

# CERIDAP

RIVISTA INTERDISCIPLINARE SUL  
DIRITTO DELLE  
AMMINISTRAZIONI PUBBLICHE

Estratto

FASCICOLO

4 / 2025

OTTOBRE - DICEMBRE

# Judicial Review of AI (assisted) Decision-making

*Guy Seidman*

DOI: 10.13130/2723-9195/2025-4-51

*L'Intelligenza Artificiale è nell'agenda di ricerca di tutti. Gli studiosi di diritto amministrativo, me compreso, non fanno eccezione. In questo contributo, illustro la mia prospettiva sulle sfide dell'Intelligenza Artificiale nel mio campo di ricerca. I governi sono generalmente assenti e lo sono doppiamente. In primo luogo, perché hanno abbandonato la loro responsabilità di regolatori finalizzata a garantire che tutto ciò che viene immesso sul mercato sia sicuro da utilizzare per il pubblico. In secondo luogo, perché i governi hanno rapidamente adottato l'Intelligenza Artificiale per le proprie attività, riconoscendone il potenziale in termini di maggiore efficienza. La sfida non sta nella politica, ma nei dettagli: le azioni del governo influenzano la vita, gli interessi e i diritti delle persone. È fondamentale che l'utilizzo dell'Intelligenza Artificiale non crei danni inutili. L'IA è uno strumento utile, anche nel governo. L'intelligenza artificiale è un valido strumento, anche nella pubblica amministrazione. Ma senza procedure chiare per gestirla, ci si può aspettare utilizzi dannosi che potrebbero suscitare l'opposizione dei cittadini rispetto ad un loro diffuso utilizzo.*

*Artificial Intelligence is on everyone's research agenda. Administrative law scholars, including myself, are no exception. In this paper, I present my perspective on the challenges of Artificial Intelligence in my field of research. Governments are generally absent and doubly so. Firstly, governments have abandoned their responsibility as regulators to ensure that all products placed on the market are safe for public use. Secondly, because governments have rapidly adopted Artificial Intelligence for their activities, recognising its potential for greater efficiency. The challenge lies not in politics, but in details: government actions influence people's lives, interests and rights. It is essential that the use of Artificial Intelligence does not cause unnecessary damage. AI is a useful tool, even in government. Artificial Intelligence is a valid tool, even in public administration. But without clear procedures to manage it, one can expect harmful uses that could arouse citizens' opposition to their*

*widespread use.*

*Summary: 1. Introduction.- 2. The Rise and Rise of AI.- 2.1. Along came Chat GPT, and the world loved it!.- 2.2. What AI IS, at present.- 3. How Dangerous is AI, today?- 4. Send in the Flaws.- 4.1. One: Algorithmic bias.- 4.2. Two: Hallucinations.- 4.3. Three: Explainability.- 5. How Can You Possibly Trust AI?- 5.1. One: The Trust and Reliance Gap.- 5.2. Two: Who Let the Dogs Out?- 5.3. Three: What Does It Mean “to Be Human”?.- 6. Enter the Government.- 6.1. Current AI Legal Regulation.- 6.2. AI’s Allure to Governments.- 6.3. Application: Reality Defies Imagination.- 7. What Can Be Done: Preliminary Thoughts.- 7.1. Is human discretion the problem or the solution?- 7.2. What does the public think about Government use of AI?- 7.3. Back to Explainability.*

## **1. Introduction**

*Artificial Intelligence (AI) means<sup>[1]</sup>, at its core, the ability of computational systems to perform tasks typically associated with human intelligence<sup>[2]</sup>. AI has been receiving serious academic attention for many years but has really boomed with the development of Generative AI (Gen. AI) in the current decade<sup>[3]</sup>. It seems fair to say that AI is now the “hottest” topic in all of academia, not least in law. Many scholars and researchers, from innovative graduate students to the “pillars of the discipline”, throw their hats into the ring, explaining what AI *is*, opining about its immediate impact, making predictions about its future use and proposing coping strategies<sup>[4]</sup>. By now, a vast amount of literature is already available<sup>[5]</sup> – probably more than any human can read<sup>[6]</sup> – and much more is being written on this topic at this very moment. Add to this that I am hardly a tech-expert, just an administrative lawyer, and you could well wonder how this piece came about.*

*My answer is this: I am troubled, like everyone else, about how “the machine”, AI, might impact humans, from jobs gone to human capacities lost.*

*In reality, at least for now, I am mostly worried about humans and how they use AI especially in the context of my field, administrative law. This topic concerns the legal means available to citizens to get remedy and review of alleged official*

wrongdoings and to ensure oversight of government operations. And the humans I worry about are the bureaucrats, and those who review their decisions. I understand officials “embrace of AI” capabilities to read and analyse the vast quantities of information present in the decision-making process. I am worried about their ability to retain human oversight in making decisions without deferring<sup>[7]</sup> to AI’s suggestions. The growing power and capabilities of AI make it not only increasingly enticing to use but also make it difficult to resist adopting its outcomes. I worry less that AI will make government officers’ jobs redundant than that administrators, presented with AI “suggestions”, might find it difficult to fully understand and overrule, and might find it easier, “more efficient” to rubber stamp AI’s conclusions; this, in turn, might make judicial (and other public) oversight difficult, if not impossible, if only because the adjudicator, a human for now, might not be presented with the full record of how the official decision was made.

This paper looks at the rapid rise of AI, the core challenges it poses to private, corporate and government use, and considers the ways to ensure that humans – officials, citizens, adjudicators – can maintain control, discretion, and oversight over AI suggestions. The paper is made up as follows: *Paras. 2 ff.* look at the dramatic, exponential recent rise of AI; *Paras. 3 ff.* ask how dangerous (to humans, to humanity) AI is at present; *Para. 4* looks at some of the acknowledged imperfections and well documented flaws of AI; *Paras. 5 ff.* look at why humans are willing to use AI, despite being aware of its flaws; *Paras. 6 ff.* look into official use of AI, and its particular dangers; the concluding *Paras. 7 ff.* make some initial suggestions as to how we – as humans, citizens, government agents – can ensure that AI is used to its greatest advantage while realizing its least potential harm.

## 2. The Rise and Rise of AI

### 2.1. Along came Chat GPT, and the world loved it!

AI is hardly a new issue, and the intellectual ideas and myths at its heart date back centuries. In the 1950s, AI moved from science fiction into reality, as serious scientific work started in the field<sup>[8]</sup>. Yet the field really took off only in recent

years. As computing power has increased in the past two decades, AI has greatly advanced technologically to the level that it can now increasingly be put into practice, and “do things” in the “real world”. Since the early 2000s, machine learning has come to be applied to a wide range of problems in academia and industry. In the last few years, AI advanced even further through deep learning and Gen. AI. This last term means, in brief and simple lay terms, a form of AI that uses vast databases to generate, produce, texts, images, videos and other forms of data. Furthermore, access to AI is no longer the purview of scientists and experts. In recent years AI has been made available, by private – mostly US – firms, to any person or entity with a computer or smartphone linking them online. It can now be used, at least in its basic form, virtually by anyone, anywhere, anytime, and for free. Not only are there no paywalls, but use of AI also now requires no real technological expertise just a *prompt* – a natural language text describing what task AI is asked to perform.

The world has embraced AI with open arms. Since 2022, individual and organizational use of AI has dramatically increased: OpenAI’s suite of Gen. AI tools reached one hundred million users in only two months. In 2025, an estimated 378 million people use AI, globally<sup>[9]</sup>.

It is fair to say that many people now use AI daily for work, study or play. For finding out facts AI did not so much replace such tools as Google – as it took them over<sup>[10]</sup>. A 2025 study across 47 countries shows that 66% of people report intentionally using AI on a regular basis for personal, work or study reasons, 38% of people report doing so on a daily or weekly basis. The rate is highest in India and Nigeria with 92% of people reporting regular use of AI systems, an “average” 68% in Israel, 60% in Italy, 53% in the USA, down to 43% in the Netherlands<sup>[11]</sup>. And it is not just the public that is harnessing the power of AI – corporations are increasingly using AI tools. A recent McKinsey global survey report found that 78% of firms reported using AI in at least one business function in 2024, up from 55% in 2023. Organizations reported using AI most frequently in IT, marketing and sales and service operations<sup>[12]</sup>.

## 2.2. What AI IS, at present

I do not want to raise expectations. Even the most laymen-friendly explanations I

have read do not shed much light on the inner workings of Gen. AI, which seem to baffle even the best of researchers. Which is why, to explain how the current generation of AI works and what its impact on humans can be, I have chosen two categorizations of AI.

*The first categorization* looks at the current techniques used by Gen. AI. In general, all types of Gen. AI rely on data to effectively “*learn*” how to generate content – but each is built around a different methodology, or technique of how they create human-like content – from words, to pictures, from video to code, and each can be used to create a somewhat different content:

- Large Language Models (LLM): this is the foundational technology behind such AI tools as ChatGPT. It consists of neural networks that are trained on huge amounts of text data, allowing them to learn the relationship between words and then predict the next word that should appear in any given sequence of words. LLMs have made it possible for computers to understand natural language inputs, enabling them to translate languages, analyse sentiment and generate images and voice from text. This is the form of AI that raises ethical concerns around bias, deepfake, hallucination, misinformation, and the use of intellectual property to train algorithms.
- Diffusion Models: these are used to generate image and video in such AI tools such as Dall-E. They work in a process known as “*iterative denoising*”. When they receive a text prompt is they generate a random “noise”, a random scribble if you will. This is then refined using its training data to understand what features should be included in the final image. At each step “noise” is removed until a new image is created, one that matches the text prompt and is not already in its training data.
- Generative Adversarial Networks: this is slightly older technology that emerged in 2014. It basically pits two different algorithms against each other: the *generator* attempts to create realistic content, while the *discriminator* attempts to determine whether it is real or not. Each algorithm learns from the other, until the generator creates synthetic content – text and images – that come as close as possible to being “real”. This technology is still used for computer vision and natural language

tasks.

- Neural Radiance Fields: this is the newest technology. It became available in 2020 and uses deep learning to create representations of 3D objects, *i.e.*, aspects of an image that cannot be captured by a real camera. «*This technique, pioneered by Nvidia, is being used to create 3D worlds that can be explored in simulations and video games, but also for visualizing robotics, architecture and urban planning*».

Hybrid Models of Gen. AI: these combine the various existing Gen. AI. methods. Drawing on the strengths of different approaches, the hybrid approach unlocks new possibilities for applications.

So much for the current state of Gen. AI. It should be noted that Gen. AI is constantly evolving. New methodologies and applications are emerging resulting in ever more advanced AI systems. «*The next decade is likely to bring groundbreaking applications that will transform industries and reshape how we interact with technology*»<sup>[13]</sup>.

*The second* categorization looks at a specific example of what AI can do, if allowed, to facilitate human activity. It concerns *autonomous vehicles* (AVs). I chose this example not just because AVs have such a clear potential to make everyday life easier but mostly because AVs have been thoroughly discussed and analysed before the public and their potential dangers exposed. The result is that unlike AI, the introduction of AVs to the market is delayed not (just) for technological reasons but owing to cautious regulators and a concerned public.

The potential here, the dream, is that AI will fully control the vehicle, read and understand road conditions, interact with any other cars on the road, and taxi the passengers to her destination. The potential for good here is enormous – making driving safer, faster, less stressful, more on this shortly. What I want to talk about is how the range between fully controlled driving – on a “dumb” car, and a fully automated drive has already been fully marked out, and is increasingly, at least technologically, feasible.

The Society of Automotive Engineers prescribes six levels of autonomous driving technology demonstrating the potential move from fully human controlled cars (*Level 0* – no driving automation), to fully AI controlled ones (*Level 5* – full driving automation, which might «*do away with steering wheels and pedals, as*

*there will be no requirement for them»).* In between lie the following levels: *Level 1*, where the car can go beyond passive assistance and offer at least one feature that provides elements of steering, braking, or acceleration support to the driver, in limited situations. Examples include automated lane keeping and part assistance. *Level 2*, which introduces advanced driver assistance systems. These combine various level 1 systems and can step in and support the driver in certain situations but only in approved areas of the roads. These features are considered “hands-off” rather than a full-self-drive system. *Level 3*, that introduces Conditional Driving Automations. When such systems are engaged, the car will steer, brake and accelerate entirely by itself, meaning the driver can take her hands off the wheel and her eyes off the road. The car is fully autonomous. *Level 4*, that allows for High Driving Automation – a fully self-driving car, where the driver, now effectively a passenger, still has an option to take over. This is the level at which the current experimentation with “robotaxis” takes place. It is allowed only in limited “geo fenced” areas – where highly-detailed HD maps are available – and when weather conditions permit and even then, speeds may also be limited<sup>[14]</sup>.

### **3. How Dangerous is AI, today?**

The question on many people’s minds is how much danger AI poses to us, both as individuals and as a society. It is very likely that you have read scary warnings about AI. Reports suggest that AI will eliminate a huge number of jobs – perhaps yours, perhaps mine<sup>[15]</sup>. Moreover, some great thinkers from Steven Hawking and Bill Gates to Yuval Noah Harari believe that AI will become too powerful for humanity to control<sup>[16]</sup>.

I dare not offer a definite answer, but I want to share a few hunches, which make me favor a less dramatic narrative. Let me explain why.

*First*, I would divide the assessment of the risks posed by ai – and hence the question of how much we should trust AI, be willing to use it, or call for its regulation – into two: the threat *currently* posed by AI and the dangers that lie *in the long run*.

On the latter – let me say that in principle I shy away from long term guesstimating and futurology, especially when it spells doom and gloom, even when it comes from some of our greatest minds. Humanity is resourceful and resilient, and it has proved time and again that it seeks life and can back off from the brink. Just recently, scientists moved the so called «*Doomsday Clock... closest ever to apocalypse – at 89 seconds to midnight*»<sup>[17]</sup>, but I believe humanity will never run out this clock. We have not seen wartime use of the atom in 80 years, threats notwithstanding. Similarly, humanity might find it difficult to take collective action to curb emissions causing global warming, but I am truly convinced that a solution – social, economic or technological will be found<sup>[18]</sup>.

In 1798 Thomas Malthus scientifically and convincingly demonstrated that the human population is limited by food growth. At much the same time along came the (first) industrial revolution and arguably sprang humanity out of the Malthusian trap<sup>[19]</sup>. Foretelling what lies behind the hill, in AI's future, is well beyond my pay grade. I want to hope that AI will not make humans obsolete<sup>[20]</sup>. What I firmly believe in is the human survival instinct, individual and collective. Thus, I expect humanity to unite, gracefully or not, around an acceptable outcome on AI, even if it would limit technological development<sup>[21]</sup>.

On the former – the threat posed by AI at its present state of development – I think we are in a better (if still imperfect) position to assess the potential dangers that lie ahead.

