REVIEW ARTICLE

**OPEN**

# Social and juristic challenges of artificial intelligence

Matjaž Perc [1,2], Mahmut Ozer[3,4] & Janja Hojnik[5]

## ABSTRACT

Artificial intelligence is becoming seamlessly integrated into our everyday lives, augmenting our knowledge and capabilities in driving, avoiding traffic, finding friends, choosing the perfect movie, and even cooking a healthier meal. It also has a significant impact on many aspects of society and industry, ranging from scientific discovery, healthcare and medical diagnostics to smart cities, transport and sustainability. Within this 21st century 'man meets machine' reality unfolding, several social and juristic challenges emerge for which we are poorly prepared. We here review social dilemmas where individual interests are at odds with the interests of others, and where artificial intelligence might have a particularly hard time making the right decision. An example thereof is the well-known social dilemma of autonomous vehicles. We also review juristic challenges, with a focus on torts that are at least partly or seemingly due to artificial intelligence, resulting in the claimant suffering a loss or harm. Here the challenge is to determine who is legally liable, and to what extent. We conclude with an outlook and with a short set of guidelines for constructively mitigating described challenges.

## Introduction

A broad body of literature anticipates that in the years to come intelligent objects will overtake more and more jobs that people have traditionally performed, from driving, diagnosing diseases, providing translation services, drilling for oil, to even milking cows, to name just some examples (Russell and Norvig, 2016). In 1999 a British visionary Kevin Ashton coined the term 'Internet of Things' (IoT) to describe a general network of things linked together and communicating with each other as computers do today on the Internet (Araujo and Spring, 2015; Miller, 2015). The connection of objects to the Internet makes it possible to access remote sensor data and to control the physical world from a distance (Kopetz, 2011, p. 301). Data communication tools are changing 'tagged things' into 'smart objects' with sensor data supporting a wireless communication link to the Internet (Weber, 2009, p. 522; Ngai et al., 2008, p. 510; Gubbi et al., 2013, p. 1645; Chabanne et al., 2013). With IoT manufacturers can remotely monitor the condition of equipment and look for indicators of imminent failure outside normal limits (e.g., vibration, temperature and pressure). This means that the manufacturer can make fewer visits, reducing costs and producing less disruption and higher satisfaction for the customer (Wilkinson et al., 2009, p. 539). Remote diagnostics, where complex manufactured

[1] Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, Maribor 2000, Slovenia. [2] Complexity Science Hub Vienna, Josefstädter Straße 39, Vienna 1080, Austria. [3] Ministry of National Education, Ankara 06420, Turkey. [4] Center for Artificial Intelligence and Data Science, Istanbul Technical University, Istanbul 34469, Turkey. [5] Faculty of Law, University of Maribor, Mladinska ulica 9, Maribor 2000, Slovenia. Correspondence and requests for materials should be addressed to M.P. (email: matjaz.perc@um.si)

products are monitored via sensors may not, however, only be important for repairing industrial machines but also for human health, such as remote control of pacemakers (Stantchev et al., 2015). The widespread use of WiFi and 4G enable the communication with smart objects without the need of a physical connection, such as to control customers' home heating and boiler from their mobile or laptop. Mobile smart objects can move around and GPS makes it possible to identify their location (Kopetz, 2011, p. 308). This technology facilitates the development of so-called connected or automated cars that enable the driver automatic notification of crashes and speeding, as well as voice commands, parking applications, engine controls and car diagnosis. It is foreseen that trucks will soon no longer need drivers, as computers will drive them, without the need for rest or sleep. Moreover, each Philips or Samsung TV comes nowadays with an application called 'Smart TV', which consolidates video on demand function, the Internet access, as well as social media applications (Kryvinska et al., 2014). Objects are thus becoming increasingly smart and consequently autonomous.

However, autonomous objects will also cause accidents, invade private space, fail surgeries and fail to diagnose cancer, and even engage in war crimes (Yoo, 2017, p. 443). As autonomous objects will become more and more commonplace on streets, on the skies, in households and in the workplace, their social and legal status will only grow in importance. Considering that autonomous objects are not a matter of "if" but rather of "when" such technology will be introduced, the regulatory dimension might be decisive in this respect, as is the prior identification of socially challenging situations where not only the user, but also others may be adversely affected. If the activity of autonomous objects, and more generally of artificial intelligence, is not properly regulated, it will not be broadly accepted as a more efficient and safe alternative to human controlled objects or human decision making. However, the autonomy we give to machines may render many established legal doctrines obsoleted, and more importantly, affect what we judge to be "reasonable" human activity in the future.

Modern businesses and technological developments thus need to be followed by appropriate regulation that will control the associated hazards and thus enable the industry to flourish. At the same time, regulation has to leave enough flexibility so that law does not restrict technological development. Considering that the industry and the consumers are getting increasingly smart, smart regulatory solutions need to follow (Oettinger, 2015), establishing the right balance between safety, liability and competition on one side, and innovation and flexibility on the other. In this respect, regulatory requirements can either restrict technological development, in particular if liability for potential errors is strict or if taxation encourages human workforce, or boost it, if the standard of liability is set so that safety of computer performance is compared to the safety of certain human activity, such as driving.

