 and which has over 170 projects driven by data gathering.<sup>10</sup> Examples of successful projects include one partnership between the city and local businesses and corporations on the Utrechtsestraat, a major shopping street. The Climate Street initiative includes energy-efficient lighting, waste reduction and recycling stations at tram stops. It has helped cut energy use on the Utrechtsestraat by 10%.<sup>11</sup> City-Zen, which stands for city zero carbon energy, was another successful partnership. The project sponsored the use of smart, future-proof energy grids and the retrofitting of buildings to be more sustainable. As a result, Amsterdam will save 59,000 metric tons per year in carbon dioxide: the equivalent of removing 12,000 cars from the road.<sup>12</sup>

At the other end of the scale is the smart home. You might have some smart home devices. A smart thermostat perhaps like Nest or Hive, or a smart home security system like Ring, or maybe just a smart hub

<sup>9</sup> Xiaomin Mou, ‘Artificial Intelligence: Investment Trends and Selected Industry Uses’ *EM Compass*, Note 71 September 2019: <https://www.ifc.org/wps/wcm/connect/7898d957-69b5-4727-9226-277e8ae28711/EMCompass-Note-71-AI-Investment-Trends.pdf?MOD=AJPERES>

<sup>10</sup> <https://amsterdamsmartcity.com/>.

<sup>11</sup> <https://amsterdamsmartcity.com/updates/news/what-makes-a-city-smart-tech-for-society-recap>

<sup>12</sup> Gabriel Jimenez, ‘What We Can Learn from the World’s Top 3 Smart Cities’ *Medium*, 7 March 2017: <https://medium.com/startup-grind/top-things-we-can-learn-from-a-city-tech-e474ef54bdf7>

or speaker system like Amazon's Echo system. The growing ubiquity of 'smart' and the use of the term AI or artificial intelligence in our lives captures the role that algorithmic decision-making plays in our lives.<sup>13</sup> We often refer to the so-called intelligent assistants in these devices as AI Assistants and nearly every major technology company has their own AI Assistant, from Apple's Siri, Amazon's Alexa and Google's Assistant (Google seem to be not very creative with names) through Microsoft's Cortana (taken from the Halo gaming series) to IBM's Watson. Through these tools companies compete for our attention and vitally for our data.<sup>14</sup> AI assistants create for these companies a two-sided market – they sell us personalised products and services while offering advertisers access to our data, our habits, and our tastes.<sup>15</sup> As a result, they are extremely valuable, and companies compete to get their assistant into our households by offering their technology to us often at a steep discount.

However, as we shall return to later in this lecture, the way these algorithmic assistants, who to be honest are not very intelligent, operate changes the way we make day-to-day decisions. While currently it might not be important if that decision is which road to follow while driving from Den Haag to Breda, or which restaurant in Leiden to book for dinner, as these algorithmic systems become more embedded in the operation of commerce, local and regional governance, the allocation of public resources, and even today on the battlefield, concerns become more acute and the theme of this lecture is how these

<sup>13</sup> Reuben Binns, 'Algorithmic Decision-making: A Guide For Lawyers' (2020) 25 *Juridical Review* 2; Omer Tene and Jules Polonetsky, 'Taming the Golem: Challenges of Ethical Algorithmic Decision-Making' 19 *North Carolina Journal of Law & Technology* 125 (2017).

<sup>14</sup> Noura Abdi, Kopo M. Ramokapane and Jose M. Such, 'More than Smart Speakers: Security and Privacy Perceptions of Smart Home Personal Assistants' *USENIX Symposium on Usable Privacy and Security* 2019: <https://www.usenix.org/system/files/soups2019-abdi.pdf>

<sup>15</sup> Kyuhong Park, Chanhee Kwak, Junyeong Lee and Jae-Hyeon Ahna, 'The effect of platform characteristics on the adoption of smart speakers: Empirical evidence in South Korea' 35 *Telematics and Informatics* 2118 (2018); Frank MacCrorry and Evangelos Katsamakos, 'Competition of Multi-Platform Ecosystems in the IoT' Available at SSRN: <https://ssrn.com/abstract=3737414>.

systems have the capacity to impact and even replicate human autonomy and agency and what we might do about it.

## 2. LAW: THE PRODUCT OF HUMAN AGENCY

Let us now turn our attention away from the simple social changes driven by the technologies that harness the process we call datafication, sometime called AI, machine learning, or smart systems. We will return to these technologies later but for the moment I want us instead to think about law and law-making.

To ask an impossibly broad question which cannot be fully answered here let us start by thinking: what is law? This is of course a philosophical, or to be precise jurisprudential, question. Your view as to what law is, in jurisprudential terms, reflects your own philosophy. Historically the dominant theory was natural law. Natural law espouses that law is founded upon, and reflects, values intrinsic to human nature. According to natural law theory, all people have inherent rights, conferred not by an act of legislation but by ‘God, nature, or reason.’<sup>16</sup> Natural law theory is often seen to be closely aligned to ‘theories of ethics, theories of moral politics, and theories of religious morality.’<sup>17</sup> Natural law was the ancient Greek philosophy of law. Plato discussed the *idea of the good*: the truth of all things, and also the light in which they shone forth, and became evident to intelligences human and divine,<sup>18</sup> while Aristotle recorded that ‘general laws [are] those based upon nature. In fact, there is a general idea of just and unjust in accordance with nature, as all men are in a manner divine, even if there is neither communication nor agreement between them.’<sup>19</sup> Despite being the dominant form of jurisprudence for mil-

<sup>16</sup> Hans Kelsen, *General Theory of Law and State* (Routledge, 2005), 392.

<sup>17</sup> Mark Murphy, ‘The Natural Law Tradition in Ethics’ in Edward N. Zalta (ed), *The Stanford Encyclopaedia of Philosophy* (Summer 2019 ed) (Metaphysics Research Lab, 2019): <https://plato.stanford.edu/archives/sum2019/entries/natural-law-ethics/>.

<sup>18</sup> Plato, *The Republic* (2<sup>nd</sup> ed) (trans. Desmond Lee) (Penguin, 2007), 508 a–c.

<sup>19</sup> Aristotle, *Rhetoric* (trans. Hugh Lawson-Tancred) (Penguin, 1991), 1.13.2.

lennia, adopted by among others Stoic scholars such as Cicero<sup>20</sup> and Seneca,<sup>21</sup> and theologians such as Thomas Aquinas,<sup>22</sup> natural law has over recent centuries lost out to legal positivism as the dominant jurisprudential school. The connection between law and morality has it seems been lost: maybe it is impossible to imagine modern lawyers and lawmakers as purely moral or good actors. A few natural law scholars remain. Probably the most famous contemporary proponents of natural law are the Australian legal philosopher John Finnis or Lon Fuller, the American legal philosopher who criticized legal positivism and defended a secular and procedural form of natural law theory in his 1964 classic text *Morality of Law*.<sup>23</sup>

The dominant jurisprudential theory currently is legal positivism. Positivism, unlike natural law, does not view law as an extension of morality or ethics. Thus, while Lon Fuller in the famous Hart-Fuller debate of 1958 wrote ‘to me there is nothing shocking in saying that a dictatorship which clothes itself with a tinsel of legal form can so far depart from the morality of order, from the inner morality of law itself, that it ceases to be a legal system’,<sup>24</sup> positivists have little difficulty finding that law can exist apart from morality or ethics.

The fathers of positivism are Hobbes and Bentham. Hobbes wrote that ‘law in general is not counsel, but command’,<sup>25</sup> and that ‘all laws, written and unwritten, have their authority and force from the will of the commonwealth, that is to say, from the will of the representative.’<sup>26</sup> The purging of morality from law was perfected by John Austin who described law as a social fact which reflects relations of power and

<sup>20</sup> Cicero, *The Republic and The Laws (De Re Publica)* (trans. Niall Rudd) (OUP, 2008), 3.33.

<sup>21</sup> Katja Vogt, ‘Seneca’, in Edward N. Zalta (ed), *The Stanford Encyclopaedia of Philosophy* (Spring 2020 ed) (Metaphysics Research Lab, 2020): <https://plato.stanford.edu/archives/spr2020/entries/seneca/>.

<sup>22</sup> Thomas Aquinas, *Summa Theologiae*, I.II 91,2: <https://aquinas101.thomisticinstitute.org/st-iaii-ae-q-91#FSQ91OUTP1>.

<sup>23</sup> Lon L. Fuller, *The Morality of Law* (Yale University Press, 1964).

<sup>24</sup> Lon L. Fuller, ‘Positivism and Fidelity to Law – A Reply to Professor Hart’ 71 *Harvard Law Review* 630 (1958), 660.

<sup>25</sup> Thomas Hobbes, *Leviathan* (Penguin, 2017), ch. XXVI, ¶ 2.

<sup>26</sup> Hobbes, *Leviathan*, ch. XXVI, ¶ 10.

obedience. For Austin this presents a twofold view: (1) that law and morality are separate; and (2) that all human-made (positive) laws can be traced back to human lawmakers.

Austin describes two forms of positive law:

1. The positive laws accounted natural: These are those which are common to all political societies in the character of positive laws, and beings palpably useful to every society, have their counterpart in the shape of moral rules in every society, political or natural.
2. Laws accounted positive, as opposed to natural, are not common to all political societies, or, if common, have not their counterpart in the moral rules of all societies (political or natural).<sup>27</sup>

The former reflects the old science of natural law, the latter the new science of positive law. Although very few legal positivists would recognise Austin's definition today, his rather dated model having been comprehensively picked apart by the great Herbert Hart in his 1958 Essay *Positivism and the Separation of Law and Morality*,<sup>28</sup> his part of the famous Hart-Fuller debate, Austin made a vital contribution to our understanding of law and law-making. He clearly, and for the first time, set out that law could be man-made rather than derived from God or from innate values of morality. As he noted: 'Laws in the literal or proper sense are rules laid down for an intelligent being by an intelligent being having authority over him, and in this sense the term Law comprises: Laws set by God to men; and Laws set by men to men.'<sup>29</sup>

As noted, pure Austinian positivism is rarely recognised today. In the 1950s Herbert Hart developed his 'concept of law' which would be published in book form in 1961.<sup>30</sup> Here he critiqued the simplicity of Austin's authority argument and introduced his famous Rule of

<sup>27</sup> John Austin, *Lectures on Jurisprudence: Or the Philosophy of Positive Law*, (1869, Reprint Forgotten Books 2015), Lecture XXXII.

<sup>28</sup> H.L.A. Hart, 'Positivism and the Separation of Law and Morality' 71 *Harvard Law Review* 593 (1958).

<sup>29</sup> Austin, *Lectures on Jurisprudence*, Part I, Section I.

<sup>30</sup> H.L.A. Hart, *The Concept of Law* (Clarendon, 1961).

Recognition, the most pre-eminent of Hart's secondary rules. This rule is not normally set out expressly but can be induced from the conduct of those whose task it is to operate the rule system. Hart explains the rule of recognition as follows:

the rule of recognition exists only as a complex, but normally concordant, practice of the courts, officials, and private persons in identifying the law by reference to certain criteria. Its existence is a matter of fact.<sup>31</sup>

Other contemporary schools of thought exist, such as legal interpretivism, the school most closely associated with Ronald Dworkin.<sup>32</sup> Interpretivism can be distinguished from positivism as interpretivism does not see law as a set of given data, conventions, or facts, but instead what lawyers aim to construct or obtain in their practice.<sup>33</sup> Equally interpretivism is not a contemporary version of natural law for in interpretivism law is not immanent in nature nor do legal values and principles exist independently and outside of the legal practice itself.<sup>34</sup> This is the opposite of the main claim of natural law theory.

While further jurisprudential schools of thought exist including critical legal studies,<sup>35</sup> feminist legal studies<sup>36</sup> and transnational or polycentric law,<sup>37</sup> by focusing in on the two historically leading schools, on natural law and positivism, we see some vital differences but also some commonalities. The key focus of dispute between the schools of thought is between the morality of laws: the focus of the Hart-Fuller debate. While natural lawyers view morality as central to law, positivists break the connection, this is something we will return to in this lecture. Interpretivism which represents one of many branch-

<sup>31</sup> Hart, *The Concept of Law*, 110.

<sup>32</sup> Ronald Dworkin, *Law's Empire* (Belknap Press, 1986).

<sup>33</sup> David Plunkett and Timothy Sundell, 'Dworkin's Interpretivism and The Pragmatics of Legal Disputes' (2013) 19 *Legal Theory* 242.

<sup>34</sup> Thom Brooks, 'Between Natural Law and Legal Positivism: Dworkin and Hegel on Legal Theory' 23 *Georgia State University Law Review* 513 (2006).

<sup>35</sup> Mark Kelman, *A Guide to Critical Legal Studies* (Harvard University Press, 1987).

<sup>36</sup> Joanne Conaghan (ed), *Feminist Legal Studies* (Routledge, 2009).