The starting point is that, to my best understanding, AI is still in its infancy. A work in progress. The AI tools currently available to the public are not a fully developed, finished product. They are beta versions<sup>[22]</sup>, under constant development and update<sup>[23]</sup>.

Perhaps AI tools should not have been released at this early stage, but they were. The fact is that AI tools, whether in free or for-pay versions, proved so much better than what was previously available to the public, that they were pounced on. In the short time that AI has been with us, many of its limitations have become very clear. Still, individuals, corporations and governments are all increasing their use of AI even though they admittedly neither fully understand how they work nor fully trust their outcomes.

*Second:* Being a lawyer, one of the oldest academic disciplines and one of the least changed, I tend to “explain away” new technologies rather than get blown away

by them. What I mean is that I believe that the traditional, often very ancient legal disciplines and categories can handle virtually all new technological developments and that to date, all the innovations I have seen only incrementally improve already available devices. I have studied autonomous operating weapons, dreading the idea that a missile will decide when and where to strike, until I realized this is essentially what the much reviled and much older weapon, the landmine, does. Smartphones have not really changed the technology at our disposal, they only made them more portable, just like the laptop did with desktops. In this sense, there is good reason why there is no “law of the internet” – like there was not “law of the horse”<sup>[24]</sup>. Yes, new technology will require many legal, social, economic, even physical, adaptations. Yet I think most of the public policy ramifications are highly foreseeable for anyone to see, a point I recently tried to make, with Aviv Gaon, about AVs<sup>[25]</sup>.

*Third:* With this in mind, I carefully estimate that there will be no independent new “AI Law” and that current structures for creating ethical, social and legal rules will be able to cope with AI. We humans can make ethics rules, decide on liability, require insurance, and demand good oversight over the activation of AI systems. This is, unless one of three things happens: *one*, AI reaches a level where it outsmarts humans; *two*, humans give up, opt out of AI use; *three*, humans make some really bad choices.

About the first two options, I have it on good authority that we should not be concerned. On the first concern, I take the word of eminent philosopher *Luciano Floridi*, who writes this about LLMs: «*They do not think, reason or understand; they are not a step towards any sci-fi AI; and they have nothing to do with the cognitive processes present in the animal world and, above all, in the human brain and mind, to manage semantic contents successfully.... However, with the staggering growth of available data, quantity and speed of calculation, and ever-better algorithms, they can do statistically – that is, working on the formal structure, and not on the meaning of the texts they process – what we do semantically, even if in ways (ours) that neuroscience has only begun to explore»*<sup>[26]</sup>.

This suggests that, at least for now, that science fiction fears that AI might become self-aware, sentient, a real form of intelligence and life, and come to threaten humanity, are premature<sup>[27]</sup>. It is to be hoped that AI will continue evolving, but not in a dangerous direction, and that humans will preempt,

prevent and preclude such abilities<sup>[28]</sup>.

On the second point, I would listen to philosopher *Andy Clark* who rebuts the «*techno-gloom*» and the fear that «*new technologies are making us stupid*». Clark advises us to bravely embrace AI, writing that: «*As human-AI collaborations become the norm, we should remind ourselves that it is our basic nature to build hybrid thinking systems – ones that fluidly incorporate non-biological resources. Recognizing this invites us to change the way we think about both the threats and promises of the coming age*»<sup>[29]</sup>.

I think that human experience – from the invention of the wheel and the discovery of fire to the industrial revolutions (of which we are said to be now at the fourth) – shows that most humans have found a way to make good use of new technology. Jobs were lost, others were created, and most people found a way to adapt. Even smart phones, a useful yet highly complex machine, have reached a penetration level of 90% in Italy<sup>[30]</sup>.

Which brings us to the third point. The choices that humans make. People, working in a private, corporate or official position, using AI might be tempted to lower their natural skepticism and critical thinking, their independent judgment, and choose to adopt the suggestions made by AI. It is now possible to use AI to create very realistic fake pictures and videos. To me, this is just one more page in the eternal battle between cheaters and forgers, on the one hand, and their intended marks and law enforcement, on the other. Anything of value was likely the subject of attempted forgery – coins and banknotes, antiques, passports – and now it is “deepfakes”. As ever, people must be vigilant in who and what they trust and not fall prey to potential fraudsters.

Personally, I would place deepfakes near the bottom of my list of concerns. Anyone posting fake data today is met with a very sceptical public, that is already self-selective in its choice of “reliable” news and is suspicious of any other information source. It seems like the public finds it impossible to agree on any core fact – be it the election results or that the earth is round. One more fake item? One more conspiracy theory? One more unsubstantiated argument? We can take it!<sup>[31]</sup>.

Let me qualify this optimism in two ways.

*First*, as the all-knowing narrator, you may ask yourself «*who are those people who take the results of an AI prompt as true and authoritative?*». I think the answer

would be that most of us, whether pressed for time, or impressed by the huge databases AI combs through, or the fluency of its answer, might come to rely on AI, with little or no fact checking. If this is not outright deference to the “all knowing” machine, it is at least reliance and trust of it. More on this shortly.

*Second*, who exactly are the “we”, who use AI? I think we must differentiate between private and public sector use of AI. As a private individual or corporation, I am free to make “bad choices” and lose my money, unless the law forbids me to do so. In some states gambling, drug use and similar habits are legal. In others – not. But government agents are different. Not only is it the role of government to make sure AI is safe enough to use – but it also must be more careful than private individuals, being the trusted agent of the people, in choosing to use AI. They must, then, be extra careful and vigilant, but it is my impression that this might not be the case, not just because of the “usual” agency problems – the difficulty to make government agents accountable for wrongdoings, let alone mistakes of judgment – but perhaps because AI really does look like mana sent from heaven to massively overworked government agents.

To me, human deference to the results of AI – this seemingly all-knowing technology – is what I would consider the greatest risk to humanity, especially for the millions of people working in all levels of government and exercising official judgment and the billions of citizens they work for. I would urge all people to use AI, but to retain control, oversight, judgment and caution, especially given the many flaws of AI that are already publicly known.

#### **4. Send in the Flaws**

The fact is, that anyone using Gen. AI is – or should be, by now – fully aware of its many flaws and shortcomings. The McKinsey report noted that many organizations are ramping up their efforts to mitigate Gen. AI related risks: firms declared that «*their organizations are actively managing*» Gen. AI risks related to inaccuracy, cybersecurity, intellectual property infringement, equity and fairness and explainability<sup>[32]</sup>. This is only a small part of a laundry list of known AI risks<sup>[33]</sup>, risks have the potential to infringe on established human right and interests<sup>[34]</sup>. In this paper I want to focus on just three of known problems with

AI.

#### 4.1. One: Algorithmic bias

The term refers to systematic and repeatable harmful tendencies of AI, which create socially unfair outcomes, privileging or discriminating against categories of people.

Algorithmic bias has been observed, *inter-alia*, in search engine results and social media platforms, with impacts ranging from inadvertent privacy violations to reinforcing social biases of race, gender, sexuality, and ethnicity<sup>[35]</sup>. It is truly troubling to find algorithms that reflect “systematic and unfair” discrimination, and the outcry against it very valid<sup>[36]</sup>.

But if AI had feelings, it might be anger rather than shame at this accusation, for two reasons. Firstly, who is it that taught AI? Who designed it, and then coded, collected and selected the data on which the algorithm is trained? Secondly, why is “the pot calling the kettle black?”<sup>[37]</sup>. An algorithm is not biased at its core, but humans are. Thus, we can strive for the evolution of “mindful AI” systems, built with awareness and consideration of their ethical, moral, and societal implications<sup>[38]</sup>. Humans, however, will retain their over 160 distinct cognitive biases<sup>[39]</sup>, of which they can, at best, be aware.

Otherwise said, we might be able to re-train and re-program AI facial recognition systems so that they would be “race blind”, but can we do the same for “real” people?<sup>[40]</sup>.

My take from this is that we seem to ask AI to operate at a higher standard of social fairness and ethics than that of the average human. It seems to me that we expect AI to meet standards of conduct, of care, that we set for the highest echelons of professionals such as judges or medical doctors. They too are likely biased in many ways, but are most likely, by training, by public expectations, to operate in the most equitable way towards all the people who require their professional judgment. Can we say the same about corporate or government officers?

## 4.2. Two: Hallucinations

This is where what AI generates in response to a prompt is not simply inaccurate,<sup>[41]</sup> but contains false or misleading information, presented as true facts.<sup>[42]</sup> What we hear about this issue the most in law is when students and lawyers submit work products based of made-up sources – results of AI legal search that refer to laws, cases, academic articles etc. that simply do not exist, that are completely fictional<sup>[43]</sup>. To me, this speaks less of AI's imperfect discretion than about humans' decision-making and discretion. Law firms are aware of the cost-saving potential of AI and are willing to embrace it, mostly without event<sup>[44]</sup>. When lawyers submit to court made-up sources – what they are is careless, negligent lawyers<sup>[45]</sup>.

How is it that legal professionals get caught time and again doing the same embarrassing mistake? My preliminary answer is this: AI seems to many people as a fast, economical, substitute for doing one's own legal research or for paying a human research assistant to do it. Even professionals can be cheap and lazy. Moreover, they might be time pressed and so fail to double check their submissions, especially when AI results are so nicely presented and have such an authoritative tone.

All of this is neither an excuse nor a novelty. Looking for shortcuts is hardly new. People have always plagiarized<sup>[46]</sup> and been sloppy in fully checking the data they rely on. Now, it seems that «*everyone is cheating their way through college*»<sup>[47]</sup>. This is unacceptable in college and surely unacceptable in professional lawyering. The standard for legal research outcome has not changed and there will always be negligent students and lawyers who fail to meet it. I do not see this as a major threat since humans (still, potentially) have full oversight over the outcomes of AI and a choice over how they use them – and humans can still double-check the facts presented to them. What this means is that people should realize that AI – should they choose to use it<sup>[48]</sup> – is an imperfect tool, with benefits and costs<sup>[49]</sup>, and that it does not absolve them from exercising full oversight, as they would a human underling carrying out legal research on their behalf. PS: a recent empirical study found that even professional, for pay, AI-powered legal research tools like Lexis+ AI, Westlaw AI, and Ask Practical Law AI still “hallucinate” 17% to 33% of the time<sup>[50]</sup>. In short: *trust – but verify*<sup>[51]</sup>.

### 4.3. Three: Explainability

This is likely the greatest challenge facing our use of – and reliance on – AI. What does it mean? Explainable AI, interpretable AI, or explainable machine learning refers to the efforts to allow humans to retain intellectual oversight over the reasoning behind the decision or predictions made by AI, making them more understandable and transparent.

The difficulty is the so called “black box” tendency of machine learning, and the fact that even AI’s designers cannot fully explain how it has reached a specific decision. Some machine learning models use a “white box” model, which provide results that are understandable to the experts in the domain.

*So why not make all AI models “white boxes”?* there are several reasons. *First*, because by making an AI system explainable, we reveal more of its inner workings – and allow competitors insight into a private, proprietary system. *Second*, because making AI explainable leaves it with less discretion, and allows “gaming” the results. *Third*, because the technical complexity of AI systems remains a barrier to explainability. *Fourth*, because even if we understand AI, we may not trust their decision.

I see explainability as a particularly great challenge in the context of administrative law.

If I, a private individual or firm, choose to use AI, I must be accountable for doing so. I should consider the fact that I do not understand how AI works or how it reaches its outcomes and that I cannot explain it to others with whom I may share AI results. But if I understand current social and legal norms, one is not expected to *understand* AI, nor to *explain* how it reached its outcomes. One is only required to be transparent about the use of AI and to make a fair effort to oversee, verify, check, its results, much as one would a human assistant’s output.

I believe the case should be different for government use of AI when acting in both its core capabilities – as regulator and actor.

I would expect government *regulators* to make sure use of AI is safe for the public to use. If it is not safe to use – I would expect government to prohibit its use, and if AI poses any risks, to issue the proper warnings. To do this and set citizens mind at peace government regulators need to fully understand AI, but do they?

In their capacity as *actors*, as direct providers of goods and services – can government ensure the public that the AI products that it develops or buys and then uses are safe? Can it ensure the public that government officials are applying a safe protocol for the use of AI and not overlying on its outcomes? Furthermore, should such questions come before an oversight body within government or the judiciary, can officials explain, in every case under review, how AI outcomes were produced, then how and why they were used by the human decision maker – *i.e.*, why the official chose to used AI products and to what extent did they affect the operative outcome in the case?

*I fear that is not how things now happen and a call for action is the focal point of this paper.* In this vein, let me thank Prof. Marc Rotenberg<sup>[52]</sup>, for sharing with me that he prefers to avoid the term “explainability” because it suggests, to lawyers, a rationale for predetermined outcomes. Rather, he would use the term “contestability”, by which he stresses the legal oversight and the requirement of fairness and due process.

## 5. How Can You Possibly Trust AI?

Given AI’s known shortcomings and established flaws, and that government regulators currently do not provide any assurances to the public as to the full extent of the harm that AI can potentially cause, the question is: how is it that individuals, firms and even governments, use AI on a vast scale and seem to trust it and endorse AI’s outcomes? <sup>[53]</sup>.

I think the answer is nuanced on many levels, beginning with the understanding that people neither fully understand nor fully trust AI. What they do is employ their best judgment in using AI outcome, even without government guidance. The result, as in any human activity, is that most people just have the intelligence, common sense and discretion to avoid serious mishaps. As long as people do not have to accept AI outputs and can double check their veracity and accuracy, the risk posed by AI remains acceptable. People take on crazy TikTok challenges, but luckily, only few end up killing themselves doing so<sup>[54]</sup>. Moreover, most societies (still) do not completely bar humans from engaging in risky behaviour – be it smoking, driving, or binging on sugar, fat and alcohol. In sum, it is likely that most legal professionals, even if they use AI, will neither rely on nor hand in fake

AI legal authorities. That is not what professionals do.

I think people use AI because it is the optimal choice now available to them, in terms of time and money. I do not want to minimize the impact of the huge economic interests that stand behind AI. Big technology companies (and governments) invest immense sums of money in AI and try to make it available for as many people as possible to use, even for free. But is this the most accurate explanation? Clearly a vast number of people think AI is a tool they want to explore and make use of, that AI shortens laborious processes and makes their work and learning missions easier, and many did so before big tech enticed them. Just to remind, innovator OpenAI that came up with Chat GPT, was not *big tech* at its inception, but rather – almost conversely – an organization aimed at developing “safe and beneficial” AI<sup>[55]</sup>. The masses immediately took to what was offered and embraced it. I think Big Tech followed popular demand in this case<sup>[56]</sup>. Again, for now, humans retain full control over *when* and *how* they use AI, for themselves or others. They may not fully understand it but can decide for themselves the risk level they are comfortable with. The greater the perceived risk, the less likely people are to trust and use AI. Let me elaborate.

