

Extract from the book by Luca Bolognini  
“A.I. Artificial Insanity – Reflections on the resilience of human intelligence”  
(ISBN 8849855834 - 2018)  
<https://books.google.com/books?id=3IBiDwAAQBAJ>

# **A.I. - Artificial Insanity**

## **[Extract]**

Food for thought and reflections on the resilience of human intelligence

By Luca Bolognini ([l.bolognini@istitutoprivacy.it](mailto:l.bolognini@istitutoprivacy.it))

Extract from the book by Luca Bolognini  
“A.I. Artificial Insanity – Reflections on the resilience of human intelligence”  
(ISBN 8849855834 - 2018)  
<https://books.google.com/books?id=3IBiDwAAQBAJ>

### **Extract from Chapter 3 - Artificial Intelligence and Artificial Insanity: catch me if you can**

[...]

Let me sketch you a picture of extreme IoTisation and artificialization (my apologies in advance to the *piranhas* for this overly simplistic, jargon-free description): imagine that, within a few short years, all objects will have a soul. “Digital animism”, *credo ut intelligam* (I believe so that I may understand). Anything from a fork to a toothbrush, from a pillow to an armchair, from a banister to a t-shirt, from a bottle to a glass, from a nappy to a pair of Bermuda shorts, from a comb to a headband, from a ring to an earring, from a toy to a shoe, from a bra to a pair of glasses, from rosary beads to a vase. From a gull-drone to a robot-lifeguard. From a sex-toy to a condom, from a key to a lock, from a blender to an ice bucket. From a syringe to a vial, from a catheter to a drip, from a scalpel to a plaster. More and more again: even the objective parts of our own human bodies, from our blood to our cell tissue, could be IoTised, made intelligent and interlinked on the Internet. All these animated “things” will become smart, equipped with the ability to process information, data and results, and to communicate with each other and with humans. The intelligence, the brain, of these things could be either localised, inside them, or available remotely, in servers (computers) stored far away inside data centres (buildings that contain thousands of servers) following the paradigm of cloud computing. A cloud of small computers which, by processing together fragments of data dynamically and elastically, and adapting themselves to the strains and peak demands of faraway objects-devices, will rain down intelligence on otherwise stupid “things”. A sort of “phone-a-friend lifeline”, rendering seemingly simple and innocuous gadgets intellectually powerful. Our lives will be immersed and pervaded, invaded and transformed into a bio-digital whole, so much so that today there is more talk of BIoT than IoT.

Am I exaggerating? Maybe so, but – having studied these issues from a legal and ethical standpoint – I would not be so sure, if I were you. For me, artificial

insanity is more fascinating and more frightening than artificial intelligence. The moment the machines “go haywire”, cracking up under the complexity of their super-intelligent calculations and heightened sensitivity. As long as they are rational, we will be able to control them: it is their irrationality, their artificial mental imbalance, even just their whims, that will be the real challenge. Months ago, a media sensation was created when the news broke that two Facebook robots had apparently started speaking a new unknown language of their own invention: this reportedly led to them being shut down by FB researchers. Newspapers all over the world picked up the story and let their fervid imaginations run wild (it appears what really happened was much less sci-fi: the chatbots did not make up a new language while chattering away amiably, but simply stalled due to a programming bug). Nonetheless, we can and we must expect similar things to happen for real in the future. Can you imagine the risks to human life that might be posed by a spiteful or petulant robot, a love-struck machine, a neurotic refrigerator, a schizophrenic car or a temperamental thermometer?

Elon Musk, the founder of Tesla and other brilliant initiatives, said in 2017 that in 2037 having a car with a steering wheel in your garage will be like owning a horse. Debates have begun on major ethical and moral dilemmas: if a self-driving car has to decide who to save, in an accident where it can choose to brake or swerve, who should it save?

Yet, new technology lawyers still continue to overdebate basic issues, “trifling” matters I would say, which could be managed through tighter compulsory insurance schemes. For example, we get tangled up in knots over civil liability (in a nutshell, who should pay for damage) in accidents caused by self-driving cars. The car designer? The software? The vehicle manufacturer? The owner? The passenger (who, at that point, is no longer the “driver”)? There is even one school of thought, defined as “zoological”, which holds the car owner accountable as is the case with pet owners. That would be all well and good, if that were the heart of the matter. The problem lies elsewhere. I see it in the relationship between individuals, objects and power, or rather powers. Let me explain, trying to keep it simple.