In the European Union in particular, there are delicate discussions taking place on who should be competent to set the rules in this respect, Member States or EU institutions. Moreover, it is also important that this regulatory process does not by-pass democratic governance principles and that industry is included in the regulatory process, as well as that self-regulation replaces legislation where possible, so that only general regulatory requirements are set by the public authorities and the market defines the technical solutions (Bräutigam and Klindt, 2015, pp. 100–106; Weber and Weber 2010, p. 23).

In what follows, we will briefly review the basic principles behind the workings of artificial intelligence, then focus on the social and juristic challenges in more detail that can emerge as a result, and finally proceed with conclusions and guidelines as to how these challenges might be successfully overcome.

## Under the hood of artificial intelligence

In addition to smart algorithms, information and data drive artificial intelligence (Kersting and Meyer, 2018). Data mining pervades social sciences, and it enables us to extract hidden patterns of relationships between individuals and groups, thus leading to a more and more seamless integration of machines and algorithms into our everyday lives. Indeed, data science and artificial intelligence are becoming increasingly popular, with applications that range from medical diagnostics and match making to predictions and classification (Xia et al., 2013).

Machine learning, for example, is a subset of artificial intelligence that allows software applications to become more accurate in predicting outcomes without being explicitly programmed to do so. Three types of learning are possible, namely supervised, unsupervised, and reinforcement learning. When supervised learning is considered, the most common approach entails using artificial neural networks, which are inspired by real biological neural networks. For artificial neural networks to learn how to classify the data, the original data set is usually separated into two sets whereby the first set is used for training while the other set is used for testing. Due to its good classification capabilities, artificial neural networks have been used widely in diagnostics, and many attempts have been made to develop different architectures and algorithms to improve the performance of such networks. For example, Erkaymaz and Ozer (2016) investigated the impact of the small-world network topology, introduced by Watts and Strogatz in 1998, on the performance of artificial neural networks with notable improvement.

Conversely, unsupervised learning proceeds without a teacher, also known as self-organisation and a method of modelling the probability density of inputs (Hinton and Sejnowski, 1999). In case of artificial intelligence, and more specifically machine learning, there is thus no training data set based on which the algorithm would learn how to best classify the data. Notable algorithms used in unsupervised learning are hierarchical clustering, k-means methods, autoencoders, deep belief networks, and self-organising maps. In reinforcement learning, there is a feedback between the actions of the algorithm and the maximisation of a reward function. Typically, such algorithms entail agents that try to maximise some notion of cumulative reward, with the focus on finding a balance between exploration and exploitation of current knowledge (Kaelbling et al., 1996).

The fast advances in machine learning and artificial intelligence, paired with readily available computing power to execute these algorithms on the smallest of devices, lead to new breakthroughs in scientific discovery and commercial applications across many field and sectors (Kersting and Meyer, 2018). At the same time, this technology is creating social and legal challenges that have to do with data accessibility and integrity, privacy, safety, algorithmic bias, the explainability of outcomes, and and transparency (Przegalinska, 2019).

## Social challenges of artificial intelligence

Preceding regulation and any legal action that may follow is the identification of situations where artificial intelligence is likely to be particularly challenged when it comes to making the right decision. Some situations are of course very clear cut. A movie recommendation system should obey parental restrictions and not serve up R rated or NC-17 rated content to a child. Likewise, an autonomous vehicle should not crash into a wall for no apparent reason. But oftentimes situations are far less clear cut, in particular when not only the user but also others are involved.

Social dilemmas are situations where what is best for an individual is not the same, or is even at odds, with what is best for others. Already in the early 80's Robert Axelrod (1981, p. 1390)

set out to determine when individuals opt for the selfish option, and when they choose to cooperate and thus take into account how their actions would affect others. Of course, cooperation is a difficult proposition because it entails personal sacrifice for the benefit of others. According to Darwin's fundamental "The Origin of Species" (1859), natural selection favours the fittest and the most successful individuals, and it is therefore not at all clear why any living organism should perform an altruistic act that is costly to perform but benefits another. In Axelrod's famous tournament, the so-called tit-for-tat strategy prove to be the most successful in the iterated prisoner's dilemma game. The strategy is very simple. Cooperate first, then do whatever the opponent is doing. If the opponent was cooperative in the previous round, the strategy of tit-for-tat is to cooperative. If the opponent defected in the previous round, the strategy of tit-for-tat is to defect. This is very similar to reciprocal altruism in biology. Recent research has also explored the impact of cognitive bias and punishment on cooperation in social dilemma experiments (Wang et al. 2018, Li et al., 2018), and indeed ample theoretical research has also been devoted to discovering what might promote cooperation in repeated social dilemmas in general (Wang et al., 2015, Perc and Szolnoki, 2010, Perc et al., 2017, Tanimoto 2018, Ito and Tanimoto, 2018).