## 5.1. One: The Trust and Reliance Gap

People do not fully trust AI and use it with caution. We know this from polls that show that the high rate of use does not fully translate into trust of AI. Even though the figures have been inching upwards, in advanced economies only half of people view AI systems as trustworthy and only 40% are willing to rely on AI system’s output and share information with such systems. In contrast, people in emerging economies have significantly more trust in AI systems. Moreover, the perceived trustworthiness of AI systems has *decreased* since their wide public introduction in 2022, declining from 63% of people viewing AI systems as trustworthy, to 54% in 2024<sup>[57]</sup>. People’s willingness to rely on AI systems has also declined from 52% to 43% and the share of people feeling worried about AI systems has risen from 49% to 62%.

This data, write the researchers «*likely reflects that with increased use and exposure to AI systems, people have become more aware of their capabilities and limitations, prompting a more considered reliance on these tools*»<sup>[58]</sup>. In fact, it can

also suggest that people are aware of the reports that AI is getting *less* rather than *more* accurate because of what is called “model collapse”<sup>[59]</sup> – or, in lay terms, inbreeding.

Nonetheless, people use AI, and what the KPMG report suggests to me is that while AI has cleared the core “willingness to use” threshold in most humans, it still triggers major concerns and ambivalence, especially in advanced economies.

The researchers linked this data to another clear finding – 70% of people across the globe believe that AI regulation is necessary. Only 43% believe that current laws are adequate while the remaining 57% disagree or are unsure that current regulation provides sufficient safeguards for AI use<sup>[60]</sup>. To me, this means, very clearly, that “the people” want to maintain human control over AI, and to the extent that they cannot ensure this themselves – by choosing whether and to what extent to use AI – they expect government to “do its job” and protect them. However, to put this in context and explain why this is interpreted as such a clear public sentiment – a staggering 83% of people reported not being aware of any laws, regulation or policy that apply to AI in their country.

The researchers read this data as meaning that «[t]here is a strong public mandate for international and national regulation of AI»<sup>[61]</sup>. I read it to mean that people feel unsafe, or uncomfortable in using AI. They are happy to be allowed to use AI, effectively without regulation, as part of their autonomy in a liberal democracy. Yet people seem to have a hunch that they are exposed to great risks, which they cannot face alone. Thus, they would like governments to consider regulations to ensure their safety even though it might result in limiting their currently free access to AI<sup>[62]</sup>.

The question now is what course should governments take regarding both citizen use of AI, and their own? How can they make sure AI is allowed to grow to its full potential while humans maintain sufficient control and oversight over it?

The main flaws of AI are now well documented and much discussed. This means that anyone using AI has had a fair warning and should be aware of its deficiencies. What seems to me as the correct protocol of action or standard of care for any human who chooses to make use of the AI outcomes, rely on them or submit them for others to use (and there are many millions of such persons) is this: *first*, exercise human oversight and double check the results; *second*, be transparent and let those to whom you submit something produced with the aid

of AI know that AI was used<sup>[63]</sup>.

## 5.2. Two: Who Let the Dogs Out?

With the understanding that AI is a powerful and highly imperfect tool I wondered how it was ever allowed to be publicly released before fully developed and vetted for safety. This is especially baffling in the developed world, where strong independent regulatory agencies exist for the specific purpose of protecting the public welfare, and where the use of even a food additive is strictly controlled<sup>[64]</sup>. How has AI “escaped the lab” and entered “real-life” so easily? Is there any serious chance, for all the talk, of any substantial AI regulation? Let me offer some intuitive answers.

*First*, I think most of the many AI tools available<sup>[65]</sup> are not considered either by the public or by regulators as inherently dangerous<sup>[66]</sup>. Rather, they are seen as benign, internet-related assisting tools – a new generation of search engines like Google or a toolkit like Microsoft’s “Office”, if on steroids. AI is viewed as an organizational and creative tool, allowing us to make short shrift of tasks such as creating documents and presentations or summarizing massive data sets, that previously required time, effort and likely the help of expensive professionals. In fact, we now see a trend towards greater use of AI for personal and professional support, rather than more technical uses, with the first spot taken by use of AI for therapy and companionship<sup>[67]</sup>.

That this poses danger to some workers whose job might become redundant is clear, but this comes part and parcel with every new, disruptive technology. Scribes gave in to typewriters; “office” software suite made secretarial jobs redundant. In short, AI offers great advantages along with what was thought – and is increasingly thought less so – as acceptable risk.

*Two*, it is possible that some regulation of AI will yet come into effect, but the fact is that AI burst into the scene and was widely adopted very quickly and with virtually no regulation. As this article will be published in Italy, let me remind of an event now in “ancient history”: in March 2023, Italy became the first Western country to ban ChatGPT. A month later, the AI chatbot was «*back online in Italy after installing new warnings for users and the option to opt-out of having chats be used to train ChatGPT’s algorithms*»<sup>[68]</sup>. It is very difficult to ban online

products that *the people really want*<sup>[69]</sup>.

*Three*, it seems to me that people are willing to use AI despite being increasingly aware of its downsides, as long as they feel they are in control, *i.e.*, that it is their choice to use AI, their discretion on how to use AI, and their oversight over AI's outcomes.

Where things are different, to my understanding, is when AI is used not in a personal capacity but by agents – by proxies carrying out activities we cannot (or choose not to) do ourselves, and especially where these functions are inherently dangerous. Here, crossing the threshold to using AI is a different matter. Let me give three examples: airplanes, autonomous vehicles, medical practitioners.

Aviation: airplanes are now mostly flown by computers. Human pilots control take-off and landing and oversee the rest of the flight. It is probably technologically feasible to fully automate civil (and military<sup>[70]</sup>) aviation, leading to significant savings. But, as the public reaction to a famous covid-era joke (the commercial airplane's captain allegedly announcing – «*this is the Captain speaking, I am working from home today...*») suggests – the public has not crossed the trust threshold and is unwilling to fly on an airplane that does not have a human at its helm<sup>[71]</sup>. This is despite statistics clearly showing that human error is the most likely cause for a crash<sup>[72]</sup>. Closer to home are AVs.

Vehicles: AI now enables fully autonomous cars, which would be more efficient and safer than the human-driven cars of today. However, the revolution is much delayed, not just for the sheer technological complexity of AVs but through human fear. We are aware of the shortcomings of human drivers yet remain deeply uncomfortable about giving full control over our safety, property, life and limbs, to a machine. It is not good enough for such a machine to be *better* at driving than any human has even been – we are already well beyond this threshold – the machine needs to be virtually foolproof, and this is a very tall order. Thus, the testing of robotaxis has proved that self-driving cars are very safe, dramatically outperforming human drivers in almost every crash scenario<sup>[73]</sup>. Yet many people remain anxious<sup>[74]</sup>, Automated Driving Systems remain under scrutiny<sup>[75]</sup> and news headlines keep playing to human concerns and fears<sup>[76]</sup>.

AI in Medicine: there is no doubt now that AI can significantly help in medicine<sup>[77]</sup>. One key area is diagnostics: AI can read and interpret scans better than any human medical practitioner, for much the same reason AI can identify

humans by picture better than any human can – it can better scan each pixel and make comparisons with a virtually limitless database of previous scans. The fact that AI outperforms humans<sup>[78]</sup>, however, does not mean that people would be willing to cut the human practitioner out of the loop, even if the AI is the better diagnostician. In fact, while the use of AI is very likely to become routine in the future and will significantly improve medical treatment<sup>[79]</sup>, at present, people are uncomfortable even with the idea that doctors use AI to help manage their care<sup>[80]</sup>, let alone that AI makes decisions about their medical treatment. Current standards require that medical professionals maintain oversight over AI suggestions<sup>[81]</sup>.

### 5.3. Three: What Does It Mean “to Be Human”?

There is a huge volume of literature on this question, especially in the age of intelligent machines<sup>[82]</sup>. To me, a key feature of humans is our strong bias in favor of our fellow humans, which takes precedence over any “non-human” be it animal or machine<sup>[83]</sup>, even if more capable at smelling or analysing data than humans.

In some legal cultures – Europe especially – people expect government to oversee all goods and services and regulate away danger to the extent possible. In other countries – the U.S. especially – people cherish their liberty to take risks and abhor regulations controlling guns or mandating vaccinations.

However, in all countries, all people must make many daily assessments as to the risks presented to them – we do not live in a sanitized world – and try to avoid them. Which is, I think, why people are willing to use AI for such functions as research, writing or creating pictures, which they consider “low risk”. This is much like people use telephones, email and the internet, despite the risk of scams or theft of intellectual property or even their identity. Having said this, people understand that they need both regulation of- and human oversight over- AI functions requiring greater expertise and potentially carrying higher risk to life and property like driving, flying or making medical decisions. Even though it is really a computer that flies the plane and AI that analyses medical data, when it comes to risking our life, we want a human, one with more knowledge than us, if likely less expertise than AI – to have the final say.

*To me a primary factor* is the human bias or even *chauvinism* in favor of humans – the latter term being defined as an «*unreasonable belief in the superiority or dominance of one's own group or people, who are seen as strong and virtuous, while others are considered weak, unworthy, or inferior*»<sup>[84]</sup>. Humans favor human judgment – theirs and that of their fellow human – despite of what they clearly know: the long list of psychological faults and biases that lurk behind the human decision-making process and the relatively frailty of the human senses. A clear example of this is the credence we give human testimony in court. We know that innocent people often plead guilty, but we still call confession the “queen of evidence”<sup>[85]</sup>. Similarly, we are aware of how the limitations of human senses and memory produce imperfect witnesses<sup>[86]</sup>, and yet, we are intuitively more comfortable with a human eyewitness presenting her account than with video footage of the event, much like we prefer human pilots and medical professionals over much better-informed and often truly superior machines.

*To me a second factor* is *agency*, which divides our functions between the things we can do by ourselves and the things that we need other people – of the private or public sector – to do for us. If we act on behalf of ourselves, we have autonomy, authority and control. Most people still fear AVs but some, the more audacious ones, are willing to participate in pilot programs and some even end up trusting robotaxis<sup>[87]</sup>. This is a matter of personal choice. Where we need to be piloted, invest our funds, get a haircut or medical advice, or be governed – we cede control to others. It seems to me that we have a strong preference, at least in high stakes decisions, to have a fellow human in the loop. Where AI assists in the process, we trust our fellow humans to be better able to judge the veracity and validity of AI’s recommendation and help us decide whether to act on it. And herein is the problem I see concerning government use of AI:

When we, as lawyers, use AI, we are the highly trained professionals, who must uphold a standard of conduct and thus, verify the results of AI – before we use them on behalf of clients. We expect medical professionals and airline pilots to maintain a high standard of professionalism. To check and double check AI input and never defer to AI.

*What about government agents?* Can we really trust them, at all levels of government to have the professionalism, and the time, the motivation, not to defer to AI – not to rubberstamp its recommendations? Furthermore, with the

difficulty explaining AI decision-making – will the supervising authorities, within the administration and in the judiciary, be able to observe the difference between human oversight and deference?

And there is another interesting wrinkle – AI is changing the standards of knowledge available to ordinary humans, and consequently, their behaviour. With the advent of the internet, then of powerful search engines like Google, and now, of AI, the extent of knowledge open and freely available to the public is the greatest in human history. Much of the data is presented or explained in clear, simple terms. This is why virtually any layperson can now, thanks to *Dr. Google*, find answers to medical and just about any other question she might have. Expertise is brought down to earth. The remaining advantage of experts is that they are better equipped to sift through data, place the information in context and, through experience and training, likely have better skills. The basic information – on how to cook like a chef, build a house – or a bomb – is now available to all. MDs, master builders and chefs are still likely to achieve better results.

With the Google power multiplier, and now with AI's enhanced powers, we can expect people's knowledge to change in various ways: first, *we* today know much more than previous generations, if only because so much data is freely available, well explained, to all; two, in considering the standard of care for non-negligent behaviour, we can – and perhaps must – now take the depth of freely available knowledge into account, when considering what people *should know* and *what enquiries they should make*; in the future doing one's own research might not be enough: for example, it is possible that *not using* AI in legal research and merely employing "human" skills, might be considered insufficient, under par<sup>[88]</sup>. To make this personal – I have not used AI in writing this paper. Soon, it might be substandard academic research *not to* use AI.

In many legal contexts we seek a good median of judgment and call it "the reasonable person". We also distinguish between the standard of care we ask of different people: the person acting in a private individual capacity is different from a commercial or government agent; we set different standards for lay people and for experts, and then again making different demands based on levels of professionalism – a barber, a chef, a banker, an MD, a pilot. All of this can be converted, using the ideas expressed in the previous paragraph to set the

standards according to the “reasonable” internet / Google / AI search, using it as the new “objective” a proxy to what we should expect each category of a “reasonable” person to know and how we expect them to act.

I would not be the first to suggest this. In a recent book *Valentin Jeutner* makes the wholly reasonable scholarly assessment that the fictional human standard of it “the reasonable person” would soon be replaced by the “reasonable algorithm”<sup>[89]</sup>, *i.e.*, the “reasonable AI”, and why not? The worst AI biases are no worse than human biases, and AI has access to more information, better processing ability and unlimited memory. With proper instructions, AI can do as good or better job than most human decisionmakers.

The saving grace is that, at least for now, we maintain human oversight and final decision-making power over AI, relegating AI mostly to an advisory role, but will this remain the case, as AI becomes ever more powerful, prevalent and persuasive?

## 6. Enter the Government

As noted earlier, governments look at AI from two perspectives: in their protective, regulatory capacity, seeking to ensure that AI is safe for the people to use and in their functional capacity, looking into using AI tools for their own activities, to be able to provide faster, cheaper and better services to the public. In both categories, governmental policy response has been underwhelming, even concerning.

### 6.1. Current AI Legal Regulation

AI has clearly jumped the gun on government regulation, but this is not to say that countries have not adopted laws and national policies dealing with AI. The European Union (EU) was the first national (or supra-national) entity to regulate AI in its AI Act, touted as «*the world’s first comprehensive AI law*»<sup>[90]</sup>. However, many countries are monitoring the development of AI and considering the need for regulation. Even the U.S., traditionally not a staunch regulator of technology, now considers the right course on AI<sup>[91]</sup>. The Organization for Economic Co-operation and Development (OECD) has created a database of national AI policies that provides a live repository of more than 1,000 AI policy initiatives

from 69 countries, territories, and the EU<sup>[92]</sup>. To these we should add previous national and supra national regulation such as the EU's General Data Protection Regulation (GDPR), of 2016 that went into effect in 2018<sup>[93]</sup>.