<sup>37</sup> Roger Cotterrell, 'What Is Transnational Law?' (2012) 37 *Law & Social Inquiry* 500.

es of legal realism focuses not on morality but on the acts of legal actors. What is clear is that since the demise of Stoic/theological natural law in the 19<sup>th</sup> Century, all legal theorists, even natural lawyers like Fuller who presented a secular/procedural form of natural law and Finnis who recognises that ‘a theory of natural law claims to be able to identify conditions and principles of practical right-mindedness, of good and proper order among persons, and in individual conduct’,<sup>38</sup> recognise that law is the product of human agency, not a product of God or nature.

Thus, the starting point for this examination is that law, however you define it, is formed in the crucible of human experience, human thought, and human interaction. Law might be seen along with the concept of morality or ‘right’ as the ultimate expression of human experience. Some laws will reflect what we call morality. These are the laws recognised by the ancient Greeks and the Stoics: the natural laws. As Austin put it ‘the positive laws accounted natural are those which are common to all political societies in the character of positive laws, and, being palpably useful to every society, have their counterpart in the shape of moral rules in every society.’<sup>39</sup>

In the earliest human culture, we knew attacks on the person such as assault, murder, and rape were intrinsically wrong: they were morally wrong. Law in all cultures developed a common prohibition on these. As society developed commerce and property, we developed common property laws such as ownership, theft, and trespass. Natural law theory associated law with morality as they grew from a common root: human experience, human culture, and human values. As our societies became more complex, more administrative, as human experience and human interaction became more complex, we moved away from the moral roots of law and theology. We became more secular as simple theological morality ebbed and more complex ethical frameworks developed: today the world is less black and white and more shades of grey. The same is true of law. Austin fired a starting pistol but now we see that law is more abundant than the natural or moral

<sup>38</sup> John Finnis, *Natural Law & Natural Rights* (2<sup>nd</sup> ed) (OUP, 2011), 18.

<sup>39</sup> Austin, *Lectures on Jurisprudence*, Lecture XXXII.

law principles that regulated criminality, property, and family relationships. Law is now corporate, administrative, and regulatory. It is permissive and restrictive. It is sometimes paternalistic and at other times libertarian. At all points though, law is about agency – the human capacity to act independently and to make our own free choices. As Jeremy Webber observes, ‘Law is consciously created’<sup>40</sup> and is the distillation of the collective agency of a society, group, or culture. The recognition of, and respect for, the rule of law is the ultimate distillation of this principle: the clear spirit of human choice in the purest form.

### 3. DATAFICATION AND HUMAN AGENCY

What impact does datafication and the increased deployment of artificially intelligent machine learning systems have on human agency and on the rule of law? In the second part of this lecture, we will examine whether the adoption of such systems pose a threat to the principle set out in the first part that: ‘law, and the rule of law, is the clear spirit of human choice in the purest form.’

Firstly, let us dispel a myth: there is no such thing as AI as a singular concept. Artificial Intelligence is an anthology term for all types of technology which seek to mimic biological systems intelligence. Colloquially, the term is often used to describe machines (normally computers) that mimic cognitive functions that humans associate with the human mind, such as learning and problem solving.<sup>41</sup> One researcher, Larry Tesler, a computer scientist who worked at Xerox PARC, Apple, Amazon, and Yahoo! over the course of his career, famously quipped that ‘Intelligence is whatever machines haven’t done yet.’<sup>42</sup> This puts one in mind of the famous Arthur C. Clarke quote

<sup>40</sup> Jeremy Webber, ‘Legal Pluralism and Human Agency’ (2006) 44 *Osgoode Hall Law Journal* 167, 177.

<sup>41</sup> Stuart J. Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (4<sup>th</sup> ed) (Prentice Hall, 2020), 1-2.

<sup>42</sup> Think Automation, *Tesler’s theorem and the problem of defining AI*: <https://www.thinkautomation.com/bots-and-ai/teslers-theorem-and-the-problem-of-defining-ai/>.

that ‘any sufficiently advanced technology is indistinguishable from magic.’<sup>43</sup>

This is a problem for a researcher in the field of emerging AI. What looks like intelligence now, may look like simple processing later. When the IBM programme Deep Blue defeated Gary Kasparov in 1997 it was initially greeted as a success for ‘cognitive computing’; however, now much more advanced computer chess programmes are viewed as an application of overwhelming brute force rather than cognition. Computers can beat human grandmasters not because they think, but because they can process more information more quickly. This is an application of so-called If This Then That (IFTTT) processing. IFTTT is algorithmic decision-making in a closed system: it can be incredibly complex and can make it look as if a computer is acting with cognition.<sup>44</sup> Essentially though, all IFTTT does is give a set of instructions to a computer to direct an output. It is not dissimilar to driving a car – inputs from the user, or the environment, directly drive responses from the machine. So at the simplest level, if someone approaches my door my smart doorbell starts recording and sends a notification to my phone; at a more complex level, if traffic sensors record slow-moving traffic on a route into the city and weather reports indicate fog, it can lower the speed limits on the road and re-route traffic by sending messages to automatic road signs and updating advice to satellite navigation systems; or to take an example from the so-called smart battlefield – the Israeli Harop drone is a fully autonomous loitering weapon which can patrol an area of over 1000 km and which when it detects RADAR emissions can autonomously target the source of these emissions.<sup>45</sup>

Today much discussion of ‘true’ Artificial Intelligence focuses on Machine Learning. Machine Learning operates quite differently to IFTTT. It more closely mimics the human brain by seeing patterns rather than just processing data and following instructions. Machine Learning

<sup>43</sup> Arthur C. Clarke, *Profiles of the Future* (revised edition) (MacMillan, 1973).

<sup>44</sup> Mireille Hildebrandt, ‘Algorithmic regulation and the rule of law’ *Philosophical Transactions of the Royal Society A* 376:20170355.

<sup>45</sup> [https://defense-update.com/20090110\\_harop.html](https://defense-update.com/20090110_harop.html).

Algorithms (MLAs) use statistics to find patterns in massive amounts of data and then check whether these patterns are correct.<sup>46</sup> In essence the machine brain learns much the same way as a human brain in its developmental phase. Is that pattern a lemon or an apple? Is it edible? A human child will process the world about them in a pattern-forming fashion. First look at something – do I recognise it? Then if possible, touch it and finally taste it. When the child eats the lemon, they record it to be not easily edible and remember the patterns to allow them to recognise lemons in future. Essentially human brains learn by processing vast amounts of information to create patterns.<sup>47</sup> If you had never encountered a lime before you would probably process it as being a ‘green lemon’. Indeed, in French citron vert is the name given to a lime. Machine Learning replicates this pattern-learning process by something called backpropagation (or backprop).<sup>48</sup> I won’t go into detail here as to how backprop works but at a basic level it works by building large networks of processors known as neural networks. You then train that network to recognise patterns based upon an initial training set. For example, you might want the network to recognise when a shadow on an x-ray may be cancer. You collect a dataset of millions of x-ray images of possible cancer patients and label them as cancer or not cancer depending upon the diagnosis made. You then manually encode the first training set showing the network the differences between the two patterns. Now you start to backprop. You give the network unlabelled datasets and ask it to determine – cancer or not cancer? At the end of a round of determinations the system checks its score, just as a human teacher would with a pupil. Did it identify the tumours on the training images? Did it correctly spot the non-cancerous shadows? It then revisits any it got wrong and tries to see why it got them wrong. By going through it-

<sup>46</sup> Karen Hao, ‘What is machine learning?’ *MIT Technology Review*, 17 November 2018: <https://www.technologyreview.com/2018/11/17/103781/what-is-machine-learning-we-drew-you-another-flowchart/>.

<sup>47</sup> Mark P. Mattson, ‘Superior pattern processing is the essence of the evolved human brain’ (2014) 8 *Frontiers in Neuroscience* 265.

<sup>48</sup> Robert Hecht-Nielsen, ‘Theory of the Backpropagation Neural Network’. In: Harry Wechsler (ed) *Neural Networks for Perception* (Elsevier, 1992); Yves Chauvin and David E. Rumelhart (eds) *Backpropagation: Theory, Architectures, and Applications* (Lawrence Erlbaum, 1995).

erations of this process very quickly Machine Learning can outperform humans at spotting patterns – although over the years humans have helped. If you have ever done an online CAPTCHA or Completely Automated Public Turing test to tell Computers and Humans Apart you have been an active part of Machine Learning training. Early CAPTCHAs often used unclear text and numbers. This helped train Google’s book digitisation AI.<sup>49</sup> By showing the same words to multiple users, it could verify that a word has been transcribed correctly by comparing multiple attempts from multiple users across the world. More recently CAPTCHAs have been used to train autonomous vehicles – this is why you may have been asked to highlight all the road signs or buses in an image.<sup>50</sup> The problem now is that the AI is so well-trained that CAPTCHAs can be defeated by AI in 99% of cases. As a result, new CAPTCHAs now often ask you just to click you are not a robot as humans have a different mechanical response to bots, or they may ask you to play a little image-based game which AI finds more difficult to replicate.<sup>51</sup>

What hopefully has become clear is that all forms of AI from the more basic IFTTT to more sophisticated forms of Machine Learning replicate biological thought patterns through the processing of massive amounts of data. While the process of pattern recognition applied in Machine Learning can produce outcomes which exceed human experts in fields such as medicine,<sup>52</sup> law,<sup>53</sup> and finance<sup>54</sup> (humans are emi-

<sup>49</sup> Josh Dzieza, ‘Why CAPTCHAs have gotten so difficult’ *The Verge* 1 February 2019: <https://www.theverge.com/2019/2/1/18205610/google-captcha-ai-robot-human-difficult-artificial-intelligence>.

<sup>50</sup> Mark Healy, *Captcha if you can*: <https://www.ceros.com/inspire/originals/recaptcha-waymo-future-of-self-driving-cars/>.

<sup>51</sup> Josh Dzieza, ‘Why CAPTCHAs have gotten so difficult’.

<sup>52</sup> Michael A. Froomkin, Ian Kerr and Joelle Pineau, ‘When AIs Outperform Doctors: Confronting the Challenges of a Tort-Induced Over-Reliance on Machine Learning’ 61 *Arizona Law Review* 33 (2019).

<sup>53</sup> Andrew D. Martin et al., ‘Competing Approaches to Predicting Supreme Court Decision Making’, 2 *Perspectives on Politics* 761 (2004); Theodore W. Ruger et al., ‘The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decision-Making’ 104 *Columbia Law Review* 1150 (2004).

<sup>54</sup> Marcos Lopez de Prado, *Advances in Financial Machine Learning* (Wiley, 2018).

nently fallible) it is doing something different to the human expert. Much like Deep Blue in 1997 it defeats the human because it has access to more data, and it has more ‘experience’ as a result. What it lacks is intuition. This is often called the frame problem. When learning an AI makes incremental changes of values – it refines its pattern recognition rather than ‘rethinking’ it. McCarthy and Hayes interpret this problem as ‘the problem of blocking the vast number of inferences about what had not changed as the result of some action A while allowing the small number of inferences about what had changed as a result of A.’<sup>55</sup> Biological intelligence can draw on multiple resources simultaneously to solve a problem whereas Machine Learning, as yet, cannot. So, if I were to train a Machine Learning system in a virtual simulation to open a virtual box which is booby trapped it would learn through iterations how to prevent the traps going off. The box may have a glass on top that falls and breaks when the box is moved. The AI software would recognise that pattern and the next time it would remove the glass before moving the box – this is binary cause and effect, and Machine Learning is good at this – a bit like playing chess. However, finding the box is now taped closed the AI would likely drop or force the box open, likely damaging the contents. Again, this is a cause and effect learning pattern. A human would instead look for something sharp such as scissors to open the box because we can access unrelated knowledge – that sharp objects can open tape.<sup>56</sup>

The reason for this discussion is to demonstrate that computer AI is very A and not terribly I. No doubt technology will continue to advance towards the eventual AI goal of Human Level Machine Intelligence but for the moment it is important to recognise that AI *imitates* human thought by processing vast datasets and establishing patterns within that data – *this is not how humans think*.<sup>57</sup>

<sup>55</sup> John McCarthy and Patrick J. Hayes, ‘Some Philosophical Problems from the Standpoint of Artificial Intelligence’, in *Machine Intelligence* (4<sup>th</sup> ed) Donald Michie and Bernard Meltzer, (Edinburgh University Press, 1969), 463–502.

<sup>56</sup> See also Steven Pinker, *How the Mind Works* (Penguin, 1997), 14–15.

<sup>57</sup> Pinker, *How the Mind Works*.

The process used in all forms of AI, whether it be simple IFTTT or more complex Machine Learning is datafication: the reduction of the complexity of the world to data values.<sup>58</sup> Computers in general, and AI in particular, are not good with noisy or jaggy data – outliers in particular are not well managed by AI. Complexity becomes numerical values and patterns and choices become mathematical processes. Datafication, although helpful for binary machines such as computers which like to process in yes/no or 1/0 trees, is extremely unhelpful for humans.<sup>59</sup> Our brains are less equipped for this form of decision-making and as a result the risk of datafication is that human decision-making is slowly replaced by algorithmic decision-making, or at least algorithmically supported decision-making.