But what about artificial intelligence, and especially one-off situations where the 'machine' has to determine whether to act in favour of the owner (or user), or in favour of others. This was brought to an excellent point by Bonnefon et al. (2016, p. 1573), who studied the social dilemma of autonomous vehicles. Inevitably, such vehicles will sometimes be forced to choose between two evils, such as running over pedestrians or sacrificing themselves and their passenger to save the pedestrians. The key question is how to code the algorithm to make the 'right' decision in such a situation? And does the 'right' decision even exist? Research found that participants in six Amazon Mechanical Turk studies approved of autonomous vehicles that sacrifice their passengers for the greater good and would like others to buy them, but they would themselves prefer to ride in autonomous vehicles that protect their passengers at all costs. Put differently, let others cooperate, i.e., sacrifice themselves for the benefit of others, but we would prefer to be spared.

This is nothing if not a brutally honest outcome of a social dilemma situation involving us, humans. We are social, and we are compassionate, and we care for one another, but in rather extreme situations Darwin still has the best of us. It is important to understand that cooperation is the result of our evolutionary struggles for survival. As a species, we would unlikely survive if our ancestors around million years ago had not started practicing alloparental care and the provisioning for the young of others. This was likely the impetus for the evolution of remarkable other-regarding abilities of the genus *Homo* that we witness today (Blaffer Hrdy, 2009). Today, we are still cooperating, and on ever larger scales, to the point that we may deserve being called "SuperCooperators" (Martin and Highfield, 2015). Nevertheless, our societies are also still home to millions that live on the edge of existence, without shelter, food, and without having met the most basic needs for a decent life (Arthus-Bertrand, 2015).

So what can we expect from artificial intelligence in terms of managing social challenges, and in particular social dilemmas? We certainly have the ability to write algorithms that would always choose the prosocial, cooperative action. But who wants to drive a car that may potentially kill you to save the lives of others. According to Bonnefon et al. (2016), indeed not many of us. Hence their conclusion, "regulating for utilitarian algorithms may paradoxically increase casualties by postponing the adoption of a safer technology". We thus have the knowledge and the ability to program supremely altruistic machines, but we are simply too

self-aware, too protective of ourselves, to then be willing to use such machines.

This in turn puts developers and engineers into a difficult position. Which is either to develop machines that are save but very few would want to buy, or to develop machines that may kill many to save one and will probably sell like honey. Nevertheless, the situation may not be as black and white, as artificial intelligence itself may learn how best to respond. Indeed, a recent review by Peysakhovich and Lerer (2018) points out that, because of their ubiquity in economic and social interactions, constructing agents that can solve social dilemmas is of the outmost importance. And deep reinforcement learning is put forward as a way to enable artificial intelligence to do well in both perfect and imperfect information bilateral social dilemmas.

Well over half a century ago Isaac Asimov, an American writer and professor of biochemistry at Boston University, put forward the Three Laws of Robotics. First, a robot may not injure a human being or, through inaction, allow a human being to come to harm. Second, a robot must obey the orders given it by human beings except where such orders would conflict with the first law. And third, a robot must protect its own existence as long as such protection does not conflict with the first or the second law. Later on, Asimov added the fourth law, which states that a robot may not harm humanity, or, by inaction, allow humanity to come to harm. But this does not cover social dilemmas, or situations, where the machine inevitably has to select between two evils. Recently, Nagler et al. (2019) proposed an extension of these laws, precisely for a world where artificial intelligence will decide about increasingly many issues, including life and death, thus inevitably facing ethical dilemmas. In a nutshell, since all humans are to be judged equally, when an ethical dilemma is met, let the chance decide. Put in an example, when an autonomous car has to decide whether to drive the passenger into a wall or overrun a pedestrian, a coin toss should be made and acted upon accordingly. Heads it's the wall, tails it's the pedestrian. No study has yet been made as to what would potential buyers of such a car make of knowing such an algorithm is embedded in the car, but it is safe to say that, fair as it may be, many would find it unacceptable.

Ultimately, the problems that arise when a machine's designer directs it toward a goal without thinking about whether its values are all the way aligned with humanity's, or when the machine is designed to "SuperCooperator" standards, rather harming the user than others around, we need good regulation and a prepared juristic system to tackle the challenges. This, however, leads us to a new set of challenges, namely those that are purely juristic.

## Juristic challenges of artificial intelligence

Considering its multifaceted character, artificial intelligence inherently touches upon a full spectrum of legal fields, from legal philosophy, human rights, contract law, tort law, labour law, criminal law, tax law, procedural law etc. In fact, there is hardly and field of law not affected by artificial intelligence. While in practice AI is just beginning to come into its own in terms of its use by lawyers and within the legal industry (Miller, 2017), legal scholars have been occupied with AI for a long time (journal "AI and Law", for example, dates back to 1991).

One of the most exposed legal issues related to law and AI concern patentability, joint infringement, and patent quality (Robinson, 2015a, p. 658). Internet of things (IoT) relies on communication between two or more smart objects and consumers and it is challenging whether inventors of certain types of IoT applications will be able to overcome the test for patent eligibility. Moreover, even if they obtain patents on new methods and protocols, the patents may still be very difficult to enforce against multiple infringers (Robinson, 2015b, p. 1961).