If you look at these regulations you will find that they do cover major AI concerns, including explainability: the AI Act speaks of it in the context of transparency<sup>[94]</sup>; notes the importance of explainability where law enforcement authorities use AI<sup>[95]</sup>; and notes system providers' duty to explain and document "the choices made", "when identifying the most appropriate risk-management measures" in high-risk AI systems<sup>[96]</sup>.

Most significantly, EU's AI law established a «*Right to explanation of individual decision-making*»<sup>[97]</sup>, which enhances a similar "right of explanation" contained in the GDPR<sup>[98]</sup>. Initial analysis suggests that while the AI act has introduced what seems like an important remedial tool, although its effectiveness remains to be seen, and «*will depend on complementary procedural safeguards and future legal developments – particularly regarding contestability, standard-setting, and participatory governance mechanisms*»<sup>[99]</sup>. Yet there are some encouraging signs. In a recent case, a telephone customer was denied service because of a credit decision by Dun & Bradstreet Austria GmbH (D&B). D&B did not provide the customer with meaningful information about the logic involved in the automated decision-making. The *Court of Justice of the European Union* ruled against D&B, stating that it must provide information so that the individual could understand which of her data was used in the decision and that trade secrecy concerns should not stand in the way of this right<sup>[100]</sup>.

These are, of course, welcome developments. However, my concern comes at a point further down the line: how exactly are people, especially government officials, making use of these AI systems, even if they formally declare and explain their use of AI? How carefully and critically are individual official users looking at AI output? Do they have the time, the skill, the knowledge to critically evaluate AI's suggestions, or will they mostly automatically endorse and rely on AI's recommendations. Here is an example: how often do you decide to ignore the driving suggestions of algorithms such as WAZE and how often do you simply follow the proposed route? Now, how likely is a busy customs or immigration official who gets a prompt from an algorithm that a person is "suspect" – or more dangerously, that a person is "non-suspect" – to overrule the

prompt based on his or her experience?

*The Economist* writes: «*Regulators are focusing on real AI risks over theoretical ones. Good*»<sup>[101]</sup>. True, but in my opinion the real risks require practical guideline, and now. What we need is a clear set of “marching orders”, especially in relation to government use of AI. This would explain, regarding state officials, who exactly have the legal power to authorize the use of each type of AI tool; which government officials, at what rank and with what power and discretion to their badge, are allowed to use AI systems? How extensive must their knowledge of how AI operates be? What oversight mechanisms will be placed over the use of AI products, and will such a review be conducted within the bureaucracy or outside of it, *i.e.*, by judicial, oversight, and will this oversight be done by humans or by AI systems, or both?

I think explainability is key. It would be very difficult, not legitimate, for a government official to say: «*I made this discretionary decision based on, in consultation with (or, in fact, in deference to) this technological wonder whose way of operation I do not understand*». That would be tantamount to admitting that she did not employ her own discretion making the decision entrusted to her. If she cannot explain AI to herself, how can she explain and justify relying on AI to anyone else? How do we expect the person on whose matter the decision was taken, the overseeing adjudicator, or the public to trust a government decision-making, entrusted to some mechanical *oracle*?

## 6.2. AI's Allure to Governments

When AI became globally freely available online to the public, I expected governments around the world to act first and foremost as regulators. I expected them to check out AI, its advantages and risks, and decide whether it is a product or service that can be safely made available to their citizens. Ensuring the safety and wellbeing of the people is, to my understanding, a government's prime directive. This, however, is not what happened as governments turned out to be unable (too slow?) or unwilling (for many possible policy reasons) to significantly limit or oversee AI's rollout. It is clear now that the AI Cat<sup>[102]</sup> is not about to return to the proverbial bag anytime soon.

What I did not foresee was governments' enthusiasm to embrace AI in carrying

out their own operations. Gradually it sunk in. The advantages of AI – its ability to process huge amounts of data, study and analyze it and then make deductions and recommendations, the very advantages that caught the attention of everyone from university students to corporate CEOs – work supremely well in government settings. Bear in mind that “The State” is the largest entity operating in any country. It is a behemoth. It has the most employees<sup>[103]</sup>, it produces the most data, it buys and sells the most goods and services.

AI is a dream tool for government in carrying out its functions in at least two respects: one, it gives each of the many layers of administrative decision-makers the power to use AI as to assist in decision-making. AI enables each official to take into account more data than any single human can possibly consider in a reasonable timeframe; two, AI has the potential to make redundant clerical positions such as those registering and archiving data, analyzing it etc. The dream of making government decision-making not just faster and cheaper – but also more fully informed – now seems achievable with AI.

I think Luciano Floridi hits the nail on its head in this respect when he suggests that we should not consider AI as a form of *intelligence* but «*as a new form of Agency without Intelligence, or Artificial Agency*»<sup>[104]</sup>. The most basic form of agency, he writes, *natural agency*, is defined by physical interaction with the environment. *Biological agency* introduces autonomy and adaptability, while *animal social agency* brings some collective behaviours. *Human agency* adds purposefulness, conscious deliberation and cultural meaning-making. Artificial agency «*represents a novel form of agency emerging from the interplay of programmed objectives and learned behaviours. At its core, AA is a computational, goal-driven form of agency defined by human purposes*»<sup>[105]</sup>. It is different from other, living, social, forms of intelligence. «*Artificial agents can maintain continuous operation without metabolic limitations*»<sup>[106]</sup>.

This reasoning we have seen before. This is the core of Kantorowicz’s argument in his seminal book “The King’s Two Bodies” which explains the medieval transition in governance theories from the monarch’s human, corporeal, living, “body natural” to the incorporate, undying, “body politic”<sup>[107]</sup>. The analogy is that we could now again be seeing a transition from government by *humans* (albeit democratically accountable officials rather than a sovereign king), to a decision that is attributed to- or at least is greatly informed by- a constructive, non-

human, intelligence.

Humans may be in control, for now. Yet we humans – private, corporate or government – might find ourselves soon unable to do without AI apps, as we did with cars and smartphones. What advantages human have – discretion, compassion, experience – are offset by the many limitations of the human mind and body and can be much better replicated by AI. How far will we allow this process to go, especially in government? How can we ensure continued human control of AI “advice”? How can we make sure that there are humans who understand the functioning of AI and can explain to others how AI reached its conclusion? How can we maintain public confidence that there are humans who retain discretion and oversight over AI and can overrule an ‘efficient’ solution in favor of a “human one”?

### **6.3. Application: Reality Defies Imagination**

In discussing government use of AI, I was going to use the following hypothetical.

A new school year is approaching. Your 6-year-old is about to start first grade. You want him to attend school A: it has a good reputation, is close to your house, and it is not only the school that you attended but also where your son’s best friend will be going. However, the municipality decided your son will attend school B. When you ask for an explanation, you are told that the decision was made, by a high-ranking school district official, at the recommendation of a new AI tool, which reviewed the details of all prospective first graders in the area over the past 20 years and considered *«a large number of other complex variables»*.

This scenario seemed to me realistic, if unpleasant. I hope school districts make serious enquiries and consultations, even with AI, before deciding where to send each school child. My concerns in this scenario would have been – who authorized the use of AI? who constructed the system? What exactly are these *«large number of other complex variables»*, and, most significantly, who *really* made the decision? If I, as a parent, wanted to challenge this decision, would there be a living human official who could explain how the decision was made and assure me that she reviewed and considered AI’s recommendation, rather than rubberstamped this decision and thousands more?

It turns out that in raising such concerns, we are chasing reality. Governments are aggressively looking for ways to harness AI's abilities to make faster, cheaper, less human-dependent administrative decisions, and the trend is hardly going away. In the US, President Trump was a big supporter of AI already in his first term, issuing an executive order on AI in 2019 (codified into law in 2020)<sup>[108]</sup>, and he is even more enthusiastic about AI in his second term<sup>[109]</sup>, raising concerns that his executive orders are «*dismantling of AI oversight, prioritizing commercial dominance*»<sup>[110]</sup>. But then, President Biden was heading in the same direction. The Washington Post reported in October 2024 that the White House was «directing the Pentagon and intelligence agencies to increase their adoption of artificial intelligence, expanding the Biden administration's efforts to curb technological competition from China and other adversaries»<sup>[111]</sup>. In the UK, one of the earliest countries to put AI into government service, both policy and problems became apparent by 2020, leading *The Economist* to warn that the British government's turn to algorithms to rule is «*weakening the public sector*»<sup>[112]</sup>. Showing, again, that there is no political divide on AI, Britain's recently elected Labour government issued a policy statement calling for AI to replace the work of government officials where it can be done to the same standard. The working mantra was that «[n]o person's substantive time should be spent on a task where digital or AI can do it better, quicker and to the same high quality and standard». The Labour government is willing to face stiff union opposition – unions being one of its core constituencies – for the prospect of major savings<sup>[113]</sup>.

When we look at what governments are *actually* using AI for, we find that they have gone far beyond my tame, even naive, hypothetical. With limited transparency and oversight, use of nascent AI systems in government has already had some ill effects, and started a lively public debate, about the possibility of ethically harnessing AI power. Here are a few examples, from the harshest, core-sovereign powers to the mundane workings of government<sup>[114]</sup>, plus a few of the greatest scandals, to date:

*One: Military uses of AI:* countries obviously want to use their AI capabilities to make their military better, more efficient – cheaper to operate, safer to their personnel and more effective against their foes. Israel has been uniquely open about its practices. It appears that private, commercial, «*U.S. tech giants have quietly empowered*» Israel allowing the Israeli military to use AI «*to sift through*

*vast troves of intelligence, intercepted communications and surveillance to find suspicious speech or behaviour and learn the movements of its enemies*». The Israeli military «*has called AI a ‘game changer’ in yielding targets more swiftly*»<sup>[115]</sup>. Military use of AI raises many concerns, but also the hope that its ethical use might lead to narrower, more accurate targeting, thus limiting so called “collateral damages”<sup>[116]</sup>.

*Two: Police use of AI: «AI is transforming policing, sometimes in dramatic ways. Face recognition, predictive policing, and location-tracking technologies – once the stuff of science fiction – now are being adopted by law enforcement agencies large and small»*<sup>[117]</sup>. Many of these tools are controversial and the effectiveness still unclear, but police is already using them to investigate and deter crime, notably to identify individuals, track people’s location and movements, manage and analyse crime, but also analyse emotions and predict crime<sup>[118]</sup>. This is not pure theory, and worrying reports abound: such as that British police has contracted a controversial US tech giant to buy «*AI tool that looks at people’s sex lives and beliefs*»<sup>[119]</sup>, or that the city of New Orleans is now considering legalizing police use of “facial surveillance”, a step that would make it the first American city to formally allow facial recognition – a very controversial tool – for surveilling residents in real time<sup>[120]</sup>. You don not really need to ask “what could possibly go wrong?” because we already know. The US has already witnessed wrongful arrests based on mistaken AI facial recognition<sup>[121]</sup>. It turns out that since African Americans are subject to arrest at a disproportionate rate, facial recognition systems that rely on mug shot databases are likely to include an equally disproportionate number of African Americans<sup>[122]</sup>. Additionally, facial recognition software is significantly less reliable for Black and Asian people, who, according to a study by the National Institute of Standards and Technology, were 10 to 100 times more likely to be misidentified than white people<sup>[123]</sup>. These and various other AI tools are now being tested and used by police forces in many nations around the world, including India and Thailand<sup>[124]</sup>. Even in liberal democracies, police and military use of AI raises concerns and calls for clear rules and oversight. Imagine what they can do in the service of illiberal autocratic regimes<sup>[125]</sup>.

*Three: Civilian use of AI: there is also extensive use of AI for the more mundane, less dramatic everyday work of government*<sup>[126]</sup>: one typical use of AI is to improve

documentation: to reduce paperwork, digitize paper documents, improve search and extraction of data and accelerate processing. Where this becomes more concerning, meaning human oversight is clearly still needed, are such tasks as document translation and document drafting. A *second* typical use of AI is to answer the public. Chatbots are often used for “first contact” with the public, reducing the real-life interaction with a human administrator<sup>[127]</sup>. This is an efficient solution to screen frequently asked questions, fill out forms or schedule meetings. However, it surely cannot replace all government “customer service” and must be supplemented with the option of contacting a human official. A *third* typical use of AI is to allow decision makers to better analyse data, identify patterns and predict outcomes. This is especially important when authorities need to prioritize resources: AI tools analyse patient data and make assessments and predictions that help hospitals rank patient care; AI can analyse data to help predict fire or crime and thus optimize firefighter and police preparation. This use is more controversial for many reasons including the concern of AI bias and the sheer risk to life involved and so clearly demands robust human oversight.

And there have been many a fiasco. Consider these examples:

- In the UK More than 900 sub-postmasters were wrongly prosecuted between 1999 and 2015 after faulty software said money was missing from their Post Office branch accounts. «*It has been called the UK’s most widespread miscarriage of justice*»<sup>[128]</sup>.
- Still in the UK, in 2020, the government replaced human teachers with an algorithm in grading A level exams. Almost 40% of students received lower grades than they expected. The outcry forced scrapping of this plan – the argument was bias, but it is more likely that this was an intentional programming of the algorithm, meant to correct for observed (human) grade inflation.<sup>[129]</sup>
- In 2021, the entire Dutch cabinet resigned in 2021 «*after a year and a half of investigations revealed that since 2013, 26,000 innocent families were wrongly accused of social benefits fraud partially due to a discriminatory algorithm*»<sup>[130]</sup>.
- In 2023, the U.S. State of Nevada used AI to estimate and predict the number of children who would struggle in school. The complex

algorithm «*cut the number of students deemed ‘at risk’ in the state by 200,000, leading to tough moral and ethical questions over which children deserve extra assistance*»<sup>[131]</sup>.

- In 2024 France it has been reported that «*Algorithms policed welfare systems for years. Now they’re under fire for bias: Human rights groups have launched a new legal challenge against the use of algorithms to detect error and fraud in France’s welfare system, amid claims that single mothers are disproportionately affected*»<sup>[132]</sup>.
- Finally, a report on “Authoritarian Technology” told the public this story in 2021: «*Who’s homeless enough for housing? In San Francisco, an algorithm decides. Replacing human decision-making with a computerized scoring system is hurting California’s most vulnerable residents*»<sup>[133]</sup>.

In similar tenor, legal scholar and philosopher *Nathalie A. Smuha* has raised the alarm about public authorities’ increasing use of algorithmic systems to apply and enforce the law. She calls this «*Algorithmic Regulation*». Reliance on such algorithmic systems, she warns, «*can undermine the law’s protective power and instead lead to rule by law*»<sup>[134]</sup>.

## 7. What Can Be Done: Preliminary Thoughts

We need to talk more about how we use AI, for private, commercial and government activities, and we must do so publicly and honestly.

AI offers huge benefits for humanity, and not just in potential but already in practice.