This increases the risk of automation bias, whereby there is over-reliance on machine outputs at the expense of human knowledge and experience, causing human skillsets to decline dramatically.<sup>60</sup> This leads to technological lock-in or the ‘computer says no’ problem. This occurs when humans without question apply the outcome of the machine element of a decision-making process both as they over-trust the machine’s judgement as being somehow better than their own, and as a shield, assuming, usually correctly, that it is more difficult to critique the logical and objective decision of the machine over the subjective, possibly emotional, decision of the human. Experience shows that by simply adopting the position of the machine the human operator can deflect any critique or challenge. The outcome can be narrow, inflexible, unjust decisions which can be difficult to challenge as logically they are flawless.

To reduce risks of error or bias on the part of employees most businesses now operate computerised customer systems. A clear example

<sup>58</sup> Mark Lycett, “Datafication”: making sense of (big) data in a complex world’ (2013) 22 *European Journal of Information Systems* 381; Jeremy Knox, Ben Williamson and Sian Bayne, ‘Machine behaviourism: Future visions of “learnification” and “datafication” across humans and digital technologies’ (2020) 45 *Learning, Media and Technology* 31.

<sup>59</sup> Matthew Cobb, *The Idea of the Brain* (Profile, 2020).

<sup>60</sup> Rebecca Crootof, “Cyborg Justice” and the Risk of Technological Legal Lock-in’ 119 *Columbia Law Review Forum* 233, 236.

of this is in banking. If you have ever visited a bank for a loan over the last twenty years you will know the system. An employee sitting in front of a computer asks a series of questions – some are very straightforward data questions such as how much you earn. Others might be more complex. They might ask ‘do you smoke?’ The true answer ‘only when I have had too much to drink’ cannot be entered onto a binary yes/no IFTTT tree so the employee must input yes even though the answer is ‘not really’. At the end a numerical score determines whether you are a good risk for a product. When the computer says ‘no’ and you ask the human operator to explain why not, they refer to the output of the computer: ‘computer says no’. They can provide no further explanation. Human agency has been diminished at the cost of a crude algorithmic calculation which was reached by imperfectly reducing you, the applicant, to data points. This is sensible for the bank. They want consistent, objective, results every time. However, it dehumanises both the applicant and the bank employee. Before the datafication of banking you could speak to the employee and they, using their expertise and training, would make a risk calculation based on everything you said to them, not just that which could be reduced to the thirty data points the software asks for. So, the fact for instance that as someone who in the past has had difficulty managing your finances you undertook a financial management course at your local college could be considered by the human but cannot by the programme that has no box to input that data.

Examples abound of a rush towards datafication and the ceding of human decision-making towards algorithmic decision-making and algorithmic regulation. As the world becomes more complex, we seek to make sense of it through greater use of data. Datafication, often referred to as big data, is the key science of the next twenty years (and the last ten) and underpins Machine Learning, the key formulation of what we currently call AI. In my 2015 LSE Inaugural Lecture *Open the Pod Bay Doors HAL: Machine Intelligence and the Law* I set out three ways in which datafication had changed the way humans think and act.<sup>61</sup>

<sup>61</sup> <https://www.youtube.com/watch?v=YyNeE8iLkxw>.

The most basic of these will be familiar. *Assisted decision-making* occurs where digital technology replaces a previous analogue technology or artefact to make our decision-making process simpler. This is the digitisation of reference texts including encyclopaedias, phone books and even newspapers. These all become instantly accessible through search engines. So, if in 1992 you were asked who won the 1972 Eurovision Song Contest, assuming you did not know the answer, you had to wait until you could consult a reference text to find out. Now it is a matter of seconds (and Google) to find that it was Luxembourg with the song ‘Après toi’, performed by Vicky Leandros. As life has become more complicated, we rely more on technology to help us filter all the information available to us and to present to us the most relevant information to make the best decision we can. We cannot possibly hope to retain at our fingertips all the information we need to manage our complicated lives, so we use digital devices to order and manage this data. Essentially this is about empowerment. We have the ability now to have in our hands our diary, our messages, our maps, the most powerful encyclopaedia known to man, our telephone, our photographs, our entertainment media and much more.

There is a second more developed version of this called *supplementary decision-making*. This takes place when technology offers us information or processes, which simply were not available before the widespread adoption of digital technologies. At the most basic level these are health apps and allied technologies which monitor a variety of variables such as activity, food intake, heart rate, blood pressure etc. and which then give us health advice.<sup>62</sup> In theory it would have been possible to monitor all these variables before health apps and wearables but the lack of portability of health monitors meant this was not a practical possibility. A greater application of supplementary decision-making technologies is now fitted to our vehicles where several safety systems assist our driving such as anti-lock and automatic braking, electronic stability control (including traction control), adaptive headlamps and collision avoidance systems.

<sup>62</sup> Luke Dormehl, *The Formula: How Algorithms Solve all our Problems ... and Create More* (WH Allen, 2014), Chapter 1.

The third category is truly *autonomous decision-making*. This is where we completely outsource decision-making to the intelligent agent and the human no longer contributes to the decision-making process. Although still uncommon we are seeing more of these systems being developed and deployed. Many modern cars not only have driver assist technologies but now have driving management systems that cannot be turned off or overridden by the driver. We are seeing the emergence of autonomous vehicles, drones, and robots. By 2020 the Komatsu ‘Frontrunner’ autonomous haulage system had moved over three billion tonnes of material at eleven mining sites using over 250 trucks 24 hours a day with no accidents or casualties.<sup>63</sup> Meanwhile Amazon have opened thirty Amazon Go stores which employ sensors throughout the store and Machine Learning systems to allow customers to enter, select items and leave without ever encountering a human operative.<sup>64</sup>

All these systems undermine human agency and human autonomy in some way. When we think of assisted decision-making: the selection and presentation of data, decided by algorithms, affects our ability to make fully autonomous decisions. The information we receive is pre-filtered. Various writers give this process different names – internet activist Eli Pariser calls it the Filter Bubble,<sup>65</sup> while Cass Sunstein referred to it as digital Balkanisation.<sup>66</sup> One difference with the pre-digital age is how this filtering is done. Prior to the internet and Google, you would have used a library catalogue or book index to access information. These were prepared by highly skilled humans – librarians and indexers –, now this is done algorithmically by the reduction of information to data points.

<sup>63</sup> Komatsu, *Komatsu customers surpass 3B metric tons moved autonomously*, 30 July 2020: [https://home.komatsu/en/press/2020/others/1206865\\_1845.html](https://home.komatsu/en/press/2020/others/1206865_1845.html).

<sup>64</sup> Maggie Tillman, ‘Amazon Go and Amazon Fresh: How the “Just walk out” tech works’ *Pocket-lint*, 4 March 2021: <https://www.pocket-lint.com/gadgets/news/amazon/139650-what-is-amazon-go-where-is-it-and-how-does-it-work>.

<sup>65</sup> Eli Pariser, *The Filter Bubble: What the Internet Is Hiding from You* (Penguin, 2011).

<sup>66</sup> Cass Sunstein, *Republic.com* (Princeton University Press, 2001), 65-69.

The role of algorithmic intermediaries in human decision-making has become a pre-eminent concern of governments and regulators after the Facebook algorithm was exploited in the 2016 UK Brexit referendum campaign and the 2016 US Presidential elections.<sup>67</sup> It is clear that exploitative, perhaps even disinformative, campaigns were run in these campaigns and as a result words such as ‘fake news’, ‘astroturfing’, and ‘cyberturfing’ have entered the mainstream.<sup>68</sup> Internet users have as a result become aware that digital platforms use algorithms to manage what they see. Some people argue this is manipulative, while the platforms contend it is about providing the best user experience. As US law professor Tim Wu points out in his book *The Attention Merchants*, these businesses sell our attention to advertisers, they need to keep us on their sites and engaged for as long as possible – telling (or selling) us the information we want to receive is more likely to hold our attention.<sup>69</sup> The best user experience therefore in the view of these platforms is one which encourages us to use their product, a point made repeatedly by social psychologists.<sup>70</sup>

But it is not only in the limited sphere of online social interaction that datafication is reducing human agency. Datafication is also widely employed in the offline world. This lecture has already referred to the use of credit scoring systems in banking and finance, something

<sup>67</sup> UK Parliament Digital, Culture, Media and Sport Committee, *Disinformation and fake news: Final Report*, HC 1791, 18 February 2019; European Commission, *The digital transformation of news media and the rise of disinformation and fake news*, JRC111529, April 2018; Congressional Research Service, *Social Media: Misinformation and Content Moderation Issues for Congress*, R46662, 27 January 2021.

<sup>68</sup> Mark Leiser, ‘Astroturfing, “CyberTurfing” and other online persuasion campaigns’ (2016) 7 *European Journal of Law and Technology* 1; Johan Farkas and Christina Neumayer, ‘Disguised Propaganda from Digital to Social Media’ in Jeremy Hunsinger, Lisbeth Klastrup and Matthew Allen (eds) *Second International Handbook of Internet Research* (Springer, 2018).

<sup>69</sup> Tim Wu, *The Attention Merchants: The Epic Struggle to Get Inside Our Heads* (Atlantic, 2017).

<sup>70</sup> David Greenfield, ‘The Addictive Properties of Internet Usage’ in Kimberly S. Young and Cristiano Nabuco de Abreu (eds) *Internet Addiction: A Handbook and Guide to Evaluation and Treatment* (Wiley, 2007); Uichin Lee et al., ‘Hooked on Smartphones: An Exploratory Study on Smartphone Overuse among College Students’ (2014) *Proceedings of the 32nd Annual ACM conference on human factors in computing systems* 2327; Li Shih-Ming and Chung Teng-Ming, ‘Internet function and Internet addictive behavior’ (2006) 22 *Computers in Human Behavior* 1067.

we have been familiar with for years, but in addition law enforcement is trialling a number of algorithmic decision-making systems such as PredPol, a predictive policing algorithm which claims to be able to predict crime hotspots and to direct police resources to areas where crimes are likely to occur,<sup>71</sup> COMPAS (or Correctional Offender Management Profiling for Alternative Sanctions), a system which uses an algorithm to assess potential recidivism risk,<sup>72</sup> and HART (or Harm Assessment Risk Tool), which like COMPAS assesses the risk of recidivism.<sup>73</sup> Lawyers are also rushing to replace years of law school training with algorithms designed to reduce the variety and challenges of the legal profession, and the skills they have developed through years of education and legal training, to a set of data points. If you work in the law, you can't have missed the rise of LegalTech systems such as Luminance, which offers natural language processing document review,<sup>74</sup> or Libryo, which automates regulatory compliance.<sup>75</sup> These (and many more) are replicating tasks usually carried out by legal secretaries, paralegals and trainee and junior lawyers.

You might think all of this is interesting but simply the march of technology. No one wants to be Ned Ludd destroying the tools of future development. In most of these examples the societal risks are minimal. The exception is the use of AI in law enforcement where errors and biases can lead to a loss of liberty. Studies of these tools unfortunately do not paint a happy picture.

<sup>71</sup> <https://www.predpol.com/about/>

<sup>72</sup> Northpointe, *Practitioner's Guide to COMPAS Core*, 19 March 2015: <https://assets.documentcloud.org/documents/2840784/Practitioner-s-Guide-to-COMPAS-Core.pdf>; Jeff Larson, Surya Mattu, Lauren Kirchner and Julia Angwin, 'How We Analyzed the COMPAS Recidivism Algorithm' *ProPublica* 23 May 2016: <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>.

<sup>73</sup> Geoffrey Barnes and Lawrence Sherman, Helping police make custody decisions using artificial intelligence, *Research Horizons*, 26 Feb 2018: <https://www.cam.ac.uk/research/features/helping-police-make-custody-decisions-using-artificial-intelligence>; Marion Oswald, Jamie Grace, Sheena Urwin and Geoffrey Barnes, 'Algorithmic risk assessment policing models: Lessons from the Durham Constabulary HART model' (2018) 27 *Information and Communications Technology Law* 223.

<sup>74</sup> <https://www.luminance.com/>.

<sup>75</sup> <https://libryo.com/>.