Furthermore, as collecting and analysing data is progressively spreading from software companies to manufacturing companies, which have started to exploit the possibilities arising from collection and exploitation of potential data, so that added value can be created (Bessis and Dobre 2014; Opresnik and Taisch 2015, p. 174; Opresnik et al., 2013), this information explosion (also called 'data deluge') unlocks various legal concerns that could stimulate a regulatory backlash. While it is claimed that data has become the raw material of production, and a new source of immense economic and social value (Polonetsky and Tene, 2012, p. 63), Big Data has been identified as the 'next big thing in innovation' (Gobble, 2013, p. 64), 'the fourth paradigm of science' (Strawn, 2012, p. 34) and as 'the next frontier for innovation, competition, and productivity' (Manyika and Bughin, 2011). On the other hand, however, open questions range from who is entitled to use this data, can data be traded and, if so, what rules apply to this. To prevent diminishing the data economy and innovation, 'smart' regulation is needed to establish a balance between beneficial uses of data and the protection of privacy, non-discrimination and other legally protected values. The harvesting of large data sets and the use of modern data analytics presents a clear threat for the protection of fundamental rights of European citizen, including the right to privacy (Brkan, 2015; Lynskey, 2014).

Thirdly, ICT is changing the role of the consumer 'from isolated to connected, from unaware to informed, from passive to active' (Prahalad and Ramaswamy, 2004). This process is sometimes also called 'digitalisation' of the consumer (Mäenpää and Korhonen, 2015), considering that people are increasingly able to use digital services. The younger generations are grown up with digitalisation and are eagerly in the forefront of adopting new technology. This could mean that the traditional presumption in consumer law that a consumer is uninformed and thus requires special legal protection no longer holds true. Nevertheless, the change is so rapid that the pre-Internet generations hardly follow the suit and new manufacturing methods bring new dangers for consumers and so consumer law need to adapt to the new challenges.

Finally, tax policy will play a very important role in the age of intelligent objects, particularly considering that human labour costs are increasing, so that it is broadly expected that automation will lead to significant job losses. As the vast majority of tax revenues are now derived from labour, firms avoid taxes by increasing automation. It is thus claimed that since robots are not good taxpayers, some forms of automation tax should be introduced to support preferences for human workers.

The focus of this review is on tort law aspects of intelligent objects. Tort law shifts the burden of loss from the injured party to the party who is at fault or better suited to bear the burden of the loss. Typically, a party seeking redress through tort law will ask for damages in the form of monetary compensation. Tort law aims to reduce accidents, promote fairness, provide peaceful means of dispute resolution etc. (Abbott, 2018, p. 3).

According to the level of fault, torts fall in three general categories:

a. intentional torts are wrongs that the defendant deliberately caused (e.g., intentionally hitting someone);
b. negligent torts occur when the defendant's actions were unreasonably unsafe, meaning that she has failed to do what every (average) reasonable person would have done (e.g., causing an accident by speeding);
c. strict (objective) liability torts do not depend on the degree of care that the defendant used, there is no review of fault on the side of the defendant; rather, courts focus on whether harm is manifested. This form of liability is usually prescribed for making and selling defective products (products' liability).

Multifaceted character of artificial intelligence brings challenges in the field of regulating liability for damage caused by intelligent objects.

**Tort law–adapting rules on product/services liability and safety**. In relation to automated systems, various safety issues may arise, despite the fact that manufacturers and designers of robots are focused on perfecting their systems for 100 percent reliability and thus making liability a non-issue (Kirkpatrick, 2013). It can happen that robotic technology fails, either unintentionally or by design, resulting in economic loss, property damage, injury, or loss of life (Hilgendorf, 2014, p. 27). For some robotic systems, traditional product liability law will apply, meaning that the manufacturer will bear responsibility for a malfunctioning part, however, more difficult cases will certainly come to the courts, such as a situation, where a self-driving car appears to be doing something unsafe and the driver overrides it—was it the manufacturer's fault, or is it the individual's fault for taking over (Schellekens, 2015).

Similar difficulties may arise in relation to remotely piloted aircrafts (so-called civil 'drones'). In the USA, a case concerning civil drones already appeared before the courts, when US Federal Aviation Administration issued an order of a civil penalty against Raphael Pirker, who in 2011, at the request of the University of Virginia, flew a drone over the campus to obtain video footage and was compensated for the flight. First instance court decided that a drone was not an aircraft, while the court of appeal ruled to the opposite. The cases ended in 2015 with a settlement for $1.100.

The starting point for examining "computer generated torts" (Abbott, 2018) is—or at least should be—that machines are, or at least have the potential to be, substantially safer than people. Although media broadly reported on the fatality involving Tesla's autonomous driving software, it is generally accepted that self-driving cars will cause fewer accidents than human drivers. It is stated that 94 percent of crashes involve human error (Singh, 2015). Moreover, medical error is one of the leading causes of death (Kohn et al., 2000). Consequently, artificial intelligence systems, like IBM's Watson, that analyse patient medical records and provide health treatment do not need to be perfect to improve safety, just better than people.