At the same time, like any powerful, useful, tool – from a utensil to a car – it can also present dangers, if used improperly or if it falls into the wrong hands. I think that the balance is clearly in AI’s favor. AI is already a very widely used tool, and I think it is here to stay<sup>[135]</sup>. I see no realistic option of shutting it off or even significantly regulating access to it be it for private, corporate or government use, at least not in liberal- democratic societies. I think what we can and should do at this stage is to develop best practices, either as a social convention or a legal norm, i.e., voluntarily or by regulation, so that we can make the best and safest long-term use of AI<sup>[136]</sup>.

## 7.1. Is human discretion the problem or the solution?

It is easy to argue that the best way to prevent any unfortunate consequences from the employment of AI is to mandate that some form of independent, human, oversight be maintained over the operation of AI, especially when carried out by agents, private or public. It is easy to show cases where human judgment can present a simple, human, solution, to complex problems, cutting through red tape and formalities. But that would be too easy. The problem I see is that humans are part of both the solutions and the problems (albeit not the same individual humans). On this let me say:

*One:* I think it would be difficult to carry out human-control oversight over AI in practice. AI is based on very sophisticated and complex algorithms, and it is very unlikely that most people using AI either fully understand how AI operates, or can explain this to others, should they be asked to. In a way, the reason we use AI is precisely *because* it outperforms humans. None of us runs faster than a car and very few of us can fully explain how a car operates. Now try to explain why your AI run AV was involved in a traffic accident, while you, the owner, were fast asleep in the back.

It seems to me that governments get into trouble, in their use of AI systems, when they advance too fast, beyond the ability of public officers to understand and control AI, and beyond what is acceptable use of AI by the general public. One area I would flag is where governments ask AI to make highly discretionary decisions, often in fields that greatly impact civil and human rights, and do not provide assuring mechanisms for human oversight and review. One such field is welfare policies. Another is immigration. The UK uses an algorithm called “IPIC” (“*Identify and Prioritise Immigration Cases*”) to automatically identify and recommend migrants for immigration decisions or enforcement action. Rights groups already warn that IPIC, which uses biometric data, ethnicity, health, and criminal history to decide on enforcement actions «*risks ‘rubberstamped’ migrant deportations*»<sup>[137]</sup>.

Use of such systems is concerning to many, even outside the human rights community. I think the public is troubled by the many scandals concerning government use of AI, especially when the cases involve apparent AI bias and resulting discrimination. This is not to say that the public is *against* government

use of AI but that people are troubled over how it is currently happening: without clear principles for AI use, bars on over-reach, caution when employing new and untested tools, and without assurances of sufficient oversight by more senior government officials and bodies of review.

*Two:* the fact is that most malfunctions in the operation of AI, at least for now, are of human origin. One example is human “error” in teaching the algorithm, *i.e.*, in the information the AI is “fed”, which might lead to (human) bias. A second example concerns the mistakes humans make in operating AI tools. In the famous English exam grading fiasco critics pointed out that humans were to be blamed for the algorithm’s failings. It turned out that the UK government «*ignored its own guidance and refused external help, resulting in an A level algorithm disaster that disproportionately impacted students from disadvantaged backgrounds*», leading one commentator to say this: «*Many of the cases of algorithms gone rogue that we know about could have been stopped by critical reflection earlier in the process. Such reflection, however, is unlikely to come introspectively. Despite their best efforts, those developing algorithms will be prone to bias and intellectual lock-in*»<sup>[138]</sup>.

## **7.2. What does the public think about Government use of AI?**

The public clearly wants to use AI itself, if cautiously, with limited say by government. Now reverse the question to ask what the public thinks of AI use in the public sector?

A major survey conducted in the UK paints a complex picture. It suggests that the public’s view on the matter is not singular but nuanced. For each case of government use of AI people weigh the potential opportunities and benefits against the perceived risks and harms. What complicates the calculus are people’s specific understanding, trust and comfort levels with AI use. These vary depending on personal characteristics and demographics and personal experiences with AI. People are concerned that private tech companies are pressing for more use of their technology in the public sector, for pay, a concern that could be alleviated with transparency, regulation and clarity about the use of personal data. People are clear that they want explainability of AI tools, clear

evidence that they are effective and full data about their expected impact. People express the desire that those affected by AI tools have a meaningful say in shaping the decision to use them. «*The public increasingly ask for stronger governance of AI, along with clear appeals and redress processes is something goes wrong*»<sup>[139]</sup>. Is the public wrong on any of these points?

To me, public concerns can be divided into four legitimate issues. The first questions governments' motivations to use AI (*i.e.*, is AI really chosen for the better good of the public or to allow the government to use “new toys” and private firms to make money?). The second asks whether AI's overall effects are positive (*i.e.*, will AI promote public welfare more than it detracts from it?). The third investigates the extent the public is involved in the decision to employ AI in the public sector (*i.e.*, is the public told, in advance, of the plans to use AI, of what it is intended to achieve, of how it operates, of its expected costs and downsides?). The fourth is whether the public is convinced that government agents can competently operate AI (*i.e.*, that the AI system is not too complex to handle, that officials will not just be rubber-stamping AI “recommendations”, and that there would be both direct human control over AI, and human oversight procedure to review the decisions taken with AI advice).

Since the grading scandal happened, the UK has developed the *Algorithmic Transparency Recording Standard* (ATRS), which requires public sector organizations to publish information about how and why they are using algorithmic tools to support decisions<sup>[140]</sup>. This is clearly a step in the right direction, a move towards “algorithmic accountability”, an issue of huge importance and variance<sup>[141]</sup>, but more needs to be done. A recent article highlights «*an underlying issue concerning government automated decision-making systems: the lack of public scrutiny they receive across pre- to post-deployment*»<sup>[142]</sup>. The authors suggest that the British government uses at least fifty-five automated systems, many of them are «*not well understood*»<sup>[143]</sup>. Their research prompts concern over insufficient safeguards protecting the public against «*illegality manifested through government [AI] systems*»<sup>[144]</sup>. Their recommendation is to create «*pre-deployment impact assessments*» of AI intended for government use, and instituting «*algorithmic auditing as part of reinforcing the duty of candour in judicial review, so as to inform courts about specific systems and the data underpinning them, which can assist judges when ruling on matters involving*

*government [AI] systems»*<sup>[145]</sup> .

### 7.3. Back to Explainability

With this we come full circle to explainability, which I think is key to all meaningful interaction with AI – use, control and oversight. If you do not understand what AI is doing, how and why it is reaching its results and makes its suggestions and recommendations, I think you should not be dealing with AI. The level of understanding required can be debated – not everyone is a mathematical genius with technological acumen, but some form of standard, a baseline, must be set. We have succeeded in doing so with the driving license, reaching an almost universal standard for the basic qualification needed to handle a car: an understanding of how the car works and how to physically operate it, studying traffic law, driving under supervision until we learn what to expect of other human drivers and how to interact with them and how to adapt our driving to various weather conditions. Drivers of larger vehicles such as trucks and buses likely needed higher qualifications. We should probably strive to ensure that AI literacy exists within the entire population and surely set this as a prerequisite for using AI as agents, not least government agents of the people.

Explainability, meaning the ability of government agents to explain and justify their act or omissions, is vital for their own function as well as for those who review their actions, and surely, for public trust. These are not really new ideas, brought about by AI.

Furthermore, I believe, almost as an article of faith, that most people are willing to accept almost any outcome, good or bad, even mistakes and negligent acts, when these outcomes are fully and honestly explained to them, and where they can ask questions, present their viewpoint, be seen, be heard. This is why the “right to be heard” is such a fundamental rule of natural justice<sup>[146]</sup>, and a cornerstone of administrative justice. It is possible to make the same point on many basic principles of public law. Thus, the U.S. Constitution’s Bill of Rights provides, in the 1<sup>st</sup> Amendment that «*Congress shall make no law... abridging... the right of the people peaceably to assemble, and to petition the Government for a redress of grievances»*<sup>[147]</sup>, and in the 5<sup>th</sup> Amendment, that «*No person shall be deprived of life, liberty, or property, without due process of law»*. Can any of these

rights be kept if the people are not fully informed about decisions made by their representatives, especially those impacting on their vital interests? What would be the point to assemble to complain – if you do not know and understand the basic facts? How would a human adjudicator be able to provide due process of law, legal oversight over government action, without knowledge and understanding of the facts?

I think that we will see an adaptation of current legal standards for human discretionary decisions to activities involving AI. The devil, as always, is in the details. For example, we might choose to keep current administrative standards and apply them to AI. This would mean that any official action involving AI would have to stand up to the full standard of explainability required of a human decision maker – whether in explaining the action to the public, specifically members of the public whose interests were harmed, to her superiors within government or to reviewing tribunals and courts.

The follow up question is whether we will maintain the standards of review that human decisionmakers must currently meet at the review instance: whether in terms of the extent of review (from de-novo to deferential, in US administrative law) or in terms of the substantive demands that official decision must need (be non-arbitrary, reasonable, proportionate etc., as required by most other administrative law regimes).

The next question is one of proof: how would an explanation for an AI assisted-decision look like? How much detail should be given to the public? How deep the transparency? And in explaining the decision to legal instances – would we keep the current rules of evidence – will humans come and translate, explain, to the adjudicator why the AI has made a recommendation, what instructions the AI tool was given, what information it was fed, how it operated technically, and why the outcome should stand scrutiny just as it would if it were made by a human decision maker?

Perhaps new models for review and oversight would be developed. It must be said that there are great developments made these days in the field of *Explainable Artificial Intelligence* (XAI), with the aim «to bridge the gap between the complexity of advanced AI models and the need for human understanding and trust»<sup>[148]</sup>.

But, again, explaining AI is not enough, at least not in law. Courts are already

grappling with AI's impact<sup>[149]</sup>. Here is my take: I think a full understanding of AI must be part and parcel of the entire administrative process, with all sides – citizens, officials, review bodies, being able to understand, on some core level, what AI is doing and why. Explainability – or, as Prof. Marc Rotenberg terms it, “contestability”, should be baked into the process.

For this to occur, I can envision at least three different models in action. One would require AI assisted discretion in administrative law areas where decisions can potentially make significant impact on human interests to follow a co-decision model – one where a human must approve AI's decisions. A second, optimized for mass- or rapid- decision-making, would use one or more extraneous AI systems to vet the decision taken by the primary AI system. There is already a burgeoning industry of AI software's that look for AI falsifications, AI double checking AI<sup>[150]</sup>. A third and most intriguing option would be to allow the AI system to explain itself, like human decision makers would, and defend its' decision-making process. For this, we need to be sure that AI will tell truth to power, but if you can ask ChatGPT to reveal its' sources and explain itself – why not government AI tools?

1. An earlier version of this paper was presented at the ReNEUAL– New Frontiers meeting in Rome, November 2024. I have benefitted from comment made at that meeting and the insightful advice of many others, notably Ben Haklai, PHD Candidate and National Technology Officer, Microsoft Israel, Marc Rotenberg, of Georgetown law school and President of the Center for AI and Digital Policy Center for AI and Digital Policy (<https://www.caidp.org/>), Dan Hunter, the Executive Dean of the Dickson Poon School of Law at King's College London & Jonathan Boymal. My sincere thanks to Diana-Urania Galetta for shepherding this project to fruition and to Aviv Gaon and Yoram Avinor for sound advice. All views expressed are mine alone, and all comments are very welcome – gseidman@runi.ac.il. No AI was harmed in the writing of this paper.
2. For a deeply thoughtful discussion see, Catholic Church's, Dicastero per la Dottrina della Fede e del Dicastero per la Cultura e l'Educazione, *Antiqua et Nova: Note on the Relationship Between Artificial Intelligence and Human Intelligence*, 28.01.2025, available at: <https://press.vatican.va/>.
3. One impressive measure is Google Scholar's Ngram Viewer of books mentioning AI. See:  
<https://books.google.com/ngrams/graph>;  
[https://en.wikipedia.org/wiki/Artificial\\_intelligence](https://en.wikipedia.org/wiki/Artificial_intelligence).
4. See, e.g., R. Susskind, *How to think about ai – a guide to the perplexed*, 2025; M. Coeckelberg, D. J. Gunkel *Communicative AI - A Critical Introduction to Large Language*

*Models*, 2025; an example for a thesis looking at the likely impact of AI on the operation of the judicial system see, C. Bordere *La Justice Algorithmique – Analyse Comparee*.

5. Here is one measure: Lexis now offers 11.402 US law reviews and journal articles containing the words “Artificial Intelligence”; in Westlaw the figure is 9.840 (of them 875 where AI is in the title); on SSRN the figure for papers containing AI in their title is a staggering 4.757 papers (last search – 24.05.2025).
6. I did my best human effort researching this topic sans AI. Clearly, however, Chat GPT knows more than I even would on this and any other topic; but can today’s AI offer a better analysis of the current condition of AI that I? Should we Ask it to analyze itself? Is it acceptable academic practice do so do? Would it be, soon, acceptable academic practice *not* to do so?
7. On the meaning of deference, see: G. Lawson, G.I. Seidman, *Deference*, OUP, New York (USA), 2019.
8. The modern field of AI began in 1950, when Alan Turing published “*Computer Machinery and Intelligence*”. The term AI was coined in a 1956 Dartmouth workshop. See: *History of artificial*, on *Wikipedia*, available at: [https://en.wikipedia.org/wiki/History\\_of\\_artificial\\_intelligence](https://en.wikipedia.org/wiki/History_of_artificial_intelligence).
9. <https://www.edge-ai-vision.com/2025/02/global-ai-adoption-to-surge-20-exceeding-378-million-users-in-2025/>.
10. The delicate answer is that «*Google is adapting by integrating AI technologies*», see: <https://www.seo.com/ai/will-ai-replace-google/#:~:text=The%20answer%20is%20no.,unlikely%20to%20replace%20Google%20entirely>.
11. See: <https://kpmg.com/xx/en/our-insights/ai-and-technology/trust-attitudes-and-use-of-ai.html>.
12. <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai>; for EU data, see: f82f1939-8e32-4f6f-a5ba-933bb5445940\_en.
13. See: B. Marr, *The 4 Types Of Generative AI Transforming Our World*, in *Forbes*, available at: <https://www.forbes.com/sites/bernardmarr/2024/04/29/the-4-types-of-generative-ai-transforming-our-world/>. See also: [https://en.wikipedia.org/wiki/Generative\\_artificial\\_intelligence](https://en.wikipedia.org/wiki/Generative_artificial_intelligence); <https://bigid.com/blog/unveiling-6-types-of-generative-ai/>.
14. See: <https://www.imaginationtech.com/future-of-automotive/when-will-autonomous-cars-be-available/what-are-the-levels-of-autonomy-in-self-driving-cars/>.
15. See: <https://www.forbes.com/sites/jackkelly/2025/04/25/the-jobs-that-will-fall-first-as-ai-takes-over-the-workplace/>. But this might take time: see J. Burn-Murdoch, *Why hasn't AI taken your job yet?*, in *Financial Times*, available at: <https://www.ft.com/content/471b5eba-2a71-4650-a019-e8d4065b78a0>, and C. Benedikt

Frey, *AI alone cannot sole the productivity puzzle*, in *Financial Times*, available at: <https://archive.fo/eMRul#selection-2275.424-2275.464>, arguing the AI, like all new tools, will «ignite a productivity renaissance» only if it is «accompanied by breakthrough innovations» rather than reduce current workloads.