Today there is a human being (or an organisation of human beings: a company, association or body) that has control over the things and the animals belonging to them (in part also over any minors if they are their legal guardians). There are rules (laws), hanging high above the human being, like chandeliers on the ceiling of power of the States, public institutions and international organisations. If any human individual – or plural entity – breaks the rules, the public power (be it state, super-state or sub-state) punishes them.

Matters have, naturally, become more complicated with multinational companies and borderless transnational technologies making it increasingly difficult to carry out *enforcement* (impose rules and punish offenders) with just the small power of individual nations. Italy, France or Japan, by themselves, would have an extremely hard time trying to force a web giant to do or not do something, as they would trying to fine them. There would need to be international agreements, extended as widely as possible to include all countries worldwide, to make certain regulations more relevant and effective (for years there has been talk of an Internet Bill of Rights, to name one, a most enjoyable but completely unrealistic intellectual exercise). I see it at work in this period with the EU General Data Protection Regulation (GDPR), valid from 25 May 2018, which on paper extends the requirement to comply with European privacy laws to all companies established outside Europe that carry out monitoring activities or offer goods and services to people located in the EU. I say on paper, since it will not be easy for the German or Spanish authorities to slap administrative sanctions and fines on a company based in Indonesia. What do you do if that country refuses to cooperate, send in the fighter jets?

The new EU Privacy Regulation also vaunts a principle (of so-called accountability) which aims to hold data controllers and data processors liable for the logical, technical and organisational security measures they use to protect systems and personal data, and for compliance with the requirements and restrictions laid down by that Regulation. The idea is: if you (individual or company) decide whether, why and how to process or store data relating to other individuals, then you are responsible for what you do with that data and you have a

duty to demonstrate, at the request of the public authorities, that you acted properly and in full compliance with the regulations.

This already complex tableau is further disrupted by the intrusion of objects and all their potential intelligence, wiliness, stupidity and folly. One object. Two objects. Ten objects. A hundred objects. A thousand objects. A hundred thousand objects. A million objects. A billion objects. A hundred billion objects. Physical. Virtual. Physical and virtual, interlinked together. Up to now we have been accustomed to data controllers and data processors, or more generally “centres of legal imputation” (and of liability, more to the point), in every area of the law where there were and are human beings or organisms formed by human beings: a limited company is a legal entity and not a natural person, sure, but it is still formed by people, at least as far as its administrative body is concerned. Votes and decisions, around the table at Board meetings or in the General Assembly, are always taken by a natural person, a representative perhaps but a human being nonetheless. With the Internet of (Intelligent) Things, we will have to shift paradigm. Talking about *accountability* for people and organisations of people will seem like a walk in the park: imagine trying to hold an object accountable, dearest public authorities. An object that reasons, captures data, processes and transforms it, exchanges it with other objects, receives it back and processes it again. An object that makes decisions, allows one thing to happen and prevents another thing from happening. In short, an object that has a bearing on the world and an impact on us humans. Our front door won't open any more (or is flung wide open while we're away on holiday) because the algorithm decided so. Insulin is injected in double the dosage, killing the patient, because the artificial intelligence that oversees that delicate telemedicine procedure has run amok or messed up the data, or perhaps because it is sulking (“*Stress-induced burn-out and touchiness of electronic medical devices and healthcare files*”, the title of a master's thesis twenty or thirty years from now at the Biomedigital Faculty of the University of Mars).

Try catching an intelligent object. Try giving it a fine. Good luck with that. We perpetuate the illusion that there is always a human being somewhere, an owner who can be held accountable for the misdeeds of an object. Not so. The autonomy

of things may suffice. Already today, in IoTisation projects I am working on as a lawyer in the financial, industrial and retail sectors, let me tell you there are countless different human beings and/or organisations of human beings (companies) involved in the design, development, construction, distribution, management and maintenance of IoT systems and the data processing that goes on behind the scenes. Good luck figuring out who is liable for what and when. Not the most authoritative opinion from an expert privacy lawyer, I know – the piranhas will be sharpening their teeth, ready to pounce. But, frankly, it is the most sensible comment I could come up with.