If accident reduction is in fact one of the central, if not the primary, aims of tort law, legislators should adapt standards for tort liability in case of harm caused by intelligent objects in such a way that law encourages investment in artificial intelligence and thus increases safety of humans. Most injuries people cause are evaluated under a negligence standard, where a tort feasor is liable in case of unreasonable conduct. If her act was not below the standard of a reasonable person, the harm is thought to be pure matter of chance for which no one can be held accountable. When computers cause the same injuries, however, a strict liability standard applies, meaning that it does not matter whether someone is at fault for the harm caused or not. This distinction has financial consequences and discourages automation, because computer controlled objects incur greater liability for the producer or owner than people. Moreover, if we want to improve safety through broader use of automation, current regulation has the opposite effect.

As currently product's liability is strict, that is independent of fault, while human activity is measured according to the standard of a reasonable person, legal scholars claim that in order to incentivize automation and further improve safety, it is necessary to treat a computer tort feasor as a person rather than a product. It is thus defended that where automation and digitalisation improve safety, intelligent objects should be evaluated under a negligence standard, rather than a strict liability standard and that

liability for damage would be compared to a reasonable person (Abbott, 2018, p. 4). Additionally, when it will be proven that computers are safer than people, they could set the basis for a new standard of care for humans, so that human acts would be assessed from the perspective what a computer would have done and how using the computer humans could avoid accidents and the consequent harm.

Nevertheless, jurists broadly defend strict liability for intelligent objects or in some respects even broader than currently foreseen, particularly in terms of the bodies involved that could be held liable—from the producer, distributer, seller, but also the telecommunication provider, when, for example, the accident was caused due to the lack of internet connection. At the European Union level, considering that the Product Liability Directive (85/374/EEC) does not apply to intangible goods, inadequate services, careless advice, erroneous diagnostics and flawed information are not in themselves included in this directive. It is nevertheless important that when damage is caused by a defective product, used in the provision of a service, it will be recoverable under the Product Liability Directive (Grubb and Howells, 2007), regulating strict liability test (see also EU Court's decisions on Cases C-203/99, Veedfald and C-495/10, Dutrueux). Many acts by robots will thus come within the ambit of this Directive, including software that is stored on a tangible medium. This means that in case the consumer, whose car causes an accident due to malfunctioning software, or a patient, who suffers the wrong dosage of radiation due to a glitch in the consumer software may bring a claim under the Product Liability Directive against the producer of software (Wuyts, 2014, p. 5). When software is supplied over the Internet (so-called non-embedded software), however, potential defects do not fall within the scope of this directive and a specific directive on the liability of suppliers of digital content is needed.

As far as product safety regulation is concerned, Article 2(1) of Directive 2001/95 on general product safety defines the reach of the product safety regime to include any product intended for consumer use or likely to be used by consumers 'including in the context of providing a service'. Nevertheless, this does not cover safety of services (Weatherill, 2013, p. 282). It is hence for the EU Member States to adopt legislation setting safety standards for services, which is not the preferred solution in times of extensive technological development. Analysis of the suitability of existing safety regulations is, for example, needed in relation to software-based product functions that can more and more be modified after delivery (WDMA, 2016, p. 12).

Moreover, in relation to drones, the EU Commission called already in 2014 for 'tough standards' to cover inter alia safety, insurance and liability (Press Release IP-14-384). Europe has about 2500 small civil drone operators, more than the rest of the world combined. Over the last few years, businesses have cropped up around the EU that manufacture and use drones in agriculture, energy, monitoring infrastructure, photography and other industries (Stupp, 2015; Michalopoulos, 2016). The regulatory work in this field is entrusted to the European Aviation Safety Agency (EASA) that is developing the necessary security requirements, as well as a clear framework for liability and insurance (North, 2014; Henshon, 2014; Mensinger, 2015). The Transport Committee of the European Parliament adopted a report (2014/2243) calling for Europe to 'do its utmost to boost its strong competitive position' in this field. Harmonised rules at the EU level would in this respect be welcome to safeguard a single market for the drones' industry.

It is also essential to understand, however, that the more autonomous the systems are, the less they can be considered simple tools in the hands of other actors (European Commission, Action Plan, 2014, p. 59) and that overly stringent regulation,

expecting perfection instead of acceptable robot behaviour, may discourage manufacturers from investing money in innovations, such as self-driving cars, drones and automated machines (Richards and Smart, 2013; Chopra and White, 2011). Smart regulation is thus again needed, taking into account all the involved stakes.