16. To note just some of the luminaries who voiced concern. See: [https://en.wikipedia.org/wiki/Artificial\\_intelligence](https://en.wikipedia.org/wiki/Artificial_intelligence).
17. Doomsday Clock moves closest ever to apocalypse – at 89 seconds to *midnight*, in *University of Chicago News*, available at: <https://news.uchicago.edu>.
18. See, e.g., <https://thesolutionsproject.org/info/climate-change-solutions/>.
19. <https://en.wikipedia.org/wiki/Malthusianism>.
20. Cf.: A. Romero, *Why AI Can't Make Human Creativity Obsolete - Recent history gives us a powerful and definitive reason for hope*, at <https://www.thealgorithmicbridge.com>.
21. I am old enough to remember the concern over human cloning (see: *Human cloning*, on *Wikipedia*, available at: [https://en.wikipedia.org/wiki/Human\\_cloning](https://en.wikipedia.org/wiki/Human_cloning)) and am following the current calls for a 10-year international moratorium on the use of CRISPR and other DNA-editing tools to create genetically modified children (see: <https://www.statnews.com>). Truly risky scientific developments can, I think, be responsibly contained.
22. See: M. Wong *The Entire Internet Is Reverting to Beta: The AI takeover is changing everything about the web – and not necessarily for the better*, in *The Atlantic*, June 18, 2025, available at: <https://www.theatlantic.com>.
23. It is different from a product like the iPhone which was highly polished since its introduction and that has (and I write this respectfully, as an iPhone user) only been incrementally improved ever since.
24. [https://en.wikipedia.org/wiki/Law\\_of\\_the\\_Horse](https://en.wikipedia.org/wiki/Law_of_the_Horse).
25. On this see, notably: G. Seidman, A. Gaon, *The Social and Legal Impact of Autonomous Vehicles - A Future without Human Driving Series*, Edward Elgar Publishing, 2025.
26. L. Floridi, *AI as Agency Without Intelligence: on ChatGPT, Large Language Models, and Other Generative Models*, in *Philosophy & Technology*, 2023, pp. 1-2. Also see: E. J. Larson, *The Myth of Artificial Intelligence: Why Computers Can't Think the Way We Do*, Harvard, 2022.
27. We should be safe, if humans – and in the future, robots – read and uphold Asimov's laws of robotics. See: [https://en.wikipedia.org/wiki/Three\\_Laws\\_of\\_Robotics](https://en.wikipedia.org/wiki/Three_Laws_of_Robotics).
28. Then we read “*Anthropic's new AI model turns to blackmail when engineers try to take it offline*”, and visions of HAL, the computer from the science fiction classic “2001, a Space Odyssey” come to mind. See: <https://techcrunch.com/2025/05/22/anthropic's-new-ai-model-turns-to-blackmail-when-engineers-try-to-take-it-offline/>; [https://en.wikipedia.org/wiki/2001:\\_A\\_Space\\_Odyssey](https://en.wikipedia.org/wiki/2001:_A_Space_Odyssey).
29. A. Clark, *Extending Minds with Generative AI*, in *Nature Communications*, 16: 4627, 2025.
30. Compared with 48.8% of Italian households that have air conditioning (2021). See:

<https://datareportal.com/reports/digital-2025-italy>;  
<https://www.businesscoot.com/en/study/the-air-conditioning-market-italy#:~:text=According%20to%20the%20Istat%20Report,2044.2%25in%20the%20Center>.

31. See, e.g., *President Trump reposted another user's false claim that the former president had been "executed" in 2020 and replaced by a robotic clone*, at: <https://www.nytimes.com/2025/06/01/us/politics/trump-biden-conspiracy-theory.html>

32. «[T]hree of the gen-AI-related risks that respondents most commonly say have caused negative consequences for their organizations»: <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai>.

33. On privacy risks, see: I. Barberá, *AI Privacy Risks & Mitigations – Large Language Models (LLMs)*, 2025, available at: <https://www.edpb.europa.eu/>.

34. See in more detail, in the *AI Action Summit International Scientific Report on the Safety of Advanced AI*, 2025, available at: <https://assets.publishing.service.gov.uk/>, and the United Nations B-Tech Project's recent paper on generative AI, the Taxonomy of Human Rights Risks Connected to Generative AI, available at: <https://www.ohchr.org/>, which explores 10 human rights that generative AI may adversely impact.

35. [https://en.wikipedia.org/wiki/Algorithmic\\_bias](https://en.wikipedia.org/wiki/Algorithmic_bias). For some troubling examples see: <https://www.digital-adoption.com/ai-bias-examples/>; <https://www.crescendo.ai/blog/ai-bias-examples-mitigation-guide>.

36. See, e.g., *Research shows AI is often biased. Here's how to make algorithms work for all of us*, available at: <https://www.weforum.org/>.

37. *The pot calling the kettle black*, on *Wikipedia*, available at: [https://en.wikipedia.org/wiki/The\\_pot\\_calling\\_the\\_kettle\\_black](https://en.wikipedia.org/wiki/The_pot_calling_the_kettle_black).

38. M. Durmus, *Mindful AI: Reflections on Artificial Intelligence*, 2022.

39. Cognitive biases are systematic patterns of deviation from norm and/or rationality in judgment. See: M. Durmus, *List of Cognitive Biases*, at: [https://www.aisoma.de/wp-content/uploads/2021/12/Cognitive-Biases\\_V4.pdf](https://www.aisoma.de/wp-content/uploads/2021/12/Cognitive-Biases_V4.pdf); also Id., *Cognitive Biases - A Brief Overview of Over 160 Cognitive Biases*.

40. *Unmasking the Bias in Facial Recognition Algorithms*, in *MIT Sloan*, available at: <https://mitsloan.mit.edu/ideas-made-to-matter/unmasking-bias-facial-recognition-algorithms>; *Facial Recognition Software Struggles to Detect Dark Skin - Here's Why & How*, in *Mozilla Foundation*, available at: <https://foundation.mozilla.org/en/blog/facial-recognition-software-struggles-to-detect-dark-skin-heres-why-and-how/>; *Mitigating Facial Recognition Bias*, on *HyperVerge Blog*, available at: <https://hyperverge.co/blog/mitigating-facial-recognition-bias/>; *How Coders Are Fighting Bias in Facial Recognition Software*, in *Wired*, available at: <https://www.wired.com/story/how-coders-are-fighting-bias-in-facial-recognition-software/>.

41. Which is a problem in of itself – especially because there are indications that AI's accuracy is not improving over time: a May 2025 report (available at:

<https://www.businesstechweekly.com>) suggests that newer AI systems are making much more mistakes than their predecessors – with error rates of up to 79%.

42. *Hallucination (artificial intelligence)*, on *Wikipedia*, available at: [https://en.wikipedia.org/wiki/Hallucination\\_\(artificial\\_intelligence\)](https://en.wikipedia.org/wiki/Hallucination_(artificial_intelligence)).
43. Recent research suggests that it is mostly self-representing laymen who present courts with such non-existent authorities. See: M. Legg, *More people are using AI in court, not a lawyer. It could cost you money – and your case*, on *The Conversation*, available at: <https://theconversation.com>. Respectfully, I think this puts the onus on the professionals in the courtroom, especially government lawyers and judges, to be hyper vigilant and double check laymen statements – much like a medical doctor cannot accept laymen's claims based on "Dr. Google" advice.
44. See, e.g., G. Del Valle, *Why do lawyers keep using ChatGPT?*, on *The Verge*, available at: <https://www.theverge.com/policy/677373/lawyers-chatgpt-hallucinations-ai>; *Thomson Reuters Future of Professionals Report*, available at: <https://www.thomsonreuters.com/reports/future-of-professionals-report-2024.pdf>.
45. See, similarly, *F. Ayinde*, R (on the application of) v The London Borough of Haringey - Find Case Law - The National Archives, available at: <https://caselaw.nationalarchives.gov.uk/ewhc/admin/2025/1383>; there is a nice twist to this increasingly common occurrence, that lawyers provide courts with non-existent authorities, the result of AI "hallucinations" not verified by the lawyers: in *Williams v. Capital One Bank*, N.A., 2025 U.S. Dist. LEXIS 49256 - U.S. District Judge Rudolph Contreras listed AI-generated fake cases in his opinion. This means that the made-up cases are now included in a real case and in the legal databases that cover it, albeit not as valid legal authorities but as part of a cautionary tale not to rely on unverified AI outputs, creating a bit of a multiple mirroring effect. What if now these fake cases surface in legal searches as legitimate?
46. Whether you are a singer or scholar-prime minister – be sure to be original! See, e.g., <https://telegrafi.com/en/Shakira-and-Carlos-Vives-appear-in-court-on-plagiarism-charges/>; <https://nationalpost.com/news/mark-carney-plagiarism-accusations>.
47. <https://nymag.com/intelligencer/article/openai-chatgpt-ai-cheating-education-college-students-school.html>.
48. And you can expect some people to give Gen. AI up, for all its blessings, to maintain our humanity, just as we many parents now realize that screens and sweets are better kept away from their children. See: O. Han, *We Need to Chat(GPT)*, in *The New York Times*, available at: <https://www.nytimes.com/2025/06/18/learning/we-need-to-chat-gpt.html> (a 16 year old's farewell letter to ChatGPT); E. Baumgaertner, *More Screen Time Means Less Parent-Child Talk, Study Finds*, in *The New York Times*, available at: <https://www.nytimes.com/2024/03/04/health/children-screen-time.html>; *How bad is sugar for kids' health? Here's what the science says*, on *National Geographic*, available at: <https://www.nationalgeographic.com/science/article/excess-sugar-health-effects-children>.
49. In terms of learning, AI can «boost learners' performance. However, these uses do not

*promote the deep cognitive and metacognitive processing that are required for high-quality learning»* (L. Yan, S. Greiff, J. M. Lodge, D. Gaševic, *Distinguishing performance gains from learning when using generative AI*, in *Nature Reviews Psychology*, 2025, available at: <https://www.nature.com/articles/s44159-025-00467-5>); in addition, there is growing evidence of the cognitive costs of AI and that premature over-reliance on AI reduces cognitive connectivity: see N. Kosmyna E. Hauptmann, Y. Tong Yuan, J. Situ, X.-H. Liao, A. V. Beresnitzky, I. Braunstein, P. Maes, Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task, on MIT Media Lab, 2025, available at: <https://www.media.mit.edu/publications/your-brain-on-chatgpt/>. Cf. N. Carr, *The Shallows: What the Internet is Doing to our Brains*, 2011.

50. V. Magesh, F. Surani, M. Dagl, M. Suzgun, C. D. Manning, D. E. Ho, *Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools*, in *Journal of Empirical Legal Studies*, 2025, available at: [https://dho.stanford.edu/wp-content/uploads/Legal\\_RAG\\_Hallucinations.pdf](https://dho.stanford.edu/wp-content/uploads/Legal_RAG_Hallucinations.pdf).
51. Cf. Lyon, *Fake Cases, Real Consequences: Misuse of ChatGPT Leads to Sanctions*, in *NY Litigator*, 2023.
52. M. Rotenberg, University of Georgetown, Executive director of the Center for AI and Digital Policy.
53. Cf. Everyone Is Using A.I. for Everything. Is That Bad? , in *The New York Times*, available at: <https://www.nytimes.com/2025/06/16/magazine/using-ai-hard-fork.html>.
54. And even this is likely too may, especially as the question of liability is yet to be fully determined. See, e.g., L. McMahon, G. Fraser, TikTok sued by parents of UK teens after alleged challenge deaths, on *BBC*, available at: <https://www.bbc.com/news/articles>; E. Keller, *She always said, "I'm going to be famous, dad": Teen dies after viral TikTok 'dusting' challenge*, in *The Independent*, available at: <https://www.independent.co.uk>; E. Clark, *Teen Dies Doing Deadly TikTok 'Dusting' Trend: These Viral Challenges Could Cost You Your Life*, in *IBTimes UK*, available at: <https://www.ibtimes.co.uk/>.
55. OpenAI, on *Wikipedia*, available at: <https://en.wikipedia.org/wiki/OpenAI>. But then things changed in what Karen Hao describes as a descent into chaos. See: K. Hao, *Inside OpenAI's Growing Pains After Launching ChatGPT: 'Empire of AI'*, on *Business Insider*, available at: <https://www.businessinsider.com> ; Empire of AI, on *Wikipedia*, available at: [https://en.wikipedia.org/wiki/Empire\\_of\\_AI](https://en.wikipedia.org/wiki/Empire_of_AI).
56. As did public criticism of Big Tech. See, e.g., K. Brennan, A. Kak, S. Myers West, *Artificial Power: AI Now 2025 Landscape*, AI Now Institute, June 3, 2025, at <https://ainowinstitute.org/2025-landscape> (arguing that “tech oligarchs” and their influence over AI pose a major threat to the public); E. M. Bender A. Hanna, *The AI Con – How to Fight Big Tech's Hype and Create the Future We Want*, 2025.
57. «[T]his demonstrates that many are feeling less positive about the ability of AI systems to provide accurate and reliable output, and be safe, secure and ethical to use. Perceived trustworthiness decreased in 13 of the 17 countries, with the largest decreases in Israel (68% to 52%) and South Africa (76% vs. 62%)», see:

<https://kpmg.com/xx/en/our-insights/ai-and-technology/trust-attitudes-and-use-of-ai.html> (p. 35).