In 2016 two researchers, Kristian Lum and William Isaac, found that PredPol reinforces biases found in historical policing data. The locations that are flagged for targeted policing are those that were, by their estimates, already overrepresented in the historical police data.<sup>76</sup> At the time PredPol had recently been offered to the Oakland Police Department as efficient and unbiased but following the research of Lum and Isaac they declined to purchase the software.<sup>77</sup> What was going wrong? Earlier we discussed how Machine Learning learns: it looks for patterns in data (so-called learning datasets) and then looks to spot them in other data (the applied data). PredPol was trained on policing data gathered in the 2000s. Although the designers of PredPol took steps to ‘clean’ that data by for example removing individual arrest data and removing data relating to drug and prostitution enforcement (as they are more heavily influenced by whom police choose to target), the training dataset still reflected the traditional biases of early 21<sup>st</sup> Century US policing. Thus, the algorithm learned that crimes were more likely to occur in Black and Hispanic neighbourhoods which had traditionally been overpoliced.<sup>78</sup> As a result, it predicted more crimes in these areas which caused more policing to be deployed there and more crime detection in these areas. This confirmed the algorithm’s predictions and led to a confirmation bias. This led Lum and Isaac to record that: ‘Using PredPol in Oakland, black people would be targeted by predictive policing at roughly twice the rate of whites. Individuals classified as a race other than white or black would receive targeted policing at a rate 1.5 times that of whites. This contrasts with the estimated pattern of drug use by race where drug use is roughly equivalent across racial classifications.’<sup>79</sup> To be clear: the algorithm did what it was programmed to do. It worked, and it continues to work, and a number of police departments still employ

<sup>76</sup> Kristian Lum and William Isaac, ‘To Predict and Serve?’ (2016) 13 *Significance* 14.

<sup>77</sup> Emily Thomas, ‘Why Oakland Police Turned Down Predictive Policing’ *Motherboard* 28 December 2016: <https://www.vice.com/en/article/ezp8zp/minority-retort-why-oakland-police-turned-down-predictive-policing>.

<sup>78</sup> Renata M. O’Donnell, ‘Challenging Racist Predictive Policing Algorithms Under the Equal Protection Clause’ 94 *NYU Law Review* 544 (2019); Aaron Shapiro, ‘Reform predictive policing’ (2017) 541 *Nature* 458.

<sup>79</sup> Lum and Isaac, ‘To Predict and Serve?’, 18.

PredPol.<sup>80</sup> There is no objective or mathematical problem with the algorithm. Also, we should not blame the developers of the software – they took all reasonable steps to eliminate bias, aware this was a risk. The problem is simply how Machine Learning works. It replaces human bias and error with machine bias and error. *We cannot completely remove error.*

We must bear this in mind as we think about the deployment of AI in warfare. Currently there is considerable discussion on the use of Lethal Autonomous Weapons Systems on the battlefield.<sup>81</sup> This would allow for the first-time algorithms to be lawfully permitted to kill humans. Currently there are three levels of LAWS system:<sup>82</sup>

1. Supervised autonomous operation: where the machine, once activated, performs a task under the supervision of a human and will continue performing the task unless the human operator intervenes to halt its operation. This control type is often referred to as ‘human on the loop’. Examples include Israel’s Iron Dome antimissile system, the US Navy’s Phalanx Aegis-class ship defence system, and the US Army’s Patriot batteries.
2. Semi-autonomous operation, where the machine performs a task and then stops and waits for approval from the human operator before continuing. This control type is often referred to as ‘human in the loop’. This would include Reaper Drones and laser guided missiles.

<sup>80</sup> See Caroline Haskins, ‘Dozens of Cities Have Secretly Experimented With Predictive Policing Software’ *Motherboard* 2 June 2019: <https://www.vice.com/en/article/d3m7jq/dozens-of-cities-have-secretly-experimented-with-predictive-policing-software>.

<sup>81</sup> Daniele Amoroso, *Autonomous Weapons Systems and International Law: A Study on Human-Machine Interactions in Ethically and Legally Sensitive Domains* (Edizioni Scientifiche Italiane, 2020); Amitai Etzioni and Oren Etzioni, ‘Pros and Cons of Autonomous Weapons Systems’ (2017) *Military Review* 72; International Panel on the Regulation of Autonomous Weapons (iPRAW), *Focus on Human Control* (2019): [https://www.ipraw.org/wp-content/uploads/2019/08/2019-08-09\\_iPRAW\\_HumanControl.pdf](https://www.ipraw.org/wp-content/uploads/2019/08/2019-08-09_iPRAW_HumanControl.pdf).

<sup>82</sup> William C. Marra and Sonia K. McNeil, ‘Understanding “The loop”: Regulating the Next Generation of War Machines’ 36 *Harvard Journal of Law & Public Policy* 1139 (2013).

3. Fully autonomous operation, where the machine, once activated, performs a task and the human operator does not have the ability to supervise its operation and intervene in the event of system failure. This control type is often referred to as 'human out of the loop'. The only current example to our knowledge of this is the Harpy missile discussed earlier but it is believed that in particular the US Military has several such weapons in development or in deployment.

There is currently a heated debate about the legality of these weapons.<sup>83</sup> In March 2019, UN Secretary-General António Guterres urged AI experts meeting in Geneva to push ahead with their work to restrict the development of lethal autonomous weapons systems. He said that 'machines with the power and discretion to take lives without human involvement are politically unacceptable, morally repugnant and should be prohibited by international law.'<sup>84</sup> However, as yet there is no UN treaty banning such weapons and the latest meeting of the Group of Governmental Experts on Lethal Autonomous Weapons Systems in September 2020 remained split between governments that want a ban of LAWS such as Brazil, China and Egypt,<sup>85</sup> countries who believe some further regulation may be required such as the United States, Russia and South Korea but who argue it is premature now to think about how that regulation may be designed,<sup>86</sup> and countries like the United Kingdom who believe the existing international

<sup>83</sup> Kenneth Anderson, Daniel Reisner and Matthew Waxman, 'Adapting the Law of Armed Conflict to Autonomous Weapon Systems' 90 *International Law Studies* 386 (2014); Kenneth Anderson and Matthew C. Waxman, 'Law and Ethics for Autonomous Weapon Systems: Why a Ban Won't Work and How the Laws of War Can' *Columbia Public Law Research Paper* 13-351 (2013); Robert Sparrow, 'Robots and Respect: Assessing the Case against Autonomous Weapon Systems' 30 *Ethics and International Affairs* 93 (2016); Erica H. Ma, 'Autonomous Weapons Systems under International Law' 95 *NYU Law Review* 1435 (2020).

<sup>84</sup> António Guterres, 'Autonomous weapons that kill must be banned, insists UN chief' *UN News* 25 March 2019: <https://news.un.org/en/story/2019/03/1035381>.

<sup>85</sup> Human Rights Watch, *Stopping Killer Robots: Country Positions on Banning Fully Autonomous Weapons and Retaining Human Control* 10 August 2020: <https://www.hrw.org/report/2020/08/10/stopping-killer-robots/country-positions-banning-fully-autonomous-weapons-and>.

<sup>86</sup> Human Rights Watch, *Stopping Killer Robots*.

humanitarian law is sufficient to regulate the use of LAWS.<sup>87</sup> Without a UN ban it seems likely that fully autonomous weapons systems lie in our future, for as General Paul Selva, the then Vice-Chairman of the Joint Chiefs of Staff of the US Military noted in evidence to the US Senate in 2017: ‘the speed and accuracy of command & control and the capabilities that advanced robotics might bring to a complex battlespace — particularly machine to machine interaction in space and cyberspace where speed is of the essence — can lead to the conclusion that humans are just too slow to make decisions in those domains.’<sup>88</sup> In other words, once one state decides to deploy fully autonomous weapons others will have to follow to countermand the massive strategic benefits – essentially version 2.0 of the Mutually Assured Destruction military strategy of the Cold War.

Selva argues that to prevent what he calls the Terminator Conundrum we must keep humans in the loop. However, he (and others focusing on fully autonomous systems) miss the point. It is not future fully autonomous weapons we should be worried about. It is current semi-autonomous weapons and their increasing deployment to the battlefield. Earlier we discussed three ways in which datafication affected human decision-making and autonomy: (1) Assisted Decision-Making; (2) Supplementary Decision-making and (3) Autonomous Decision-making.

There I made the argument that all three systems undermined human agency and human autonomy. Even with the most basic assisted decision-making system the presentation of data, decided by unseen algorithms, affects our ability to make fully autonomous decisions. There we focused on the Filter Bubble, or digital Balkanisation. The same is also true of Human in the Loop and Human on the Loop systems. Today human actors on the battlefield are being reduced to data points which are then presented to human operators as binary information – viable or nonviable targets. Ultimately a human ‘pulls

<sup>87</sup> Human Rights Watch, *Stopping Killer Robots*.

<sup>88</sup> Colin Clark, ‘VCJCS Selva Says US Must Not Let Robots Decide Who Dies’, *Breaking Defense*, 18 July 2017: <https://breakingdefense.com/2017/07/vcjscs-selva-us-must-not-let-robots-decide-who-dies-supports-lrso/>.

the trigger' or allows the trigger to be pulled, but they do so based on pre-selected data supplied by the weapon system.

In 2014 we learned from an NSA whistle-blower that the US Joint Special Operations Command's (JSOC) High Value Targeting task force were authorising drone strikes based on the activity of SIM cards. As the whistle-blower revealed: 'We're not going after people – we're going after their phones, in the hopes that the person on the other end of that missile is the bad guy.'<sup>89</sup> The whistle-blower estimated that 90% of the drone strikes in Afghanistan relied on the NSA's phone-tracking technology, but claimed that too often the wrong people were killed.<sup>90</sup> The programme was confirmed later by General Michael Hayden, former director of the NSA and the CIA who said, 'We kill people based on metadata.'<sup>91</sup>

At each turn the process of datafication causes human decision-making, human agency, and human autonomy to be sacrificed on the altar of big data. The value of a person is reduced to the data which are gathered, processed, and then acted upon by algorithms. Whether it is the mundane filtering of which of our friends' Facebook posts is favoured by the algorithm, or which restaurant is favoured by Google; to the important such as whether an applicant qualifies for a bank loan or where to invest public funds; to the fundamental such as recommending an offender be given bail or recommending the release of an offensive kinetic weapons system – we all individually and collectively are dehumanised into data points, fed into algorithms, and processed. The resulting data is then fed to human operatives, not to question or challenge, but to act upon.

<sup>89</sup> Jeremy Scahill and Glenn Greenwald, 'The NSA's Secret Role in the US Assassination Program', *The Intercept* 10 February 2014: <https://theintercept.com/2014/02/10/the-nsas-secret-role/>.

<sup>90</sup> Jeremy Scahill and Glenn Greenwald, 'The NSA's Secret Role in the US Assassination Program'.

<sup>91</sup> David Cole, 'We Kill People Based on Metadata' *The New York Review* 10 May 2014: <https://www.nybooks.com/daily/2014/05/10/we-kill-people-based-metadata/>.

## 4. AUTONOMY AND THE DATA HUMAN

The Israeli Law Professor Michal Gal wrote about how datafication affected autonomy in 2018.<sup>92</sup> She points out that there are three rationales for autonomous choice. The first are the efficiency rationales. These argue that individuals know better than anyone else what is best for them. Therefore, the decisions they make are likely to be efficient from their point of view. This rationale focuses on the result of the act of choice, rather than on the act itself.<sup>93</sup> The second are the psychological rationales. These argue on the values inherent in the act of choice itself, regardless of the efficiency of the result. Conscious and subconscious conceptions of identity are shaped through choice, offering us opportunities to define and enhance our self-image, creating personal meaning and responsibility, and potentially increasing our happiness.<sup>94</sup> Finally, there are liberty rationales. Under this rationale, the act of choosing, in itself – not just having the ability to choose or enjoying the objects of our choices – is intrinsically valuable.<sup>95</sup> As John Stuart Mill famously argued, a person's ‘...own mode of laying out his existence is the best, not because it is the best in itself, but because it is his own mode.’<sup>96</sup> Gal believes, as do I, that algorithmic agents and processes affect all three.<sup>97</sup>

Regarding efficiency rationales, she acknowledges that algorithms may in the first instance increase efficiency especially in those cases where the algorithm reaches the same decision as the user, but simply in a more efficient way. However, she warns of a number of risks. Algorithms, in acting efficiently, may narrow markets by pushing users to make the same choice, perhaps favouring incumbents. Further she warns that algorithms may lead to digital echo chambers – a problem acknowledged earlier where the information one receives simply echoes

<sup>92</sup> Michal S. Gal, ‘Algorithmic Challenges to Autonomous Choice’ 25 *Michigan Telecommunications & Technology Law Review* 59 (2018).

<sup>93</sup> Michal S. Gal, ‘Algorithmic Challenges to Autonomous Choice’, 76.

<sup>94</sup> Michal S. Gal, ‘Algorithmic Challenges to Autonomous Choice’, 77.

<sup>95</sup> Michal S. Gal, ‘Algorithmic Challenges to Autonomous Choice’, 79.