I am talking about “smart” objects that can either be entirely virtual and immaterial (that don’t exist physically, therefore, but that still have an impact on other physical objects) or both physical and virtual. In one of the later chapters of this brief and light-hearted book, I will talk instead about objects that are necessarily physical (sensors, in particular) and how we can defend ourselves against their attacks on our data. Coming back to mere artificial intelligence, I can hypothesise scenarios in which an autonomous algorithm grants itself the legal power to manage complex organisms. We could go to a notary, establish a company or a political party and agree in writing that, after the first decision in the Deed of Incorporation made by the human founders, all subsequent deliberations will be performed by algorithm XYZ “*which is attached to this deed and is to be considered an integral and essential part of the Statute*”. That moment will mark the beginning of a new era of inhuman and autonomous companies and political parties, capable of self-operating, self-directing, and maybe even self-reproducing, just like any other informatics executables. Apparently there is at least one political movement in Italy enthralled by this prospect.

If the algorithm is a code, the code is law (thanks, Lawrence Lessig). Law and informatics blur into one another even in semantics. At a conference a few years ago, an excellent digital criminal defence lawyer from Turin, Carlo Blengino, spoke to me about “auto-installing rules” and that metaphor really stuck with me. Do laws derive solely from public authorities? Clearly not. An algorithm is a rule that can have a variety of effects on the dynamics of the outside world, it can discriminate

between good and bad things in life, and have an impact on an individual or a whole community. Remember Max Weber and his definition of State which included, as a core concept, the “monopoly of the legitimate use of physical force, in the enforcement of its order”? This has never been more obsolete, as a definition, than in the context of extreme IoT. If the something physical is made to happen in the world – a door is made to open or not open, to stick with the simplest example – on the basis of the rules of an algorithm (an algorithm that may well have been programmed and decided by other algorithms and other non-human objects), does that not correspond to the use of force, legitimate in that it enforces the order which the algorithm itself aided or advocated?

The relationship between individual – citizen, consumer, person with various roles, groups of people – and power becomes more complicated: it becomes less and less about the relationship between private (individual, company) and public (state, etc.) and more about the relationship between passive subject (individual, company, other objects, even states and public institutions) and active object (algorithm, bot, robot, inhuman entity). While it was only humans designing the algorithms for assessing income, sales and purchases, to combat tax evasion, on the basis of laws approved by parliaments made up of other humans, there were grounds for protest because there was *someone* who – in a potentially biased way – assigned a value to each commodity using their own discretion and personal criteria (if you spend too much on this or that commodity, you are flagged as an anomaly and end up on a list of presumed tax evaders: a presumption that is open to criticism, since each of us should remain free to make our own unique lifestyle and consumer choices). Just think, though: what if it was not *someone* programming the algorithm, but *something*?

## **Chapter 9 - E-War and E-Peace: towards the “rule of human law”**

One evening several months ago, my family was caught up in an almost impossible and embarrassing mission. We were all gathered in a country house, in the Apennine hills less than an hour from Bologna. We live in Rome but I take my daughters to those hills every chance I get, because for me they are part and parcel of the city that I miss so deeply, a place filled with memories and meanings that never fails to rejuvenate me. There we were, having dinner in the old stone-built dining room surrounded by all our relatives, when suddenly out of the blue Matteo started crying. Matteo actually had another name and he wasn't a real baby at all – but don't tell my youngest daughter that, she would be mortally offended and would not believe you anyway. In reality, Matteo was a new generation doll, connected to the internet via Wi-Fi, fitted with an electronic chipset that made him “evolve” day after day simulating the needs of a real new-born baby. Nothing extraordinary, we're not talking about intelligence but, once again, about advanced artificial stupidity. And so, this adorable digital doll was crying.

Everyone's first reaction, calm and good-humoured as we were, was to ask my daughter to make the crying stop. A request that fell on deaf ears, not because my daughter obstinately refused to listen but because, effectively, she had no idea what to do. All of us tried in vain to calm him down: Matteo, who was screaming and wailing with an ear-splitting realism worthy of a horror film, had no off button (once removed from his box, he was “born”). The battery was rechargeable and buried deep inside his little body, impossible to remove. I can assure you that, after the first ten minutes of exasperated attempts, our nerves were torn to shreds: evil thoughts began to crowd our minds as to how to shut the shrieking gizmo up. Finally, at the end of my tether, I carried Matteo out into the woods behind the house, far enough away so his cries could not be heard. The next morning we went to fetch him and he was still there, his battery dead, the wolves had not carried him off. But this incident made me think. I will come back to Matteo and his inconsolable bawling in the next few pages. For now, let's talk about war – which has sadly been an all too familiar refrain for the Bolognese hills over the past centuries.