58. *Ibid.*
59. Model collapse occurs when AI models are trained on data that includes content generated by earlier versions of themselves. Over time, this recursive process causes the models to drift further away from the original data distribution, losing the ability to accurately represent the world as it really is. Instead of improving, the AI starts to make mistakes that compound over generations, leading to outputs that are increasingly distorted and unreliable. See: B. Marr, *Why AI Models Are Collapsing And What It Means For The Future Of Technology*, on *Forbes*, available at: <https://www.forbes.com/sites/bernardmarr/2024/08/19/why-ai-models-are-collapsing-and-what-it-means-for-the-future-of-technology/>; I. Shumailov, Z. Shumaylov, Y. Zhao, N. Papernot, R. Anderson, Y. Gal, *AI models collapse when trained on recursively generated data*, in *Nature*, available at: <https://www.nature.com/articles/s41586-024-07566-y>.
60. The figure drops to 37% in advanced economies and rises to 55% of people in emerging economies. See, *supra* note 52, at pp. 5, 48.
61. *Ivi*, at p. 51.
62. In honesty, I see little chance for this happening any more than calls to reign in online and social media excesses.
63. Revealing this information is, arguably, the ethical and fair thing to do – however research suggests that disclosing usage of AI, practicing ethical transparency, might actually have a *negative* effect and erode trust. See: O. Schilke, M. Reimann, *The transparency dilemma: How AI disclosure erodes trust*, in *Organizational Behavior and Human Decision Processes*, 2025.
64. For example, U.S. Secretary of Health and Human Services R. F. Kennedy Jr. «*made targeting additives a key part of his “Make America Healthy Again” initiative to improve America’s food supply, setting off a wave of legislation and government action*», M. Spina, *Your Favorite Snacks May Soon Carry An Alarming Warning Label In Texas. Here’s What You Need To Know*, at <https://www.yahoo.com/news>.
65. Estimated at ca. 50,000 in 2025. See: F. Faye, *How Many AI Applications Exist Worldwide in 2025?*, on *Byte Plus*, available at: <https://www.byteplus.com/en/topic/536933?title=how-many-ai-applications-exist-worldwide-in-2025>.
66. What people do with AI outcomes is another matter. A New Zealand member of parliament presented the chamber with a fake nude picture of hers, created with AI, calling to criminalize the creation and sharing of non-consensual sexually explicit deepfakes. But what if this were a nude painting by a former painter boyfriend, would that be illegal too? What is the cause for concern, the ease of use of AI, where it took the parliamentarian “less than five minutes to make a series of deep fakes of myself”, or the unwanted pornography, a much older, long debated issue.
67. M. Zao-Sanders, *How People Are Really Using Gen AI in 2025* in *Top of Form Bottom of*

*Form Harvard Business Review, 2025.*

68. *Italy lifts ban on ChatGPT after data privacy improvements*, in *DW*, available at: <https://www.dw.com/en/ai-italy-lifts-ban-on-chatgpt-after-data-privacy-improvements/a-65469742>.
69. As many authoritarian regimes find out today. Cf. *81% of Iranian internet users bypass censorship with VPNs – Parliament*, at <https://wwwiranintl.com/en>.
70. Cf. *AI takes over Gripen E fighter jet in dogfight trial against real pilot*, in *Business Standard*, available at: <https://www.business-standard.com/world-news/>.
71. S. Rowan Kelleher, *No Pilot, No Problem? Here's How Soon Self-Flying Planes Will Take Off*, in *Forbes*, available at: <https://www.forbes.com/sites/suzannerowankelleher/2023/02/26/pilotless-autonomous-self-flying-planes/> and cf. S. Van Aarde, *The Case Against Autonomous Planes: Risks & Realities*, available at: <https://www.v-hr.com/blog/the-case-against-autonomous-planes/#>.
72. According to the NTSB investigations performed into air accidents, over 88 percent of all chartered plane crashes are attributed, at least in part, to pilot error. See: <https://www.wisnerbaum.com>.
73. Waymo's AV fleet logged 56.7 million miles showing fewer crashes that injure at intersections (96%), that injure pedestrians (92%), cyclists or motorcyclists (82%), or causing serious injuries (85%). See: <https://growsf.org/news/2025-05-02-waymo-safety/>.
74. S. Yoo, S. Lee, S. Kim, H. Hwangbo, N. Kang, *The Anxiety Consumers Feel About Using Robotaxis: HMI Design for Anxiety Factor Analysis and Anxiety Relief Based on Field Tests*, identify 19 major anxiety factors.
75. See: D. S. Fowle, C. Maple, *A robotaxi artificial intelligence safety failure*, <https://doi.org/10.1049/icp.2024.2527>.
76. See, e.g.: C. Jones, *With 51 reported autopilot deaths, how safe is Tesla's robotaxi?*, at <https://www.mysanantonio.com/lifestyle/travel/article/tesla-robotaxi-safety-austin-20351329.php>; C. Kirkham, *Tesla seeks to block city of Austin from releasing records on robotaxi trial*, in *Reuters*, available at: <https://www.reuters.com/business/autos-transportation/tesla-seeks-block-city-austin-releasing-records-robotaxi-trial-2025-06-06/>; C. Edwards, *US safety regulators contact Tesla over erratic robotaxis*, in *BBC*, available at: <https://www.bbc.com/news/articles/cg75zv4gny2o>; *Videos show driverless Tesla 'robotaxis' braking hard without warning - and on wrong side of the road*, on *Sky news*, available at: <https://news.sky.com/story/videos-show-driverless-teslas-braking-hard-without-warning-and-on-wrong-side-of-the-road-13388213>.
77. See, in detail: [https://en.wikipedia.org/wiki/Artificial\\_intelligence\\_in\\_healthcare](https://en.wikipedia.org/wiki/Artificial_intelligence_in_healthcare).
78. See, e.g., M. Lenharo, *Google AI better than human doctors at diagnosing rashes from pictures*, in *Nature*, available at: <https://www.nature.com/articles/d41586-025-01437-w>; D. McDuff, M. Schaeckermann, T. Tu, *Towards accurate differential diagnosis with large*

*language models*, at <https://doi.org/10.1038/s41586-025-08869-4>; G. Kolata, *A.I. Chatbots Defeated Doctors at Diagnosing Illness*, in *The New York Times*, available at: <https://www.nytimes.com/2024/11/17/health/chatgpt-ai-doctors-diagnosis.html>.

79. *How AI is improving diagnostics and health outcomes*, available at: <https://www.weforum.org/stories/2024/09/ai-diagnostics-health-outcomes/>.

80. K. Hetter, *Doctor explains how artificial intelligence is already being deployed in medical care*, in *CNN*, available at: <https://edition.cnn.com/2025/03/27/health/artificial-intelligence-diagnosis-technology-wellness>.

81. See, e.g., *The EU AI Act is here: requirements for healthcare organizations* (more on the EU AI Act, see *infra* in notes; also, more generally: <https://www.nature.com/articles/s43856-024-00492-0>; <https://www.nature.com/articles/s41746-024-01300-8>).

82. The discussion is wide and complex. A nice starting point is D. J. Gunkel, *Person Thing Robot - A Moral and Legal Ontology for the 21<sup>st</sup> Century and Beyond*, in *MIT*, 2023; an example of more complex argumentation is the suggestions that *cognition* is not synonymous with *consciousness*, see: N. K. Hayles, *Bacteria to AI - Human Futures with Nonhuman Symbiosis*, 2025.

83. And what if we are wrong to presume human superiority, for example, in terms of creativity and personhood over animal and machine? See on these ideas: J. Gibson, *Wanted, More than Human Intellectual Property - Animal Authors and Human Machines*, 2025.

84. *Chauvinism*, on *Wikipedia*, available at: <https://en.wikipedia.org/wiki/Chauvinism>.

85. See: *Forced confession*, on *Wikipedia*, available at: [https://en.wikipedia.org/wiki/Forced\\_confession](https://en.wikipedia.org/wiki/Forced_confession); S. Penney, *Theories of Confession Admissibility: A Historical View*, in *Am. J. Crim. L.*, 1997-1998, p. 309; cf. D. C. Flatto *Evidently Not: Why Confessions Are Excluded in Jewish Criminal Jurisprudence*, in *J. L. & Religion*, 39(2), 2024, p. 173.

86. See: M. J. Steele, J. M. Chin, C. van Golden, *Witness Preparation and the Corruption of Memory: A Survey of Australian Trial Judges*, in *Melbourne U.L.R.*, 2024, p. 152.

87. According to a recent study, robotaxi passengers rated the experience an 8.53 out of 10. See: J.D. Power study finds, *Trust in robotaxis is higher among users*, *J.D. Power study finds*, at: <https://finance.yahoo.com/news/trust-in-robotaxis-is-higher-among-users-jd-power-study-finds-193350461.html>.

88. Especially as Gen. AI picks up better and greater skills. See, e.g. S. Murphy Kelly, *ChatGPT passes exams from law and business schools*, in *CNN*, January 26, 2023, available at: <https://edition.cnn.com/2023/01/26/tech/chatgpt-passes-exams>; D. Charlotin, N. Ridi, *GenAI as an International Lawyer: A Case Study with the Jessup International Law Moot Court*, available at: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=5283722](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5283722) (studying the capacity of Gen. AI, specifically Large Language Models, to craft compelling

international legal arguments).

89. See: V. Jeutner, *The Reasonable Person - A Legal Biography*, 2024, pp. 132, 142-151.
90. See: European Parliament, EU AI Act: first regulation on artificial intelligence, at <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>. See: Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), available on *EUR-Lex*.
91. Cf. J. Thornhill, *US AI laws risk becoming more 'European' than Europe's*, in *Financial Times*, May 12, 2025, available at: <https://www.ft.com/content/aed82f47-b441-4bb3-930e-eca10585fc6d>.
92. AIPRM, AI Laws Around the World, at <https://www.aiprm.com/ai-laws-around-the-world/>.
93. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation); see: *General Data Protection Regulation*, on Wikipedia, available at: [https://en.wikipedia.org/wiki/General\\_Data\\_Protection\\_Regulation](https://en.wikipedia.org/wiki/General_Data_Protection_Regulation).
94. Recital 27: «*Transparency means that AI systems are developed and used in a way that allows appropriate traceability and explainability...*» (emphasis – added). Also see Art. 13.
95. Recital 59: «*Furthermore, the exercise of important procedural fundamental rights, such as the right to an effective remedy and to a fair trial as well as the right of defence and the presumption of innocence, could be hampered, in particular, where such AI systems are not sufficiently transparent, explainable and documented*» (emphasis – added).
96. Recital Art. 65. Also see: Art. 58, Annex IV, and Annex XI.
97. Section 86 which, generally, demands that where a person's fundamental rights might be affected by a decision using AI, the person «*shall have the right to obtain from the deployer clear and meaningful explanations of the role of the AI system in the decision-making procedure and the main elements of the decision taken*» (subsection 1).
98. See: Recital Art. 71 GDPR.
99. M. E. Kaminski, G. Malgie, *The Right to Explanation in the AI Act*, on SSRN, available at: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=5194301](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5194301).
100. Court of Justice, Judgment of February 27, 2025, C-203/22, *Dun & Bradstreet Austria GmbH*, ECLI:EU:C:2025:117. See: M. Rotenberg, *Brilliant Decision from the EU High Court on Algorithmic Transparency*, on *LinkedIn*, available at: [https://www.linkedin.com/posts/marc-rotenberg\\_db-austria-c-20322-cjeu-2025-activity-7300911470651867137-kqn7/?rcm=ACoAABSwbIIBOu1MZyb7SgW1h8ahMxAfNfXIRE](https://www.linkedin.com/posts/marc-rotenberg_db-austria-c-20322-cjeu-2025-activity-7300911470651867137-kqn7/?rcm=ACoAABSwbIIBOu1MZyb7SgW1h8ahMxAfNfXIRE). For analysis of more cases across Europe see: L. Metikoš, J. Ausloos, *The right to an explanation in*

*practice: insights from case law for the GDPR and the AI Act*, in *Law, Innovation and Technology*, 2025, p. 581.

101. *Regulators are focusing on real AI risks over theoretical ones. Good*, in *The Economist*, available at: <https://www.economist.com/leaders/2024/08/22/regulators-are-focusing-on-real-ai-risks-over-theoretical-ones-good>.
102. The AI generated cat picture below is from Pixabay.com (<https://pixabay.com>). See: [https://www.bing.com/images/search?view=detailV2&ccid=qOBw17Ep&id=DB2D00EB62BA62D383C4A70D517650A6D37E0BF3&thid=OIP.qOBw17EpebS4evqXlzJgRAHaE7&mediaurl=https%3a%2f%2fdcdn.pixabay.com%2fphoto%2f2023%2f06%2f11%2f20%2f19%2fai-generated-8056852\\_1280.jpg&cdnurl=https%3a%2f%2fth.bing.com%2fth%2fid%2fR.a8e070d7b12979b4b87afa9797326044%3frlik%3d8wt%252b06ZQdlENpw%26pid%3dImgRaw%26r%3d0&exph=853&expw=1280&q=ai+cat&csimid=607993900778612318&FORM=IRPRST&cck=BC73541390A806CCBE0B5F92D1012387&selectedIndex=1&itb=0](https://www.bing.com/images/search?view=detailV2&ccid=qOBw17Ep&id=DB2D00EB62BA62D383C4A70D517650A6D37E0BF3&thid=OIP.qOBw17EpebS4evqXlzJgRAHaE7&mediaurl=https%3a%2f%2fdcdn.pixabay.com%2fphoto%2f2023%2f06%2f11%2f20%2f19%2fai-generated-8056852_1280.jpg&cdnurl=https%3a%2f%2fth.bing.com%2fth%2fid%2fR.a8e070d7b12979b4b87afa9797326044%3frlik%3d8wt%252b06ZQdlENpw%26pid%3dImgRaw%26r%3d0&exph=853&expw=1280&q=ai+cat&csimid=607993900778612318&FORM=IRPRST&cck=BC73541390A806CCBE0B5F92D1012387&selectedIndex=1&itb=0).
103. The largest (now) private employer in Italy, ENEL, has 65,000 employees, and revenues of US \$ 148 billion (2022) (see: [https://en.wikipedia.org/wiki/List\\_of\\_largest\\_Italian\\_companies](https://en.wikipedia.org/wiki/List_of_largest_Italian_companies)); there are over 3 million government employees in Italy (see: <https://www.statista.com/topics/8540/government-employment-in-italy/#:~:text=Over%20three%20million%20people%20work,female%20employees%20are%20the%20majority..>), and Italy's tax revenue for September 2024 alone was over US dollar 169 billion (see: <https://www.ceicdata.com/en/indicator/italy/tax-revenue#:~:text=Italy%20Tax%20Revenue%20was%20reported,Sep%202024%2C%20with%20103%20observations.>).
104. L. Floridi, *Editor Letter - AI as Agency without Intelligence: On Artificial Intelligence as a New Form of Artificial Agency and the Multiple Realisability of Agency Thesis*, in *Philosophy & Technology*, 2025.
105. *Ivi*, pp. 17-18 (cite – 18; spelling – in the original).
106. *Ivi*, p. 18.
107. See: E. Kantorowicz, *The King's Two Bodies*, 1957 (see: [https://en.wikipedia.org/wiki/The\\_King%27s\\_Two\\_Bodies](https://en.wikipedia.org/wiki/The_King%27s_Two_Bodies)).
108. The issued Executive Order 13859 in February 2019. It established the American AI Initiative that identified five key lines of effort «*including increasing AI research investment, unleashing Federal AI computing and data resources, setting AI technical standards, building America's AI workforce, and engaging with international allies. These lines of effort were codified into law as part of the National AI Initiative Act of 2020*» (*Artificial Intelligence for the American People*, available at: <https://trumpwhitehouse.archives.gov/ai/>).
109. N. Lee, R. Huffman, R. Burnette, A. Gweon, *April 2025 AI Developments Under the Trump Administration*, May 2025, available at: <https://www.insidegovernmentcontracts.com/2025/05/april-2025-ai-developments-unde>

r-the-trump-administration/.