<sup>96</sup> John Stuart Mill, *On Liberty* (OUP, 2015), Chapter III.13

<sup>97</sup> Michal S. Gal, ‘Algorithmic Challenges to Autonomous Choice’, 80-91.

one's views, thereby indirectly limiting one's ability to change their mind.<sup>98</sup>

Regarding psychological rationales, she reports that freeing users from the need to engage in some otherwise burdensome decisions may increase their well-being and that more free time may increase innovation and creativity. But again, this is not without risk. She notes that employment of the algorithm could reduce people's ability to define themselves through their choices. The user's identity will instead be shaped by the choices made through the algorithm, producing beliefs, goals, perceptions, and interactions other than those which would have been created by the individual.<sup>99</sup> As Professor Richard Ford notes: 'Over time, one could say that rather than the computer profile reflecting my tastes, I reflect its tastes.'<sup>100</sup>

Finally, in addressing liberty rationales Gal argues that 'Positive freedom requires that people be able to act on their authentic or rational will. It can be argued that deferring to algorithms is generally compatible with positive freedom. The voluntary and informed decision to implement the algorithm is, in itself, an act of choice.' However, as Gal observes, to exercise positive freedom the user should be aware of his self-inflicted limitations on choice, in particular the technological limitations of the algorithm and the parameters used by it to make the choice.<sup>101</sup> This is what Professor Frank Pasquale and others call the Black Box problem<sup>102</sup> or what Barocas and Nissenbaum call the transparency paradox.<sup>103</sup> Providing the level of detail needed to enable

<sup>98</sup> Michal S. Gal, 'Algorithmic Challenges to Autonomous Choice', 84. See also Alessandro Bess, 'Personality traits and echo chambers on Facebook' 65 *Computers in Human Behavior* 319 (2016); Seth Flaxman, Sharad Goel and Justin M. Rao, 'Filter Bubbles, Echo Chambers, and Online News Consumption' 80 *Public Opinion Quarterly* 298 (2016).

<sup>99</sup> Michal S. Gal, 'Algorithmic Challenges to Autonomous Choice', 85.

<sup>100</sup> Richard T. Ford, 'Save the Robots: Cyber Profiling and your So-Called Life' 52 *Stanford Law Review* 1573 (2000), 1577.

<sup>101</sup> Michal S. Gal, 'Algorithmic Challenges to Autonomous Choice', 87.

<sup>102</sup> Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Harvard University Press, 2015).

<sup>103</sup> Solon Barocas and Helen Nissenbaum, 'Big Data's End Run around Anonymity and Consent' in Julia Lane, Victoria Stodden, Stefan Bender and Helen

users to provide genuinely informed consent would overwhelm even savvy users because the decisional parameters are volatile and indeterminate. Finally, Gal simply observes that ‘autonomy is realized through the act of making choices, and if we delegate this responsibility to an algorithm, we *ipso facto* relinquish that autonomy.’<sup>104</sup>

In *The Morality of Freedom* Joseph Raz sets out three conditions that must be satisfied if a particular choice is to count as autonomous: (1) the person must have the mental abilities to form intentions of a sufficiently complex kind, and plan their execution. These include minimum rationality and the ability to comprehend the means required to realize his goals; (2) there must be adequate options available for him to choose from; and (3) choice must be free from coercion and manipulation by others.<sup>105</sup>

It is my belief that when decisions are made for, and equally importantly about, us by algorithms (often without our knowledge), Raz’s three conditions are undermined. While it might be true that we all possess the capacity required by Raz’s first condition, there is a question as to whether we possess the requisite contextual knowledge to exercise that capacity fully. What I mean here is we can never know what information the algorithm has excluded from our view – to steal from the philosopher and lexicographer Donald Rumsfeld, we can never know the unknown unknowns.<sup>106</sup> This occurs on two levels: (1)

Nissenbaum (eds) *Privacy, Big Data, and the Public Good Frameworks for Engagement* (CUP, 2014).

<sup>104</sup> Michal S. Gal, ‘Algorithmic Challenges to Autonomous Choice’, 89.

<sup>105</sup> Joseph Raz *The Morality of Freedom* (OUP, 1986), 373.

<sup>106</sup> As Secretary Rumsfeld records: ‘in early 2002. Toward the end of one of my Pentagon press briefings, a journalist told me that “reports” were suggesting the absence of a link between Saddam Hussein’s regime and terrorists seeking weapons of mass destruction. These unidentified reports, the questioner suggested, were evidence of a lack of a “direct link”. Putting aside the substance of the reporter’s question—at least for the moment—I raised a larger point about the limits of human knowledge. I responded:

Reports that say something hasn’t happened are always interesting to me because as we know, there are known knowns: there are things we know we know. We also know there are known unknowns: that is to say we know there are some things [we know] we do not know. But there are also unknown unknowns—the

we cannot know exactly how a learning algorithm is selecting information to make available to us or how it processes that information – the transparency paradox, and (2) when decisions are made *about* us we do not know when and how decision-making values are programmed – the explainability paradox. We therefore might meet Raz’s first requirement but only in a non-contextual fashion.

The second requirement is complicated by the fact that adequacy is a relative value – how much choice affords ‘adequate options’? There is no doubt technology affords us many more options than we had in the past. When a visitor arrives in a new town and they want to find a place for dinner that night their list of available restaurants is much fuller when selected by Siri on their phone than it would have been in the pre-digital era when they would have had to make do with perhaps the local phone book and the recommendations of people they speak to (or by just walking the streets). We clearly now have more choice. But again the question is not about what can be seen but rather what is hidden from view. The focus is not on the positive presentation of information but rather the negative absence of information. As with the capacity question the challenge is the unknown rather than the known. If the visitor is able to determine what information was not chosen to be presented to them would their decisions have been different? Or in the alternative, would they have presented different information to an algorithm making decisions about them had they known how it would make that decision? This narrowing of actual choice has been discussed by Dogruel, Facciorusso and Stark in their paper *I’m still the master of the machine*.<sup>107</sup> The authors recorded that ‘(survey) participants raised concerns regarding algorithms’ impact on their decision-making and were likely to perceive them as

ones we don’t know we don’t know. And if one looks throughout the history of our country and other free countries, it is the latter category that tends to be the difficult one.’

See Donald Rumsfeld, *Known and Unknown: Author’s Note*: <https://papers.rumsfeld.com/about/page/authors-note>.

<sup>107</sup> Leyla Dogruel, Dominique Facciorusso and Birgit Stark, “I’m still the master of the machine.” Internet users’ awareness of algorithmic decision-making and their perception of its effect on their autonomy’ (2020) *Information, Communication & Society*, DOI: 10.1080/1369118X.2020.1863999.

manipulative and restricting their autonomy’ noting that one respondent recorded that ‘an algorithm patronises me in a way. I am not in charge of the decision-making anymore, or at least only indirectly’ while another responded ‘in the context of advertisements and online shopping, you’re quickly tempted to think that there is nothing else out there than these offers you are seeing.’<sup>108</sup> This demonstrates that algorithmically filtered options are not seen by human users to be the equivalent of open choice options. Returning to Raz we must challenge the adequacy of algorithmically filtered options – there may be more choice, but it might be inadequate all the same. Again, the reason for this are the unknown unknowns – the choices hidden from us by the algorithm.

This all leads to the inevitable conclusion that the process of datafication and algorithmic decision-making is at odds with Raz’s third requirement: that choice must be free from coercion and manipulation by others. The very act of datafication, the reduction of choice to data points is a manipulation of choice factors. As algorithms then process that data in a way different to biological actors, and in a way that cannot always be explained to those biological actors this means that our autonomy is undermined in four different ways when algorithms make decisions for or about us.

1. It is undermined in an internal positive way when it presents us with choices (or in fully autonomous systems takes actions) which are limited by pre-determined values we cannot observe due to the transparency paradox.
2. It is undermined in an internal negative way when it removes choices from consideration without informing us due to pre-selected parameters of values.
3. It is undermined in an external positive way when it makes decisions about us based only on information gathered about us or supplied by us by observation or data requests.
4. It is undermined in an external negative way when it makes decisions about us either with or without our knowledge based upon

<sup>108</sup> Dogruel, Facciorusso and Stark, ‘I’m still the master of the machine’, 12.

limited information about us when other more relevant information may be available from other sources.

As a result, datafication and the use of algorithms affect human agency.

5. WHY DATAFICATION IS BAD FOR THE RULE OF LAW  
(AND HUMANS) AND WHAT TO DO ABOUT IT

The final part of this lecture looks at the implications of datafication, AI, and Machine Learning for lawyers and policymakers, and asks 'are we developing the right approach to the regulation of AI and Machine Learning?'

Until relatively recently there has been little in the way of a policy/legal response to AI and Machine Learning. In general, early discussion of AI and Machine Learning tended to focus on the direct application of AI or Machine Learning to a specific activity or area of law.<sup>109</sup> Unsurprisingly there has been extensive discussion of tort liability when AI or Machine Learning goes wrong.<sup>110</sup> The earliest examples of such discussions date to the late 1980s and the use of expert systems (a type of IFTTT) but more recently a vibrant debate has developed around where tort liability may lie in a Machine Learning system where the decisions are made inside a black box. There have also been extensive discussions around AI/Machine Learning and intellectual property rights. So, when an AI creates a work of art like *Edmond de Belamy*, from *La Famille de Belamy*, which then sells at

<sup>109</sup> See e.g. Richard Susskind, 'Expert systems in law: A jurisprudential approach to artificial intelligence and legal reasoning' (1986) *Modern Law Review* 168; Jessica S. Allain, 'From Jeopardy to Jaundice: The Medical Liability Implications of Dr. Watson and Other Artificial Intelligence Systems' 73 *Louisiana Law Review* 1049 (2013); John Lightbourne, 'Algorithms & Fiduciaries: Existing and Proposed Regulatory Approaches to Artificially Intelligent Financial Planners' 67 *Duke Law Journal* 651 (2017).

<sup>110</sup> George S. Cole, 'Tort Liability for Artificial Intelligence and Expert Systems' (1990) 10(2) *Computer/Law Journal* 127; Bryan Casey, 'Amoral Machines, or: How Roboticians Can Learn to Stop Worrying and Love the Law' 111 *Northwestern University Law Review* 1347 (2017); Ryan Abbott, 'The Reasonable Computer: Disrupting the Paradigm of Tort Liability' 86 *George Washington Law Review* 1 (2018).

auction for \$432,500, who owns the copyright? So far, as authors like James Grimmelmann have pointed out, AI/Machine Learning is little advanced from any other writing or research tool like a pen, a computer, or a camera.<sup>111</sup> We have strong principles in place that the product of an AI/Machine Learning process is owned (in copyright terms) by the person who made the arrangements for the work to be produced. AI/Machine Learning does not simply ‘go off on a frolic of its own’; it is not creative in the human sense – it does not wake up at 5 am thinking ‘I’ve just had a great idea’, it produces work to order. As Grimmelmann observes, ‘It is possible that some future computer programs could qualify as authors. We could well have artificial intelligences that are responsive to incentives, unpredictable enough that we can’t simply tell them what to do, and that have attributes of personality that make us willing to regard them as copyright owners. But if that day ever comes, it will be because we have already made a decision in other areas of life and law to treat them as persons, and copyright law will fall in line.’<sup>112</sup>

There are many more academic and policy debates about AI and autonomous weapons,<sup>113</sup> AI and transportation,<sup>114</sup> AI and contracting,<sup>115</sup>

<sup>111</sup> James Grimmelmann, ‘There’s No Such Thing as a Computer-authored Work’ (2016) 39 *Columbia Journal of Law & Arts* 403.

<sup>112</sup> Grimmelmann, ‘There’s No Such Thing as a Computer-authored Work’, 414.

<sup>113</sup> Anderson, Reisner and Waxman, ‘Adapting the Law of Armed Conflict to Autonomous Weapon Systems’; Anderson and Waxman, ‘Law and Ethics for Autonomous Weapon Systems: Why a Ban Won’t Work and How the Laws of War Can’; Sparrow, ‘Robots and Respect: Assessing the Case against Autonomous Weapon Systems’; Ma, ‘Autonomous Weapons Systems under International Law’.

<sup>114</sup> Ivo Coca-Vila, ‘Self-driving Cars in Dilemmatic Situations: An Approach Based on the Theory of Justification in Criminal Law’ (2018) 12 *Criminal Law and Philosophy* 59; Jan De Bruyne and Jarich Werbrout, ‘Merging self-driving cars with the law’ (2018) 34 *Computer Law & Security Review* 1150; Maurice Schellekens, ‘Self-driving cars and the chilling effect of liability law’ (2015) 31 *Computer Law & Security Review* 506.

<sup>115</sup> Farshad Ghodoosi, ‘Contracting in the Age of Smart Contracts’ 96 *Washington Law Review* 51 (2021); Mark Giancaspro, ‘Is a ‘smart contract’ really a smart idea? Insights from a legal perspective’ (2017) 33 *Computer Law & Security Review* 825; Giovanni Sartor, *Cognitive Automata and the Law* EUI working paper LAW No. 2006/35: <https://cadmus.eui.eu/bitstream/handle/1814/6423/LAW-2006-35.pdf>.

and AI and the dispensation of justice.<sup>116</sup> However, Grimmelmann has really hit the point here – talking about AI as an abstract development in a particular area of policy or practice is to fall into what Judge Frank Easterbrook called the ‘Law of the Horse’ in 1996;<sup>117</sup> as Grimmelmann observes ‘It is a special case of a more general problem, one that presents few new twists on familiar issues.’<sup>118</sup>

As, though, this lecture has demonstrated the changes implemented by AI and Machine Learning are structural not specific. They have the capacity to undermine our systems of communication, politics, media, and culture, but above all else our system of law and the rule of law. This is because at a fundamental level the use of AI and Machine Learning to supplement, assist, or in time replace, human decision-making is a change to human autonomy, and through that to thought, society and ultimately law as a product of these.