There was a big buzz in the summer of 2017 when more than one hundred leading artificial intelligence pioneers, including the ubiquitous Elon Musk, sent an open letter to the UN urging them to prevent a “killer robot” arms race. The reasons for this letter are numerous and, reading it carefully between the lines does instil a bit of healthy fear. Firstly, a robot-soldier can fight non-stop without a break, unlike humans who have to fight in shifts due to fatigue: you could argue that a robot’s batteries also run out of energy (like Matteo’s) but they will certainly last far longer than the measly 12 human hours. This could bring a whole new level of qualitative and quantitative intensity to armed conflict, which could be fought at a scale greater than ever before and at timescales faster than we could ever imagine from the non-automated non-intelligent wars we have seen up to now. I find it interesting that their brief letter touches on the speed of these “lethal autonomous weapons”, warning that humans would be left with no time even to comprehend what is happening.

If we want to paint an even gloomier picture, we can add to the list of concerns the unpredictability of robot decisions: contrary to what many people might think, artificial intelligence is not one hundred percent predictable in its expert and advanced determinations. Each neural network includes evolutionary calculation mechanisms that make it predictable only on a probabilistic basis, exactly like a human brain. Basically, we cannot be sure of the fact that the robot will do this or that, and only this or that, because it will also “use its own head”. As long as the robot is only preparing dinner or making juice, we can let this go: but when we’re talking about RoboCop or Terminator 4.0, the risks skyrocket.

In modern warfare, today, remote-controlled drones and missiles are already in use. However, we are still at minimal levels of intelligence: the majority of the work is done by humans, the military remote-commanders (up, up all the way to the “remote-Commander-in-Chief”), while the parameters of flight and targeting – for precision strikes on specific targets – are evolved, yes, but absolutely in no way entrusted to the artificial autonomy of neural networks.

What is more, intelligent robot-soldiers by nature do not have other characteristics typical of humans, which make them even more dangerous: they do

not feel any physical pain or moral pain (not yet at least, pending an improbable, though still possible, neural evolution). This means that they lack, in principle, the ability to feel the basic, innate emotion of fear. We humans, all of us, enter this world as small, defenceless, vulnerable and sentient beings, capable of fear – and we have to build our courage up, step by step, day by day (as this is not a primary element). Instead, robots come into the world with their courage, weapons and systems already formed – but without the gift of fear. The lack of physical and moral pain, and therefore fear, makes an intelligent robot extremely stupid emotionally. Emotional stupidity translates into: no fear of getting hurt by an enemy strike, almost zero self-preservation instinct and, above all, a complete lack of conscience.

How many tragedies have been made less tragic, how many wars (or private quarrels) have come to an end sooner, for “reasons of conscience”? How many lives have been saved by conscientious objection? It is impossible to know, but something tells me – also listening to Second World War stories told by grandparents and great-grandparents or reading historical accounts of wartime events – that the human conscience has, on more than one occasion, been a providential ingredient in preventing even worse destruction and brutality. What we are used to describing as “reasonableness” actually has very little to do with reason and much more with sentiment. I say that as a lawyer, and it is no secret that we lawyers know all about conflict. With robots, we can forget about all that. Or at the very least we can only expect scant emotional sensitivity. E-peace will prove difficult.

In those last few lines, though, I appear to have contradicted myself. I realise that. Throughout the first chapters of this short book, I imagined (and feared) the advent of objects equipped with artificial intelligence so advanced as to make them prone to whims and capable of getting offended or falling in love – whereas here I am arguing that a robot cannot feel compassion, fear or other sentiments. Let me make myself clearer, then: I believe that we will see the arrival of robots capable of feeling emotions but I’m equally convinced that those emotions will be inhuman

and as such, consequently, less humanitarian. Less attentive to the needs, desires, hopes and feelings of men and women, real flesh and blood (and soul).

It is true that, throughout history, human beings have committed atrocities, without being artificial. Nevertheless, I can't help but feel that even the worst, most heinous war crimes were not so much the result of feelings of rage, jealousy, envy, greed or other negative emotions that are innate in every individual, but rather of hyper-rationalisations. After all, those phases of profiling I flagged earlier as being so dangerous for people's fundamental rights and freedoms – the perfect tools for carrying out every imaginable form of discrimination – what are they if not over-simplified and hyper-rationalised abstractions? What are they if not theoretical negations of the concrete and irrefutable complexity and uniqueness of the human race? The characteristic of Nazism, and of any form of extremist ideology, is thinking that abstract scientific and mathematical formulas can translate, in a linear fashion, into rules for society and governance. Personally, I am convinced that a certain way of interpreting religious beliefs, which transforms theology into theocracy or “holy war”, depending on the circumstances, is another excellent example of hyper-rationalisation: not irrationality, but hyper-rationality.