While intelligent objects are imitating the work of humans, as well as their legal liability, the question also arises, whether robots will be entitled to sue, be sued and also be engaged as witnesses for evidence purposes. Currently it is not possible to sue a robot as they are considered property, just like an umbrella. Intelligent objects do not have legal identity and are not amendable to sue or be sued. If a robot causes harm, the injured party have to sue its owner. However, comparing the robots to companies, for procedural purposes companies were also not treated as separate legal entities from the human owner for a long time in history (Abbott and Sarch, 2019). Nevertheless, over time legislators and courts abandoned the model of treating corporations solely as property and awarded them an independent artificial personality that allowed them to sue and be sued. In respect of the robots it will thus need to be established whether they are more like and employee, a child, an animal, a subcontractor or something else (Michalski, 2018, p. 1021).

## Conclusions and guidelines

Artificial intelligence certainly has the potential to make our lives better. It is in fact already happening, but as the adoption of any new technology, the welcoming of artificial intelligence into our lives is not without challenges and obstacles along the way. We have here reviewed some of the more obvious social and juristic challenges, for which we are nevertheless not well prepared. In particular, we have reviewed social dilemmas as traditionally demanding situations, in which we find ourselves torn between what is best for us and what is best for others around us and for the society as a whole. It is difficult enough for us to do the right thing in such situations, and now we have to essentially build machines that will, with more or less self-training, be able to do the right thing as well. The essential question is whether we expect artificial intelligence to be prosocial, or whether we expect it to be bent on satisfying an individual, the owner, or the company of which property it is. The meme "is my driverless car allowed to kill me to save others?" brings the dilemma to the point. It is relatively easy and noble to answer yes without much thought, but who would really want a car that could potentially decide to kill you to save other strangers. Research by Bonnefon et al. (2016) indicates that not many, depending of course on some details as to who might the passengers be and how many others would potentially be saved. But regardless of these considerations, one of such cars is an unlikely entry on the top of any wishing list. There are of course many similar situations that have the same hallmark properties of a social dilemma, and the answer to the question whether we want artificial intelligence to be prosocial or not certainly has no easy or universally valid answer. As is so often the case, it depends on the situation, and also on the juristic circumstances either decision would create.

Indeed, social and juristic challenges are often intertwined, and with this in mind, we have also reviewed the later in some detail. As industry and technology are changing hastily, all the involved stakeholders have to utterly consider, whether the society can adjust to this development equally fast and whether people develop the necessary working skills. While some commentators claim that EU may adopt the legislation concerning digitising industry too fast, since it is not yet known, how exactly smart industry will develop, others call for immediate response to avoid distinct legislative activities by individual states. Robotisation in

many aspects makes sense and it is thus reasonable that it gets regulatory support. However, this does not mean that it is always necessary to rush into new regulation, when amending existing legislation would suffice.

In reviewing the social and juristic challenges, we propose the following set of guidelines:

i. Improving the digital skills of the workforce for all professions and age groups requires public measures with pertinent financial support.

ii. Strict liability for the marketing of autonomous objects, claimed as necessary to protect the society from dangerous aspects of robotisation, in fact discourages investment in this field, thereby decreasing the potential of robotisation to make the society safer. This can be considered as the main regulatory paradox with respect to the introduction of AI into new applications.

iii. Before autonomous vehicles enter the roads, liability issues need to be clearly set by legislation, so that it is not left to the user to search and prosecute the liable entity in courts.

iv. Obligatory black box to record the functioning of the intelligent object and help ascertain liability in cases of potential faults.

v. No fine print. The user should be informed how the AI will react in critical situations.

vi. Last but not least, the off button should be readily accessible. Users should retain their right to decide for themselves.

We hope that the above review and guidelines will prove useful in successfully mitigating the social and juristic challenges of artificial intelligence.

## References

Araujo L, Spring M (2015) Product biographies in servitization and the circular economy. In: *Servitization: the theory and impact*. Aston Business School, Birmingham

Abbott R (2018) The reasonable computer: disrupting the paradigm of tort liability. Geo Wash L Rev 86:1

Abbott R, Sarch AF (2019) Punishing artificial intelligence: legal fiction or science fiction. SSRN 3327485

Arthus-Bertrand Y (2015) Human (movie). Bettencourt Schueller Foundation, Neuilly-sur-Seine, France

Axelrod R, Hamilton WD (1981) The evolution of cooperation. Science 211 (4489):1390–1396. https://doi.org/10.1126/science.7466396

Bessis N, Dobre C (2014) Big data and internet of things: a roadmap for smart environments. Springer International Publishing Switzerland, Basel

Blaffer Hrdy S (2009) Mothers and others: the evolutionary origins of mutual understanding. Belknap, Cambridge

Bonnefon J-F, Shariff A, Rahwan I (2016) The social dilemma of autonomous vehicles. Science 352(6293):1573–1576. https://doi.org/10.1126/science.aaf2654

Bräutigam P, Klindt T (2015) Digitalisierte Wirtschaft/Industrie 4.0, BDI. Noerr

Brkan M (2015) Data protection and european private international law. SSRN Scholarly Paper ID 2631116. Social Science Research Network, Rochester

Chabanne H, Urien P, Susini JF (2013) RFID and the Internet of Things. John Wiley & Sons, London