Therefore, it is positive to see more recently moves to develop standards and regulation for AI and Machine Learning at a general rather than a specific level. There have been to date a number of global initiatives. In December 2018, Canada and France announced plans for a G7-backed International Panel on Artificial Intelligence, modelled on the International Panel on Climate Change, to study the global effects of AI on people and economies and to steer AI development.<sup>119</sup> The panel is not empowered to propose regulations or laws for AI, instead being asked to:

<sup>116</sup> Crotoof, “‘Cyborg Justice’ and the Risk of Technological Legal Lock-in”; John Morison and Adam Harkens, ‘Re-engineering justice? Robot judges, computerised courts and (semi) automated legal decision-making’ (2019) 39 *Legal Studies* 618; Richard Susskind, *Online Courts and the Future of Justice* (OUP, 2019).

<sup>117</sup> Frank H. Easterbrook, ‘Cyberspace and the Law of the Horse’ (1996) *University of Chicago Legal Forum* 207.

<sup>118</sup> Grimmelmann, ‘There’s No Such Thing as a Computer-authored Work’, 415.

<sup>119</sup> Government of France, *France and Canada create new expert International Panel on Artificial Intelligence*, 7 December 2018: <https://www.gouvernement.fr/en/france-and-canada-create-new-expert-international-panel-on-artificial-intelligence>. See further <https://www.gpai.ai/>.

1. support and guide the responsible development, use and adoption of AI that is human-centric and grounded in human rights, inclusion, diversity and innovation, while encouraging economic growth;
2. facilitate international collaboration in a multistakeholder manner; and
3. monitor and draw on work being done domestically and internationally to identify gaps, maximize coordination and facilitate international collaboration on AI.<sup>120</sup>

The only time law is mentioned in the *Declaration of the International Panel on Artificial Intelligence* is to say that the rule of law should be promoted and protected, but as I have argued, the rule of law is fundamentally undermined by the very existence and use of AI.

In May 2019 the OECD Principles on AI were adopted.<sup>121</sup> These propose five complementary values-based principles for the responsible stewardship of trustworthy AI:

1. AI should benefit people and the planet by driving inclusive growth, sustainable development and well-being.
2. AI systems should be designed in a way that respects the rule of law, human rights, democratic values and diversity, and they should include appropriate safeguards to ensure a fair and just society.
3. There should be transparency and responsible disclosure around AI systems to ensure that people understand AI-based outcomes and can challenge them.
4. AI systems must function in a robust, secure and safe way throughout their life cycles and potential risks should be continually assessed and managed.
5. Organisations and individuals developing, deploying or operating AI systems should be held accountable for their proper functioning in line with the above principles.

In June 2019 the G20 AI Principles were published.<sup>122</sup> These com-

<sup>120</sup> Government of Canada, *Declaration of the International Panel on Artificial Intelligence*: <https://www.canada.ca/en/innovation-science-economic-development/news/2019/05/declaration-of-the-international-panel-on-artificial-intelligence.html>.

<sup>121</sup> <https://www.oecd.org/going-digital/ai/principles/>.

<sup>122</sup> <https://www.g20-insights.org/wp-content/uploads/2019/07/G20-Japan-AI-Principles.pdf>.

mitted the G20 to support ‘an enabling policy environment for human-centred AI’; to ‘respect the rule of law, human rights and democratic values, throughout the AI system lifecycle’ and to ‘commit to transparency and responsible disclosure regarding AI systems’.

By now a theme should be emerging. To regulate AI, principles will be adopted, these principles will be of an extremely high level and will focus on inclusiveness, fairness, transparency and a respect for human rights and the rule of law. These are ethical frameworks, not legal or regulatory ones. Into this developing global framework, the EU made its first contribution in April 2019 when it published its *Ethics Guidelines for Trustworthy Artificial Intelligence*.<sup>123</sup> As the name suggests this is a set of ethical principles for the development and deployment of AI. At its heart there are three principles:

1. AI should be lawful, complying with all applicable laws and regulations;
2. AI should be ethical, ensuring adherence to ethical principles and values; and
3. AI should be robust.

It is positive to have finally a focus on lawful AI, not only ethical AI; however, when the Guidelines say AI should be lawful, they mean ‘all legal rights and obligations that apply to the processes and activities involved in developing, deploying and using AI systems remain mandatory and must be duly observed.’<sup>124</sup> In short, to be ethical your development and deployment of AI must not be illegal. There is no suggestion here of the need for new laws or regulations specifically tailored to AI.

Finally, in February 2020 the EU published their key AI strategy document – the *White Paper on Artificial Intelligence – A European*

<sup>123</sup> Independent High-level Expert Group on Artificial Intelligence, *Ethics Guidelines for Trustworthy Artificial Intelligence* (European Commission, 2019): <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.

<sup>124</sup> *Ethics Guidelines for Trustworthy Artificial Intelligence*, 6.

*approach to excellence and trust*.<sup>125</sup> For the first time we get a fully developed regulatory proposal for AI and Machine Learning. The paper is in two sections, an ‘ecosystem of excellence’ and a ‘ecosystem of trust’. The latter outlines the EU’s approach for a regulatory framework for AI. This document for the first time makes significant proposals for the regulation of AI and Machine Learning. This proposes a risk-based approach to the regulation of AI with regulation being targeted at ‘high-risk’ technologies. High-risk technologies are those which meet two cumulative criteria:

1. the AI application is employed in a sector where, given the characteristics of the activities typically undertaken, significant risks can be expected to occur. These would be sectors such as healthcare, transport, energy, and parts of the public sector.
2. the AI application in the sector in question is, in addition, used in such a manner that significant risks are likely to arise. This recognises that scheduling software in healthcare is not high risk whereas MRI pattern recognition for cancerous tissue is.

Where AI/Machine Learning is viewed as high risk the White Paper recommends that such technologies are made subject to the proposed future regulatory framework for AI. This will set out systems for supervision and transparency in the programming, training, and accuracy of the AI/Machine Learning systems. There remain two weaknesses with the White Paper, indeed with all international efforts to regulate AI/Machine Learning. The first is over-reliance on ethical principles to the exclusion of legal-regulatory frameworks and the second is the risk-based approach proposed in the White Paper.

Perhaps not unsurprisingly given the massive commercial possibilities of AI and Machine Learning and the current state of development of the technologies, States and technology companies are pressing for the flexibility of an ethical rather than a legal regulatory framework. In a recent position paper fourteen EU Member States called upon

<sup>125</sup> European Commission, *White Paper on Artificial Intelligence – A European approach to excellence and trust*, February 2020: [https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020\\_en.pdf](https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf).

the EU to ‘avoid setting burdensome barriers and requirements which can be a hindrance for innovation. Instead, we should incentivise AI developers and deployers to proactively and systematically promote trustworthy AI for the benefit of our society, citizens and economy.’<sup>126</sup> This is remarkable given that as European Digital Rights (EDRi) observe, the current EU AI Framework ‘draws absolutely no red lines for even the most harmful of uses of AI that are recognised across civil society globally.’<sup>127</sup> This over-reliance on ethical standards to the exclusion of legal-regulatory standards is something that elsewhere Professor Julia Black and I have found to be worrisome.

In a retreat from regulation many have proposed the answer to AI governance may be found in ethics. This should be resisted. It leads to lists of desiderata of good behaviour, causes the law, including regulation, to be marginalised in the debate in favour of a focus on these soft forms of governance. This ‘ethics washing’ leads to significant problems. Where companies have a voluntary, ethical commitment in tension with a legal, commercial duty, it is not difficult to see why compliance with the legal duty wins out. Furthermore, the data and AI profession lacks key characteristics of professions in which a soft governance approach works – there are no longstanding norms of good behaviour, no well-established methods for translating principles into practice, and no licensing body.<sup>128</sup>

The focus on risk-based regulation is also incorrect. The technology is currently nascent, the best way to think about it is as being as the internet was in the 1990s and early 2000s. If we were asked to draft a risk-based approach to internet regulation and governance at that point how many people would have picked out Myspace.com (the forerunner to Facebook) and Amazon.com (then a bookshop) as

<sup>126</sup> *Innovative and Trustworthy AI: Two Sides of the Same Coin*. Position paper on behalf of Denmark, Belgium, the Czech Republic, Finland, France Estonia, Ireland, Latvia, Luxembourg, the Netherlands, Poland, Portugal, Spain and Sweden <https://em.dk/media/13914/non-paper-innovative-and-trustworthy-ai-two-side-of-the-same-coin.pdf>.

<sup>127</sup> EDRi, *Attention EU regulators: we need more than AI “ethics” to keep us safe*, 21 October 2020: <https://edri.org/our-work/attention-eu-regulators-we-need-more-than-ai-ethics-to-keep-us-safe/>.

<sup>128</sup> Julia Black and Andrew Murray, ‘Regulating AI and Machine Learning: Setting the Regulatory Agenda’ *European Journal of Law and Technology*, Vol 10, Issue 3, 2019: <https://ejlt.org/index.php/ejlt/article/view/722/980>,

among the riskiest technologies most in need of regulation? Beyond this there is a basic fallacious assumption that some AI or Machine Learning is riskier than others. Based on this we would no doubt list Lethal Autonomous Weapons, Autonomous Transportation, Medical Diagnostic Tools, and some Public Sector Systems as higher risk. But as this lecture has demonstrated the risk is not the output effect of the AI but rather the input effect it has on our autonomy. Siri, Alexa, and all similar AI assistive systems, as well as basic ranking algorithms in search and social media are every bit as risky to our society and our rule of law, perhaps given their ubiquity more so than so-called high-risk systems.

Finally, we must address a fundamental, normative, truth. At the outset of this lecture, we spent considerable time addressing the question ‘what is law?’ At the time this may have seemed a bit remote to the subject of the lecture but through this we established a vital, incontrovertible fact – that law is consciously created and is the distillation of the collective agency of a society, group, or culture. And from this the hopefully equally incontrovertible claim that the rule of law is the ultimate distillation of this principle: the clear spirit of human choice in the purest form. But beyond this in our journey through the development of jurisprudential thought we also charted the movement from natural law to legal positivism. We observed that positivism unlike natural law does not view law as an extension of morality or ethics. Post-Austin we have observed that while morality may be at the root of some laws, law does not have to be moral to be law. To return to my earlier words:

As our world became more complex, more administrative, as human experience and human interaction becomes more complex, we move away from the moral root of law and theology. We become more secular as simple theological morality ebbs and more complex ethical frameworks develop – the world is less black and white and more shades of grey. The same is true of law, Austin fires a starting pistol but now we see that law is more abundant than the natural or moral law principles that regulated criminality, property, and family. Law is now corporate, administrative, permissive, and restrictive. It is sometimes paternalistic and at other times libertarian.

This is the fundamental normative problem of promoting ethical standards and ethical values for the regulation of AI. They are natural law systems for a world which has evolved beyond natural law. We recognised in the 19<sup>th</sup> Century that our world had evolved beyond morals into a complex post-industrialist world where legal positivism was required. Yet as every new technology and every new system is developed the first call from industrialists and governments is that we should adopt ethical standards. This is like asking Plato or Cicero to design our modern legal system. We are adopting only part of the regulatory palette and are ignoring the possibilities of positive legal-regulatory regimes.

We cannot let ethics substitute for legal positivism with AI. AI and Machine Learning will likely be the biggest changes we all experience in our lifetimes. Already as we have seen they have key implications for human agency and human thought, but we are just at the start.

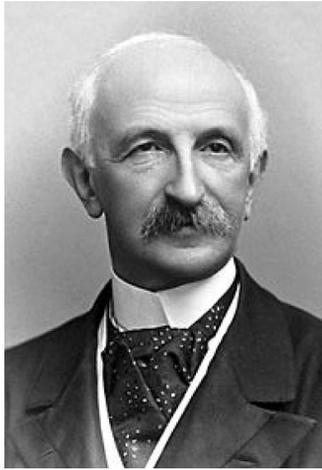
If we are to seek to control the way corporations and governments use AI and Machine Learning, then ethics cannot substitute for law or other forms of formal regulation. Unlike academic proposals, new regulatory regimes rarely land newly minted, in perfect form and onto a blank canvas: they are always situated in an existing context often thick with existing norms and rules, with existing organisational structures, and amongst actors with particular behaviours, cognitive frameworks, capacities and motivations. This is at a minimum a call for lawyers, and for regulators more generally, to get involved in the debate and to drive the discussion on from ethical frameworks to legal/regulatory frameworks and how they might be designed.<sup>129</sup> There are also risks that if we leave it to existing regimes to respond then we will end up not with a coherent system but with patchwork regulation in which there are overlaps and underlaps, with conflicting goals and logics. We need global leadership on this. The challenge of AI Regulation and Governance is a global one – just as we have the International Telecommunication Union and the International Civil Aviation Organization, we must have an International Office for AI, tasked with

<sup>129</sup> Black and Murray, 'Regulating AI and Machine Learning: Setting the Regulatory Agenda'.

the development of a positive, international, legal framework for the development and deployment of AI. This is urgent. For in twenty years the technology will be ambient, and we will have missed our chance at meaningful, modern, positive regulation of AI.