Will there be robots that believe in God? Will they wage holy wars? Or will the robots themselves become divinities? I'm not just having fun spouting delirious nonsense here: planting my feet firmly back on the ground, my point is: we human beings, and first and foremost our governments and artificial intelligence tech experts and companies (the same ones that appealed to the UN), must be careful not to dig ourselves into a hole by haplessly creating “artificial Gods”. Profile of a robot divinity: be stronger, more knowledgeable and intelligent, last longer and be more resilient and less vulnerable than humans, and be objectively capable of governing from the height of its power. Good or bad character traits are not essential.

Let's go back to Matteo. That evening up in the hills, I found it so frustrating and disturbing that there was no ON/OFF button. I wanted to switch it off, but I couldn't. There was no way to intervene and stop the artificial crying, simply because the designers and the manufacturer had not envisaged that function. This made me think about the fact that we humans should never give up our *super-admin*

function – to use tech-jargon – in our relationships with intelligent robots and algorithms. The administrator of a system is the person (or object, since it could also be inhuman) that can decide the base settings, grant or take away powers from others, and basically make life or death decisions about that system. Clearly there can be a range of administrators with differing privileges and powers, but at the top there always has to be, somewhere, a super-admin capable of tracing the code back to its source, activating or deactivating and switching on or off the entire machine. The super-admin has divine power over the whole system.

A few chapters back, I quoted the formula “rule of law” States rather too offhandedly, without taking the time to explain the concept for the benefit of non-lawyers. I will do so now, trying to keep it as simple and easy to follow as possible: we use that formula to mean that no human being – emperor, king, head of state or government – is above the law. If we turn this the other way around, it also means that all citizens in a free and democratic nation, including governors and even kings and emperors, are accountable to the law. The law is above the king. The king has to respect the law. This principle has helped many countries, over the course of history, to overcome absolute monarchies, tyrannies and dictatorships, those totalitarian regimes in which the leaders are above the rules and can bully and “lord it over” everyone.

And here a horrible (and justified) doubt creeps in: across most of the western world, we struggled to free ourselves from the Sun King and from various dictators who were “just” human beings. Now what? Are we now going to create, with our very own hands, robotic Sun Kings and tyrants? Today, that democracy-defending formula would need to be expanded upon and better specified: “rule of human law”. We should in no way accept the idea of subjecting ourselves to rules, regulations, laws, decisions and codes that are automated and artificially created. No public law should ever be generated from an inhuman algorithm. No robot and no other form of artificial intelligence should be designed without an ON/OFF button that can be controlled only by humans and not by other robots – meaning that for each robot or form of artificial intelligence there should be at least one human super-admin and definitely no artificial super-admin. Also the robots, like

the kings and other governors, have to be held accountable to human law. And each super-admin, or remote-Commander-in-Chief, in turn, must also be subject to the rule of human law.

The perfection of an automated abstract calculation is closer to insanity than the imperfection of human feelings and rules, with all their inevitable qualms and eccentricities. Highly intelligent artificial insanity is the new frontier of risk, in peacetime as in wartime. Let’s get ready to fight it, armed with just our bare hands and our keyboards.

-----

### **Complete Index of the book**

1. Grappling (Barehanded) with The Artificial
2. Keeping a Safe e-Distance
3. Artificial Intelligence and Artificial Insanity: catch me if you can
4. Digital Profiling, Freedom and Circumvention
5. The Digital Subconscious and the Snack Market
6. Physics and Metaphysics of Data Protection: “*Crowdprivacy*” and “*3D privacy*”
7. Data as money: liberalism and control of data
8. Me-Reporters, zigzagging between privacy, information and expression
9. E-War and E-Peace: towards the “rule of human law”

**The Author - Luca Bolognini** (born in 1979) is a European privacy and data protection expert. Lawyer and President of the Italian Institute for Privacy and Data Valorisation, founding partner of the international law firm ICT Legal Consulting - ICTLC, author of many articles for national and international reviews, magazines and newspapers, he has published books and essays with *Giuffrè*, *RCS Etas*, *Rubbettino*, *Corriere della Sera* and *Springer*. He teaches data law in master courses and universities in Italy and overseas, and he serves as an ethics and privacy advisor for several EU Horizon 2020 research projects on e-privacy, smart cities, Big Data and the Internet of Things. [www.lucabolognini.it](http://www.lucabolognini.it)