Chopra S, White LF (2011) A legal theory for autonomous artificial agents. University of Michigan Press, Michigan

Erkaymaz O, Ozer M (2016) Impact of small-world network topology on the conventional artificial neural network for the diagnosis of diabetes. Chaos Solitons Fractals 83:178–85. https://doi.org/10.1016/j.chaos.2015.11.029

European Commission (2014) European Commission calls for tough standards to regulate civil drones, Press Release IP-14-384, 8 April. http://europa.eu/rapid/press-release_IP-14-384_en.htm. Accessed 2 April 2019

Gobble MM (2013) Big data: the next big thing in innovation. Res-Technol Manag 56(1):64

Grubb A, Howells GG (2007) The law of product liability. Butterworths, Oxford

Gubbi J et al. (2013) Internet of things (IoT): a vision, architectural elements, and future directions. Future Gener Comput Syst 29(7):1645–1660. https://doi.org/10.1016/j.future.2013.01.010

Henshon M (2014) Case notes: FAA V. Pirker. Scitech Lawyer 11(1):26

Hilgendorf E (2014) Robotik im Kontext von Recht und Moral. Nomos Verlagsges

Hinton GE, Terrence JS (1999) Unsupervised Learning: Foundations of Neural Computation. MIT Press

Ito H, Tanimoto J (2018) Scaling the phase-planes of social dilemma strengths shows game-class changes in the five rules governing the evolution of cooperation. R Soc Open Sci 5(10):181085. https://doi.org/10.1098/rsos.181085

Kaelbling LP, Littman ML, Moore AW (1996) Reinforcement learning: a survey. J Artif Intell Res 4:237–85. https://doi.org/10.1613/jair.301. May

Kersting K, Meyer U (2018) From big data to big artificial intelligence? KI - Künstliche Intell 32(1):3–8. https://doi.org/10.1007/s13218-017-0523-7

Kirkpatrick K (2013) Legal issues with robots. Commun ACM 56(11):17–19. https://doi.org/10.1145/2524713.2524720

Kohn LT, Corrigan J, Donaldson MS (2000) To err is human: building a safer health system. National Academy Press, Washington, DC

Kopetz H (2011) Real-time systems. Springer US (Real-Time Systems Series), Boston

Kryvinska N et al. (2014) Servitization-its raise through information and communication technologies. In: Snene M, Leonard M (eds) Vol. 169, Exploring Services Science. Springer International Publishing (Lecture Notes in Business Information Processing), Geneva, pp. 72–81

Li X, Jusup M, Wang Z, Li H, Shi L, Podobnik B, Stanley HE, Havlin S, Boccaletti S (2018) Punishment diminishes the benefits of network reciprocity in social dilemma experiments. Proc Natl Acad Sci 115(1):30–35. https://doi.org/10.1073/pnas.1707505115

Lynskey O (2014) Deconstructing data protection: the 'added-value' of a right to data protection in the EU legal order. Int Comp Law Q 63(03):569–597

Mäenpää R, Korhonen SJJ (2015) Digitalization in retail: the impact on competition, In: Colin et al. (ed.) Vol. 7. Transition: IT Leadership in Transition. Aalto University. Helsinki, pp. 89–102

Manyika J, Bughin J (2011) Big data: the next frontier for innovation, competition, and productivity. The McKinsey Global Institute, New York

Martin N, Highfield R (2015) Super cooperators: altruism, evolution and why we need each other to succeed. Free Press, New York

Mensinger M (2015) "Remodeling model aircraft": why restrictive language that grounded the unmanned industry should cease to govern it. Minn L Rev 100:405–441

Michalopoulos S (2016) Commission promotes smart farming to mitigate climate change, EurActiv. http://www.euractiv.com/sections/agriculture-and-climate-change-role-new-cap/commission-promotes-smart-farming-mitigate. Accessed 2 April 2019

Michalski R (2018) How to sue a robot. Utah Law Rev 2018(5):3

Miller M (2015) The internet of things: how smart TVs, smart cars, smart homes, and smart cities are changing the world. Que, Indianapolis

Miller S (2017) Artificial intelligence and its impact on legal technology: to boldly go where no legal department has gone before Thompson Reuters. https://legal.thomsonreuters.com/en/insights/articles/ai-and-its-impact-on-legal-technology. Accessed 2 April 2019

Nagler J, van den Hoven J, Helbing D (2019) An extension of asimov's robotics laws. In: Helbing D (ed.) Towards digital enlightenment: essays on the dark and light sides of the digital revolution. Springer International Publishing, Cham, pp. 41–46

Ngai EWT et al. (2008) RFID research: an academic literature review (1995–2005) and future research directions. Int J Prod Econ 112(2):510–520. https://doi.org/10.1016/j.ijpe.2007.05.004

North D (2014) Private drones: regulations and insurance. Loyola Consum L Rev 27:334–357

Oettinger G (April 15, 2015) Europe's future is digital, Speech at Hannover Messe, *Speech 15-4772*. http://europa.eu/rapid/press-release_SPEECH-15-4772_en.htm. Accessed 2 April 2019