THE ANNUAL T.M.C. ASSER LECTURE ON THE  
DEVELOPMENT OF INTERNATIONAL LAW

*A Mission for Our Time*



INTRODUCTION

The Annual T.M.C. Asser lecture has been established in honour of the Dutch jurist and Nobel Peace Prize Laureate, Tobias Michael Carel Asser (Amsterdam, 28 April 1838 – The Hague, 29 July 1913), and his significant contributions to the development of public and private international law. It is the T.M.C. Asser Instituut's flagship lecture and its date commemorates the foundation of the Institute in December 1965.

MISSION

Tobias Asser was a man with a vision. A man who kept his finger on the pulse of his time, and who managed to shape the legal develop-

ments during his days.<sup>1</sup> In his Inaugural Address upon the acceptance of his professorship at the University of Amsterdam in 1862, Asser explained that it was his ‘vocation’ to reflect on commercial law and its ‘import’, while ‘taking into consideration the condition of society in [his] century’.<sup>2</sup> What we learn from his lecture extends beyond the field of commercial law; it shows Asser’s view of the law more generally: ‘law serves primarily to cultivate trust’.<sup>3</sup>

For its mission statement, the Annual T.M.C. Asser Lecture builds on the vision and mission of the man who has lent it his name. It invites distinguished international lawyers to take inspiration from Asser’s idea of cultivating trust and respect through law and legal institutions, and to examine what it could mean in their area of expertise today.

Current legal scholarship has uncovered the complications of Asser’s mission, and of his internationalist friends and colleagues.<sup>4</sup> It has pointed to the downside of how the international legal order took shape in spite of the good intentions of these late 19<sup>th</sup> and early 20<sup>th</sup> century liberal-humanitarian internationalists. Asser himself was well aware of the dangers of utopian idealism<sup>5</sup> on the one hand, and the dangers of a nationalistic conservative attitude towards international law on the other. Every age has different needs and pitfalls and hence, sailing between commitment and cynicism,<sup>6</sup> every age requires a different course.

<sup>1</sup> A Eyffinger, *T.M.C. Asser [1838–1913] Founder of The Hague Tradition* (The Hague: Asser Press, 2011), p. 11.

<sup>2</sup> The Inaugural Address is included in E.M.H. Hirsch Ballin (ed. and intro.), *A Mission for his Time. Tobias Asser’s Inaugural Address on Commercial Law and Commerce, Amsterdam 1862* (The Hague: Asser Press, 2012), p. 18.

<sup>3</sup> *Ibid.*, p. 22.

<sup>4</sup> See below ‘Tobias Asser in context: One of the ‘Men of 1873’.

<sup>5</sup> At the Second Hague Peace Conference, Asser himself said ‘you know I am not a Utopian’, Eyffinger, p. 5, n. 45.

<sup>6</sup> M Koskenniemi, ‘Between Commitment and Cynicism: Outline for a Theory of International Law as Practice’, in *Collection of Essays by Legal Advisors of States, Legal Adviser of International Organizations and Practitioners in the field of International Law* (United Nations, NY, 1999), pp. 495–523; also available online.

Our time, too, is in dire need of reflection. It is marked by the politics of fear, domestically as well as globally. In different ways ‘fear operates directly as a constitutive element of international law and the international ordering and decision-making processes.’<sup>7</sup> Taking note of Tobias Asser’s legacy in this context, a reorientation of the international order towards an order based on respect and trust urges itself upon us.<sup>8</sup>

Today, with international lawyers perhaps sadder and wiser, it seems more than ever to be an international lawyer’s task to examine – as Asser did in his day – how to respond to ‘the condition of society’. Mutual trust and respect are crucial to the health of any heterogeneous society, whether it is the international society or one of the rapidly growing cities across the globe. A (research) question which Tobias Asser bequeathed to us is ‘how can law serve this aim?’

In spite of well-known complications and dark sides,<sup>9</sup> in this context the Rule of Law and the principles of human rights are paramount. These may provide direction in our considerations about trust and respect in relation to challenges brought by, for example, globalisation, urbanisation, (global) migration, the atomisation of society, climate change, environmental degradation, the complexity of the traditional North-South divide, the dangers of a renewed international arms race, and the dilemmas of new global actors such as the EU.

Against this backdrop, the Annual T.M.C. Asser Lecture aspires to be a platform for a constructive, critical reflection on the role of law in dealing with the challenges and (potentially radical) changes of the global society of the 21<sup>st</sup> century.

<sup>7</sup> D. Joyce & A. Mills, ‘Fear and International Law’, *Cambridge Review of International Affairs*, 19:2 (2006), pp. 309–310.

<sup>8</sup> A. Carty, ‘New Philosophical Foundations for International Law: From an Order of Fear to One of Respect’, *Cambridge Review of International Affairs*, 19:2 (2006), pp. 311–330; also, J.E. Nijman, ‘Paul Ricoeur and International Law: Beyond ‘The End of the Subject’. Towards a Reconceptualization of International Legal Personality’, *Leiden Journal of International Law*, 20 (2007), pp. 25–64.

<sup>9</sup> D. Kennedy, *The Dark Sides of Virtue* (Princeton: PUP 2004); also, M. Koskenniemi, *The Gentle Civilizer*, infra note 21, and *The Politics of International Law* (Oxford: Hart 2011).

## BACKGROUND

In Asser's time, the cultivation of trust and respect in international relations was indeed an urgent matter. Asser's professional life spans from the second half of 'the long 19<sup>th</sup> century'<sup>10</sup> up to the eve of the First World War. It was a time of rising nationalism and mounting 'distrust and despair'<sup>11</sup> in Europe. The 19<sup>th</sup> century Eurocentric world order was to collapse only a few years after Asser's death.

In Asser's lifetime America had experienced the Civil War (1861–65) and slavery was abolished after a slow struggle. In Europe, the Crimean War (1853–56) and the Franco-Prussian War (1870–71) brought decades of peace in Europe to an end. With these wars the horrors of industrial warfare began and forever changed the destructive scale and intensity of armed conflict. In Asia, Britain and France forced China, by military means, to open up its markets for opium, on the basis of what they argued to be their sovereign right to free trade, even against the imperial government's desperate attempt to protect its dwindling population from opium addiction. A socialisation into international society and law that was to leave its mark on China's approach to international law well into our time.<sup>12</sup> In the latter days of his career, Asser actively supported the International Opium Conference (1912) to end the opium enslavement of the Chinese people.<sup>13</sup>

With the economic policies of the late 19<sup>th</sup> century the European empires spurred on the process of modern globalisation in the industrial era. Asser had a keen interest in economics and as the head of a (commercial) law practice for most of his life,<sup>14</sup> he is likely to have been especially sensitive to the process. In his view, transnational trade and commerce were crucial for societies to thrive and develop peace-

<sup>10</sup> Eric Hobsbawm's term for the period 1789–1917.

<sup>11</sup> Eyffinger, p. 67.

<sup>12</sup> S. Suzuki, 'China's Perceptions of International Society in the Nineteenth Century: Learning more about Power Politics?', 28 *Asian Perspective* (2004), pp. 115–144.

<sup>13</sup> Eyffinger, p. 79.

<sup>14</sup> Among his clients, though, were the heirs of King Leopold in the Congo heritance.

fully. In that sense, his perspective on free trade and commerce was utilitarian – in the service of ‘public welfare’.<sup>15</sup> Hence, his stance was not uncritical; transnational trade and commerce facilitated by law and legal institutions were to serve peace and justice, but not to exploit or violate ‘the inalienable rights of a free people’.<sup>16</sup>

The urbanisation of 19<sup>th</sup> century Europe prefigures that of today; it basically put much of the current global city system in place. Asser was outspoken about his love for the ‘distinguished mercantile city’ of Amsterdam: ‘[u]nder any circumstances, wherever my place of domicile, I will forever remain an Amsterdammer!’<sup>17</sup> His love of Amsterdam, however, not only sprung from the city’s tradition of international trade and commerce, but also and even more so from its tradition of openness to strangers and providing a refuge for the expelled. Being a Dutch citizen of Jewish descent, the exclusion and violence brought about by anti-Semitism in European (urban) societies must have been a matter of personal concern for someone so eager to participate in the public sphere. Nationalism, a growing sentiment in Europe, was completely alien to Asser. With his urban cosmopolitan mind-set, his thinking was transnational by nature. His vision of international and personal relations did not hinge upon fear and othering, but rather upon respect and trust.

For Asser, the role of law was vital to the emancipation of the Jewish minorities in Europe, as was the case for any minority. He worked with an integral view of the Rule of Law, to be strengthened as much in the domestic as in the international society. Asser’s dedication to citizens’ rights and the principle of legal equality is visible, for example, in his advocacy of equal voting rights for women.<sup>18</sup>

While Asser’s vision of law and legal institutions was all about the ideals of peace, prosperity and justice, he was concrete and prag-

<sup>15</sup> Hirsch Ballin, p. 19.

<sup>16</sup> *Ibid.*, p. 33.

<sup>17</sup> Eyffinger, p. 13.

<sup>18</sup> Hirsch Ballin, p. 13.

matic when aiming to shape developments in private and public international law.

Asser's commitment to international trade and commerce as a means to achieve peace and international solidarity inspired his efforts to deal with 'conflict of laws' and to promote a unification and codification of the rules of private international law. In his view, the demands of international life went beyond economic relations only, and so, being the pragmatic lawyer that he was, Asser presided over the Four Hague Conferences on Private International Law (1893–1904) which managed to produce six conventions ranging from procedural law to family law issues.

While international tensions intensified and an arms race was looming, Asser moved into the realm of public international law – albeit with a good share of realism about state conduct and the pursuit of self-interest. Together with Feodor Martens, Asser stood at the helm of the Hague Peace Conferences (1899 and 1907), which focused on international humanitarian law and the peaceful settlement of disputes. The First Conference resulted in the constitution of a Permanent Court of Arbitration (PCA). Being a prominent arbiter himself, Asser participated in the first case before the PCA. Thanks to Andrew Carnegie, who wanted to ensure a 'wise distribution' of his wealth, the Peace Palace was built and The Hague was thus granted its role of *City of Peace and Justice*.

T.M.C. Asser's mission of peace, liberty and justice defined both his academic and diplomatic work. He intended to listen to 'the voice of the conscience of [his] century' and tirelessly applied his legal genius to develop public and private international law. After decades of neutrality, he would moreover steer the Netherlands back into the diplomatic arena and towards a more prominent international position.

Tobias Asser's legacy is almost too vast for one man. No wonder his role was recognized by the Nobel Prize Committee in 1911. The

Committee portrayed Asser as ‘the Hugo Grotius of his day’.<sup>19</sup> Certainly they both aimed to strengthen the Rule of Law in a global society.<sup>20</sup>

In contemporary international legal scholarship, Professor T.M.C. Asser was one of the international lawyers Martti Koskenniemi has famously called the ‘Men of 1873’: twenty to thirty European men who were actively engaged in the development of international law and who, thanks to among others Asser and his dear friend Rolin, established the *Institut de Droit International* in 1873.<sup>21</sup> They were interested in ‘extending the mores of an *esprit d’internationalité* within and beyond Europe. ... [they were the] “founders” of the modern international law profession.’<sup>22</sup>

For the men of 1873, international law was to be social and cultural in a deep sense: not as a mere succession of treaties or wars but as part of the political progress of European societies. They each read individual freedoms and the distinction between the private and the public into constructive parts of their law. If they welcomed the increasing interdependence of civilized nations, this was not only to make a point about the basis of the law’s binding force but to see international law as part of the progress of modernity that was leading societies into increasingly rational and humanitarian avenues.<sup>23</sup>

Their liberal project was a project of reform, human rights, freedom of trade, and ‘civilization’. In their view, ‘jurists should not remain in the scholar’s chamber but were to contribute to social progress.’<sup>24</sup> Koskenniemi further cites Asser to explain the *esprit d’internationalité*:

For Asser, for instance, the tasks of the *jurisconsulte* in the codification of private international law followed “from the necessity to subordinate

<sup>19</sup> See for the Nobel Peace Prize 1911 speech: <[http://www.nobelprize.org/nobel\\_prizes/peace/laureates/1911/press.html](http://www.nobelprize.org/nobel_prizes/peace/laureates/1911/press.html)>.

<sup>20</sup> See Asser’s Address at the Delft Grotius Memorial Ceremony July 4, 1899, p. 41.

<sup>21</sup> Eyffinger; M. Koskenniemi, *The Gentle Civilizer of Nations* (Cambridge: CUP 2002).

<sup>22</sup> *Ibid.*, p. 92.