110. N. Shafabafy, A. O’neil, *America first, ethics second: The implications of Trump’s AI Executive Order, on the interpreter*, available at: <https://www.lowyinstitute.org/the-interpreter/america-first-ethics-second-implications-trump-s-ai-executive-order>.
111. *White House national security memo asks military to increase use of AI*, in *The Washington Post*, available at: <https://www.washingtonpost.com/technology/2024/10/24/white-house-ai-nation-security-memo/>.
112. B. Begehot, *How the British Government Rules by Algorithm*, in *The Economist*, available at: <https://www.economist.com/britain/2020/08/20/how-the-british-government-rules-by-a-lgorithm>.
113. R. Mason, *AI should replace some work of civil servants, Starmer to announce*, in *The Guardian*, available at: <https://www.theguardian.com/technology/2025/mar/12/ai-should-replace-some-work-of-civil-servants-under-new-rules-keir-starmer-to-announce>.
114. As noted, *supra*, the OECD has identified over 1,000 AI policy initiatives from 69 countries. See: *OECD’s live repository of AI strategies & policies*; see also: *Artificial intelligence in government*, on *Wikipedia*, available at: [https://en.wikipedia.org/wiki/Artificial\\_intelligence\\_in\\_government](https://en.wikipedia.org/wiki/Artificial_intelligence_in_government).
115. See: *How US tech giants’ AI is changing the face of warfare in Gaza and Lebanon*, on *AP News*, available at: <https://apnews.com/article/israel-palestinians-ai-weapons-430f6f15aab420806163558732726ad9> (cites); and also: *Questions and Answers: Israeli Military’s Use of Digital Tools in Gaza*, on *Human Rights Watch*, available at: <https://www.hrw.org/news/2024/09/10/questions-and-answers-israeli-militarys-use-digital-tools-gaza>; *Artificial Intelligence in the Battlefield: A Perspective from Israel*, on *Opinio Juris*, available at: <https://opiniojuris.org/2024/04/20/artificial-intelligence-in-the-battlefield-a-perspective-from-israel/>; *AI-assisted targeting in the Gaza Strip*, on *Wikipedia*, available at: [https://en.wikipedia.org/wiki/AI-assisted\\_targeting\\_in\\_the\\_Gaza\\_Strip](https://en.wikipedia.org/wiki/AI-assisted_targeting_in_the_Gaza_Strip); *Military applications of artificial intelligence*, on *Wikipedia*, available at: [https://en.wikipedia.org/wiki/Military\\_applications\\_of\\_artificial\\_intelligence#:~:text=AI%20technologies%20enable%20coordination%20of,both%20human%20operated%20and%20autonomous](https://en.wikipedia.org/wiki/Military_applications_of_artificial_intelligence#:~:text=AI%20technologies%20enable%20coordination%20of,both%20human%20operated%20and%20autonomous).
116. See: H. W, Meerveld, R. H. A. Lindelauf, E.O. Postma, M. Postma, *The Irresponsibility of not using AI in the military*, in *Ethics and Information Technology*, 25(14), 2023, at: <https://doi.org/10.1007/s10676-023-09683-0>.
117. *How Policing Agencies Use AI*, available at: <https://www.policingproject.org/ai-explained-articles/2024/9/6/how-policing-agencies-u>

se-ai (cite); also: *Policing and artificial intelligence*, available at: <https://www.police-foundation.org.uk/publication/policing-and-artificial-intelligence/#mob-menu>; *AI and policing - The benefits and challenges of artificial intelligence for law enforcement*, available at: <https://www.europol.europa.eu/publication-events/main-reports/ai-and-policing>.

118. «*Countries aren't only using AI to organize quick responses to crime – they're also using it to predict crime. The United States and South Africa have AI crime prediction tools in development, while Japan, Argentina, and South Korea have already introduced this technology into their policing*», *How Countries Are Using AI to Predict Crime*, available at: <https://swisscognitive.ch>.

119. *Police use controversial AI tool that looks at people's sex lives and beliefs*, available at: <https://www.msn.com/en-nz/news/other/police-use-controversial-ai-tool-that-looks-at-people-s-sex-lives-and-beliefs/ar-AA1GOCWV?ocid=BingNewsVerp>.

120. D. MacMillan, *New Orleans pushes to legalize police use of 'facial surveillance'*, in *The Washington Post*, available at: <https://www.washingtonpost.com/business/2025/06/12/facial-recognition-new-orleans-artificial-intelligence/>.

121. *How AI-powered tech landed man in jail with scant evidence*, on *AP News*, available at: <https://apnews.com/article/artificial-intelligence-algorithm-technology-police-crime-7e3345485aa668c97606d4b54f9b6220>.

122. <https://innocenceproject.org/news/when-artificial-intelligence-gets-it-wrong/>.

123. *NIST Study Evaluates Effects of Race, Age, Sex on Face Recognition Software*, available at: <https://www.nist.gov/news-events/news/2019/12/nist-study-evaluates-effects-race-age-sex-face-recognition-software>.

124. *In Madhya Pradesh, cops are using AI to modernise: 'We need a new type of policing'*, in *India News - The Indian Express*, available at: <https://indianexpress.com/article/india/madhya-pradesh-police-ai-use-modernise-10082265/>; *Eye spy! Delhi Police sharpens surveillance in digital age AI image; to use next-gen Netra to identify potential threats*, available at: <https://www.msn.com/en-in/money/news/eye-spy-delhi-police-sharpens-surveillance-in-digital-age-ai-image-to-use-next-gen-netra-to-identify-potential-threats/ar-AA1H0mv?ocid=BingNewsVerp>; *Pattaya Tourist Police use AI and local tips to bust human trafficking ring on Walking Street*, available at: <https://www.pattayamail.com/news/pattaya-tourist-police-use-ai-and-local-tips-to-bust-human-trafficking-ring-on-walking-street-505816>.

125. Cf. A. Cevallos, *How Autocrats Weaponize AI – And How to Fight Back*, on *Journal of Democracy*, available at: <https://www.journalofdemocracy.org>; Y. Gostoli, *Turkey's AI-Powered Protest Crackdown*, on *New Lines Magazine*, available at: <https://newlinesmag.com/spotlight/turkeys-ai-powered-protest-crackdown/>; P. Dizikes, *How an "AI-tocracy" emerges*, on *MIT News*, available at: <https://news.mit.edu/2023/how-ai-tocracy-emerges-0713>; *Authoritarian AI*, available at: <https://news.mit.edu/2023/authoritarian-ai-0713>.

<https://www.sobi.uni-passau.de/en/political-communication/research/authoritarian-ai>.

126. Deloitte US, *AI Use Cases in Government*, available at: <https://www.deloitte.com/us/en/what-we-do/capabilities/applied-artificial-intelligence/articles/ai-dossier-government-public-services.html>; *AI in Government: Examples & Challenges*, available at: <https://research.aimultiple.com/ai-government/>; *How Governments are Using AI: 8 Real-World Case Studies*, available at: <https://blog.govnet.co.uk/technology/ai-in-government-case-studies>.

127. In poor countries, such as Nigerian, chatbots based on GPT-4 have been tested for school education – and ultra-sensitive area for AI use – with surprisingly good results. See: *Can AI be trusted in schools?*, in *The Economist*, available at: <https://www.economist.com/graphic-detail/2025/05/30/can-ai-be-trusted-in-schools>.

128. *Post Office Horizon scandal: Why hundreds were wrongly prosecuted*, in *BBC*, available at: <https://www.bbc.com/news/articles/c1wpp4w14pqq>.

129. D. Kolkman, “F\*\*k the algorithm”?: *What the world can learn from the UK’s A-level grading fiasco*, available at: <https://blogs.lse.ac.uk/impactofsocialsciences/>.

130. *How a Discriminatory Algorithm Wrongly Accused Thousands of Families of Fraud*, available at: <https://www.vice.com/en/article/how-a-discriminatory-algorithm-wrongly-accused-thousands-of-families-of-fraud/> (cite); 20201217\_eindverslag\_parlementaire\_ondervragingscommissie\_kinderopvangtoeslag.pdf; also: L. M Haitsma, *Addressing discrimination in algorithmic profiling: Examining risk governance in Dutch public social security agencies*, in *European Journal of Social Security*, 2025, p. 1.

131. *Nevada Used A.I. to Find ‘At-Risk’ Students. Numbers Dropped by 200,000*, in *The New York Times*, available at: <https://www.nytimes.com/2024/10/11/us/nevada-ai-at-risk-students.html>.

132. *Algorithms Policed Welfare Systems For Years. Now They’re Under Fire for Bias*, in *WIR ED*, available at: <https://www.wired.com/story/algorithms-policed-welfare-systems-for-years-now-theyre-under-fire-for-bias/>.

133. Who’s homeless enough for housing? In San Francisco an algorithm decides, at: [https://www.codastory.com/authoritarian-tech/san-francisco-homeless-algorithm/?utm\\_source=digg](https://www.codastory.com/authoritarian-tech/san-francisco-homeless-algorithm/?utm_source=digg)

134. See: N. A. Smuha, *Algorithmic Rule By Law: How Algorithmic Regulation in the Public Sector Erodes the Rule of Law*, 2025. The United Arab Emirates «set to use AI to write laws in world first. Gulf state expects move will speed up lawmaking by 70% but experts warn of ‘reliability’ issues with artificial intelligence», available at: <https://archive.md/ubZAW#selection-1587.0-1587.124>. Other experts say this isn’t necessarily a terrible idea. <https://foreignpolicy.com/2025/05/14/ai-generated-law-uae-legislation/>

135. And you realize that when you see that even the judiciary, arguably the most conservative

branch of government, is beginning to familiarize itself with AI. See, e.g., K. Firth-Butterfield, K. Silverman, “*Artificial Intelligence – Foundational Issues and Glossary.*” *Artificial Intelligence and the Courts: Materials for Judges. American Association for the Advancement of Science*, 2022, available at: <https://doi.org/10.1126/aaas.adf0782>; European Commission for the Efficiency of Justice, 1st Report on the use of Artificial Intelligence (AI) in the judiciary, based on the information contained in the CEPEJ’s Resource Centre on Cyberjustice and AI, available at: <https://www.coe.int/en/web/cepej/-/1st-report-on-the-use-of-artificial-intelligence-ai-in-the-judiciary-based-on-the-information-contained-in-the-cepej-s-resource-centre-on-cyberjustice-and-ai>; VV.AA., *Artificial Intelligence, Judicial Decision-Making and Fundamental Rights*, 2024, available at: [https://ssm-italia.eu/wp-content/uploads/2025/02/JuLIA\\_handbook-Justice\\_final.pdf](https://ssm-italia.eu/wp-content/uploads/2025/02/JuLIA_handbook-Justice_final.pdf); AI and the Rule of Law: Capacity Building for Judicial System, available at: <https://www.unesco.org/en/artificial-intelligence/rule-law/mooc-judges>.

136. Rushing into reform without setting in place proper regulation seems is unfortunate. It seems to be the main reason for Thailand’s recent rollback of its cannabis legalization. See, e.g., *Is weed still legal in Thailand? Here’s what tourists need to know as government u-turns*, available at: <https://www.euronews.com/travel/2024/11/07/is-weed-still-legal-in-thailand-heres-what-tourists-need-to-know-as-government-u-turns>.
137. Rights groups warn UK Home Office AI tool risks ‘rubberstamped’ migrant deportations, available at: <https://www.aa.com.tr/en/europe/rights-groups-warn-uk-home-office-ai-tool-risks-rubberstamped-migrant-deportations/> (cite); Privacy International, *Automating the hostile environment: uncovering a secretive Home Office algorithm at the heart of immigration decision-making*, available at: <https://privacyinternational.org/news-analysis/5452/automating-hostile-environment-uncovering-secretive-home-office-algorithm-heart>; also: J. Maxwell, J. Tomlinson, *Experiments in Automating Immigration Systems*, Bristol, 2022.
138. Cites: *UK A-level algorithm fiasco a global example of what not to do - what went wrong and why*, available at: <https://diginomica.com/uk-level-algorithm-fiasco-global-example-what-not-do-what-went-wrong-and-why>, and D. Kolkman, “*F\*\*k the algorithm*??: *What the world can learn from the UK’s A-level grading fiasco*, cit.
139. See: I. Parker, L. Carter, Licence to build, available at: <https://www.adalovelaceinstitute.org/policy-briefing/licence-to-build/>.
140. «*This includes providing information on algorithmic tools and algorithm-assisted decisions in a complete, open, understandable, easily-accessible, and free format*», see *Algorithmic Transparency Recording Standard Hub*, in GOV.UK, available at: <https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub>.

141. For a useful early overview see: Ada Lovelace Institute, *Algorithmic Accountability for the Public Sector*, 2021.
142. R. Mackenzie, G. Scott, L. Edwards, *The Inscrutable Code? The Deficient Scrutiny Problem of Automated Government*, in *Technology and Regulation*, 2025, p. 37.
143. *Ivi*, p. 38.
144. *Ivi*, p. 42.
145. *Ivi*, p. 42 (cites; spelling – in the original), pp. 53-58.
146. [https://en.wikipedia.org/wiki/Natural\\_justice#Right\\_to\\_a\\_fair\\_hearing](https://en.wikipedia.org/wiki/Natural_justice#Right_to_a_fair_hearing).
147. For a wider discussion see: G. Lawson, G. Seidman, *Downsizing the Right to Petition*, in *Nw. U. L. Rev.*, 1999, pp. 739-766.
148. D. E. Mathew, et al., *Recent Emerging Technologies in Explainable Artificial Intelligence to Enhance the Interpretable and Understanding of AI Models for Humans*, in *Neural Processing Letters*, 2025, p. 16. Also: *Explainable AI (XAI) in 2025: Guide to enterprise-ready AI*, at <https://research.aimultiple.com/xai/>.
149. See, e.g., S. E. Biber, *Between Humans and Machines: Judicial Interpretation of the Automated Decision-Making Practices in the EU*, in H.C.H. Hofmann, F. Pflücke (eds.), *Governance of Automated Decision-Making and EU Law*, OUP, Oxford (UK), 2024, p. 186.
150. S. Kessler, *When It Comes to Spotting Fake Receipts, It's A.I. vs. A.I.*, in *The New York Times*, Sep. 6, 2025, available at: <https://www.nytimes.com/2025/09/06/business/dealbook/ai-receipts-expense-reports.html>.