Opresnik D et al. (2013) Information–the hidden value of servitization. In: Prabhu V, Taisch M, Kiritsis D (eds) Advances in production management systems. sustainable production and service supply chains. Springer, Berlin Heidelberg, pp. 49–56

Opresnik D, Taisch M (2015) The value of Big Data in servitization. Int J Prod Econ 165:174–184. https://doi.org/10.1016/j.ijpe.2014.12.036

Polonetsky J, Tene O (2012) Privacy in the age of big data: a time for big decisions. Stanf Law Rev Online 64:63

Perc M, Jordan JJ, David, Rand G, Wang Z, Boccaletti S, Szolnoki A (2017) Statistical physics of human cooperation. Phys Rep 687:1–51. https://doi.org/10.1016/j.physrep.2017.05.004

Perc M, Szolnoki A (2010) Coevolutionary games—A mini review. Biosystems 99 (2):109–25. https://doi.org/10.1016/j.biosystems.2009.10.003

Peysakhovich A, Lerer A (2018) 'Towards AI that can solve social dilemmas', In: Christopher A, Thore G, Joel L, Frans O & Karl T (eds). 2018 AAAI Spring Symposium Series, Stanford University, Palo Alto, California, USA

Prahalad CK, Ramaswamy V (2004) 'Co-creating unique value with customers'. Strategy Leadersh 32(3):4–9. https://doi.org/10.1108/10878570410699249

Przegalinska A (2019) State of the art and future of artificial intelligence, Policy Briefing, Policy Department for Economic, Scientific and Quality of Life Policies. Directorate-General for Internal Policies. European Parliament

Richards NM, Smart W (2013) How should the law think about robots?, SSRN 2263363. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2263363. Accessed 2 April 2019

Robinson WK (2015a) Economic theory, divided infringement and enforcing interactive patents. Fla Law Rev 67:1961

Robinson WK (2015b) Patent law challenges for the internet of things. Wake For J Bus Intellect Prop Law 15(4):657–670

Russell SJ, Norvig P (2016) Artificial intelligence: a modern approach. Pearson Education Limited, Malaysia

Schellekens M (2015) Self-driving cars and the chilling effect of liability law. Comput Law Secur Rev: Int J Technol Law Pract 31:506–517. https://doi.org/10.1016/j.clsr.2015.05.012

Singh S (2015) Critical reasons for crashes investigated in the national motor vehicle crash causation survey US Department of Transportation. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812115. Accessed 2 April 2019

Stantchev V et al. (2015) Smart items, fog and cloud computing as enablers of servitization in healthcare. Sensors and Transducers (1726-5479), 185(2). http://www.opus4.kobv.de/opus4-srh/frontdoor/deliver/index/docId/718/file/P_2610.pdf. Accessed 2 April 2019

Strawn GO (2012) Scientific research: how many paradigms? EDUCAUSE Rev 47 (3):26

Stupp C (2015) Europe's first civil drone law gets a boost in Parliament. EurActiv

Tanimoto J (2018) Evolutionary Games with Sociophysics: Analysis of Traffic Flow and Epidemics Evolutionary Economics and Social Complexity Science. Springer, Singapore, https://www.springer.com/la/book/9789811327681

Wang Z, Jusup M, Shi L, Lee J-H, Iwasa Y, Boccaletti S (2018) 'Exploiting a cognitive bias promotes cooperation in social dilemma experiments'. Nat Commun 9(1):2954. https://doi.org/10.1038/s41467-018-05259-5

Wang Z, Kokubo S, Jusup M, Tanimoto J (2015) Universal scaling for the dilemma strength in evolutionary games. Phys Life Rev 14:1–30. https://doi.org/10.1016/j.plrev.2015.04.033. September

Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. Nature 393(6684):440. https://doi.org/10.1038/30918

WDMA European Office (2016) "Industrie 4.0: Mastering the Transition", 10 Key Recommendations for a European Framework for the Successful Digital Transition in Industry. Brussels

Weatherill S (2013) EU consumer law and policy. Edward Elgar Publishing, Cheltenham

Weber RH (2009) Internet of things–Need for a new legal environment? Comput Law Secur Rev 25(6):522–527. https://doi.org/10.1016/j.clsr.2009.09.002

Weber RH, Weber R (2010) Internet of things: legal perspectives. Springer Science and Business Media, Springer-Verlag, Berlin Heidelberg

Wilkinson A, Brax S, Jonsson K (2009) Developing integrated solution offerings for remote diagnostics. International Journal of Operations & Production Management 29(5):539–560

Wuyts D (2014) The product liability directive–more than two decades of defective products in Europe. J Eur Tort Law 5(1):1–34. https://doi.org/10.1515/jetl-2014-0001

Xia C, Xie F, Zhang Y, Caulfield C (2013) Artificial intelligence and data mining: algorithms and applications jianhong abstract and applied analysis. Hindawi Publishing Corporation, London

Yoo J (2017) Embracing the machines: rationalist war and new weapons technologies. Calif L Rev 105:443

## Additional information