<sup>23</sup> Koskenniemi, pp. 93–94.

<sup>24</sup> *Ibid.*, p. 57.

interest to justice – in preparation of general rules for the acceptance of governments to be used in their external relations”<sup>25</sup>.

BUILDING ON TOBIAS ASSER’S VISION AND MISSION

The institution of this Annual Lecture is inspired by these ‘Men of 1873’ in general and by Asser’s social progressive, ‘principled’ pragmatism, liberalism, and ‘emancipation from legal traditionalism’ in particular.<sup>26</sup>

Drawing inspiration from the ‘Men of 1873’ is however not without complications. Part of their project was the ‘civilizing mission’, with all its consequences. On the one hand, in the early decades of the 20<sup>th</sup> century these scholars may have been hopeful about decolonisation and lifting developing countries out of poverty. Asser’s own involvement in attempts to end a most ‘embarrassing chapter of Western history’, the Opium Wars, may also be mentioned. On the other hand, international law as an instrument of civilisation has surely shown its dark sides. Today, more than ever before, we are aware of how internationalism and the Rule of Law have been the handmaidens of (economic, legal) imperialism.<sup>27</sup> Scholars have pointed to the ‘double standards’ as ‘an integral part of the ideology of democracy and the rule of law’ so visible in the application of international law even today.<sup>28</sup>

The rich and somewhat complex heritage of internationalism does not leave room for naïve ideas about international law as an instrument only for the good of liberal-humanitarian reform; if ‘[l]egal internationalism always hovered insecurely between cosmopolitan humanism and imperial apology... [and i]f there is no perspective-

<sup>25</sup> Ibid., pp. 57–58.

<sup>26</sup> Hirsch Ballin, pp. 12 and 2.

<sup>27</sup> E.g. A. Anghie, *Imperialism, Sovereignty, and the Making of International Law* (Cambridge: CUP, 2005).

<sup>28</sup> A. Carty, ‘The terrors of freedom: the sovereignty of states and the freedom to fear’, in J. Strawson (Ed.) *Law after Ground Zero* (London: Glasshouse Press, 2002), pp. 44–56.

independent meaning to public law institutions and norms, what then becomes of international law's universal, liberating promise?<sup>29</sup>

While for some this rhetorical question marks the end-point of possible legal endeavours, the Annual T.M.C. Asser Lecture hopes to be a place for reflecting critically on what lies *beyond* this question. As Koskenniemi points out, '[i]n the absence of an overarching standpoint, legal technique will reveal itself as more evidently political than ever before.'<sup>30</sup> And so, since '[i]nternational law's energy and hope lies in its ability to articulate existing transformative commitment in the language of rights and duties and thereby to give voice to those who are otherwise routinely excluded', we ask: What does the *esprit d'internationalité* mean today and what could it mean in and for the future?

PROF DR JANNE E. NIJMAN

*Chair of the Executive Board and Academic Director  
of the T.M.C. Asser Instituut, The Hague*

<sup>29</sup> Koskenniemi, p. 513.

<sup>30</sup> *Ibid.*, p. 516.



# INTERNATIONAL & EUROPEAN LAW AS A SOURCE OF TRUST IN A HYPER-CONNECTED WORLD

*Contours of the Asser Strategic Research Agenda 2016–2020*

## INTRODUCTION

The T.M.C. Asser Instituut was founded in 1965 as an interuniversity institute for international law in The Hague. Over the past 50 years, the institute has developed into an internationally renowned centre of expertise in the fields of public international law, private international law and European law.

Located in The Hague, the ‘International City of Peace and Justice’, the Asser Institute is the established location where critical and constructive reflection on international and European legal developments takes place. In the vicinity of the many Hague international (legal) institutions, diplomatic missions, and government ministries, the institute exercises strong convening power and attracts legal scholars from around the world to present and test cutting-edge ideas in their respective fields of expertise.

The Asser Institute has a strong tradition in pursuing independent research. The coming years will see the institute build on this research expertise and further strengthen its academic profile whilst fostering its orientation towards fundamental and independent policy-oriented research.

In doing so, the Asser Institute will continue to fulfil the following roles:

- A facilitator for all Dutch Law Schools that wish to collaborate with Asser in research networks and projects and/or in knowledge disseminating activities.

- A vanguard institute for the University of Amsterdam (UvA) in The Hague (for the UvA Law School in general and the Amsterdam Center for International Law (ACIL) in particular).

#### MISSION

The T.M.C. Asser Instituut aims to further the development of international and European law in such a way that it serves a cultivation of trust and respect in the global, regional, national and local societies in which the law operates.

#### CONTOURS OF THE ASSER STRATEGIC RESEARCH AGENDA 2016–2020

Pursuant to the institute's mission, the Asser Strategic Research Agenda (ASRA) 'International & European law as a source of trust in a hyper-connected world' aims to examine how law as one of the social institutions can contribute to the construction and cultivation of trust and trusting relations needed for cooperation in this large and hyper-connected world.

It will guide the further development of the institute's research capacity and it will contribute to further strengthening Asser's intellectual identity and its position at the interface of the world of legal academia and legal practice.

In the ASRA, the Asser Institute's research is structured along three research strands and an architrave. The latter deals with more general conceptual questions about trust, trustworthiness, and trust-building effects of international and European law fostering the overarching, more abstract and loosely defined normative framework. The three strands are separate but mutually interlinked:

- Human Dignity and Human Security in International and European Law
- Advancing Public Interests in International and European Law

- Adequate Dispute Settlement and Adjudication in International and European Law

HUMAN DIGNITY AND HUMAN SECURITY IN INTERNATIONAL  
AND EUROPEAN LAW

If law cannot provide a sense of human dignity and security, it falls short of the cultivation of trust. Upholding the Rule of Law and a generally high level of human rights protection contributes to the development of trust (and, arguably, vice versa). The research strand Human Dignity and Human Security in International and European Law adopts as its normative framework a human rights approach to contemporary global challenges, inter alia in the field of counter-terrorism, international criminal law, international humanitarian law, international trade, environmental protection, European private international law, and the law of EU external relations. It examines what it means to safeguard human dignity – also in relation to human security – in these areas.

ADVANCING PUBLIC INTERESTS IN INTERNATIONAL  
AND EUROPEAN LAW

Both at the European and international level, the dual impact of globalisation and fragmentation has complicated the use of legislation and regulation in safeguarding public interests. Advancing Public Interests in International and European law aims to critically examine how international and European law may further protection of public interests in different areas, ranging from the governance of sports and media in Europe to natural resources, trade, and environmental protection at the international level. Research within this strand will engage with a large set of questions centred on the potential synergies and trade-offs between different public interests and private interests. Possible normative frameworks for reconciling conflicting values are, for example, the principle of proportionality and variants of the constitutional approach.

ADEQUATE DISPUTE SETTLEMENT AND ADJUDICATION  
IN INTERNATIONAL AND EUROPEAN LAW

By effectuating the law – and thus upholding the Rule of Law –, courts, tribunals and other dispute settlement mechanisms provide fairness, security, stability and predictability. All of them values conducive to trust. Courts, tribunals and other dispute settlement mechanisms can perform this function adequately only if they, in turn, are perceived as trustworthy in speaking and enforcing the law. The research strand Adequate Dispute Settlement and Adjudication in International and European Law examines the adequacy of dispute settlement and adjudication in various areas, as diverse as foreign investment and transnational civil and commercial disputes, doping and sports more generally, cross-border civil disputes, international crimes, and classic inter-state relations.

LOOKING AHEAD

Over the period of this research agenda, the institute will:

- Conduct high-quality independent research – both fundamental research and policy-oriented research –, in order to contribute to current academic and policy debates within the scope of the aforementioned research strands.
- Increase its research capacity, especially through the promotion and fostering of PhD research in international and European law.
- Deliver research-based, cutting-edge, high-level policy-oriented meetings, (professional) education modules and public events of knowledge dissemination.
- Intensify – in areas where the institute’s research expertise can be brought to bear – its cooperation and engagement in European and international academic networks, as well as in the national, European and international arenas of policy formation and legal practice.

More information about the Asser Institute’s research & activities can be found on the website: [www.asser.nl](http://www.asser.nl).

## THE ANNUAL T.M.C. ASSER LECTURE SERIES

The Annual T.M.C. Asser Lecture is a platform for a critical, multi-disciplinary and constructive reflection on the role of law in the (potentially radically) changing global society of the 21st century, and a high-level event within the context of our research programme ‘*International & European law as a source of trust in a hyper-connected world*’.

In 2015, Professor Joseph Weiler (President of the European University Institute in Florence, and University Professor at NYU School of Law) delivered the Inaugural Annual T.M.C. Asser Lecture on ‘*Peace in the Middle East: has International Law failed?*’ in which he identified an indeterminacy issue in the legal framework of belligerent occupation that allows for different interpretations. This, according to Weiler, has turned into a political dispute about the facts, for which international law can provide no more than a roadmap.

In 2016, Onora O’Neill, Professor Emeritus of Philosophy at the University of Cambridge and crossbench member of the British House of Lords, spoke about ‘*Accountable Institutions, Trustworthy Cultures*’ and how rules are not enough. The ethics and culture of institutions, international or otherwise, are important for the trustworthiness of these institutions. This is an important argument that still resonates in these days of institutional distrust.<sup>1</sup>

In 2017, Saskia Sassen, Robert S. Lynd Professor of Sociology at Columbia University (NY), discussed the relations between globalisation, economic development and global migration in the lecture entitled ‘*A Third Emergent Migrant Subject Unrecognized in Law: Refugees from “Development”*’. She asked: ‘Is there any role for inter-

<sup>1</sup> O. O’Neill, *Accountable Institutions, Trustworthy Cultures* (The Hague, T.M.C. Asser Press 2017).

national law in the prevention of, and protection against, expulsions caused by the accelerating destruction of land and water bodies?<sup>2</sup>

In 2018, Martti Koskenniemi, Professor of International Law at the University of Helsinki and Director of the Erik Castrén Institute of International Law and Human Rights, gave the lecture '*International Law and the Far Right: Reflections on Law and Cynicism*' in which he critically reflected on the general state of international law, as well as on its role in the rise of the far right.<sup>3</sup>

The Fifth Annual T.M.C. Asser Lecture, in 2019 delivered by Anne Orford, Professor of International Law at Melbourne Law School, was entitled '*International Law and the Social Question*' and placed the social question, the value of solidarity and social justice back on the table of international lawyers.<sup>4</sup>

For more information on the Annual Lecture Series, registration and programme, please go to: [www.asser.nl/annual-lecture](http://www.asser.nl/annual-lecture), or contact [TMCAsserLecture@asser.nl](mailto:TMCAsserLecture@asser.nl)

<sup>2</sup> S. Sassen, *A Third Emergent Migrant Subject Unrecognized in Law: Refugees from 'Development'* (The Hague, T.M.C. Asser Press 2018).

<sup>3</sup> M. Koskenniemi, *International Law and the Far Right: Reflections on Law and Cynicism* (The Hague, T.M.C. Asser Press, 2019).

<sup>4</sup> A. Orford, *International Law and the Social Question* (The Hague, T.M.C. Asser Press, 2019).

## ABOUT THE AUTHOR



Andrew Murray is a Professor of Law and Deputy Head of the Department of Law at the London School of Economics where he researches and teaches in the fields of Internet and new media law, including the laws of digital surveillance and digital privacy.

Professor Murray is a leading thinker on information technology law and regulation, who focuses on regulatory design within Cyberspace and on the protection and the promotion of Human Rights within the digital environment. In 2018/19 he was the specialist advisor to the British House of Lords Communications Committee inquiry 'Regulating in a Digital World'. He has written a number of impactful books, including, *The Regulation of Cyberspace* (2007), *Rethinking the Jurisprudence of Cyberspace* (2018) and *Information Technology Law: The Law and Society* (4<sup>th</sup> edition, 2019).

Andrew Murray on his lecture *Almost Human: Law and Human Agency in the Time of Artificial Intelligence*:

‘Law is about agency – the human capacity to act independently and to make our own free choices. As Jeremy Webber observes, “Law is consciously created” and is the distillation of the collective agency of a society, group, or culture. The rule of law is the ultimate distillation of this principle: the clear spirit of human choice in the purest form.

However, the process of datafication: the reduction of the complexity of the world to data values, threatens the fabric of human agency and the rule of law. Complexity becomes numerical values and choices become mathematical processes. Human brains, less equipped for this form of decision-making, risk being replaced by algorithmic decision-making. Human agency diminishes as Artificial Intelligence ascends.

LawTech replaces lawyers; risk assessments replace actuaries; there is even the possibility of “algorithmic warfare”. In meeting this challenge, the focus to date has been to invest in AI Ethics rather than AI Regulation. The EU Commission, in its recent White Paper, recommended ethical guidelines could be the foundation for a human-centric trust framework in AI. This lecture challenges this normative assumption by identifying the foundations of datafication and addressing its challenge to human agency, and above all the rule of law.’

ISBN 978-90-6704-366-3

