

Editorial

Last month, I commented on a paper in a PLSC-style session (ie everyone that wants to attend should have read the paper in advance). To the superficial eye, the session was a failure: practically no one came. Most conference participants preferred other parallel sessions or the comforts of the lounge near the coffee machine. It was me, the two authors of the paper and two friends of the authors that filled the big room. But it turned out to be the best session since we could remember. People had actually read the paper, there was enough time for everyone to speak and listen to each other and it allowed us to steer away from questions from the audience typical to contemporary academic seminars (Does the speaker agree with my recently published paper in which I argue...?). Perhaps it was academia in its ideal form.

The paper concerned the use of Artificial Intelligence (AI) by insurance companies and took as a basic premise that modern data-technologies have unleashed a shift from causation to correlation. Already suggested during the early 2000s, this is by now something of a truism and has become even more evident with the rise of Big Data and AI. The thought is simple: what is unique to modern data processing technologies is their ability to process large quantities of data and distill patterns. These patterns uncover general correlations, some of them useful, others not. Algorithms have found nonsensical correlations, such as between the number of hurricanes in India and the amount of peanut butter eaten in the United States, but also many that have in fact been useful, such as a correlation between the colour of a person's couch and their lifestyle. There is obviously no causation – the colour of the couch does not have a negative effect on health, but there might be a correlation (eg the people with a bad lifestyle opt disproportionately often for a white leather couch).

Although I've used this divide myself in abundance, for example when working on big data, I've increasingly grown sceptical.

The most controversial reason is the one I suggested during the session. Having a philosophical background, I sometimes blurt out things that are a philosophical given, but that sound wildly imaginative to legal scholars. Many philosophers do not believe that causation exists, but find it to be a human construct. Empirically, for example, you cannot measure or prove causation. In a game of pool, you can see one ball hitting the other, and also that the other ball subsequently starts moving. But that does not prove that it is because the first ball hits the second one that the second ball moves. It is two independent phenomena that we witness, between which we ourselves provide a link. Suppose we run an experiment, say a million times. Each time ball A hits ball B, ball B starts moving. We can call it causation, but empirically we have witnessed a

correlation of a 100% with $n =$ a million.¹ Also from a theoretic perspective, it is unsure what causation actually suggests. Supposedly, Descartes's famous *cogito ergo sum* was first formulated as *cogito, sum*. The question is whether it is possible to think of a thing that thinks but does not exist? If that is possible, the fact that you are cannot be logically deduced from the fact that you think. If it is not possible, the question is whether the statement contains information at all; it might be like saying that the fact that there is a yellow ball implies that there is a ball (ie does 'yellow' have its own existence or is it logically dependent on being an attribute to a physical object?).

Quantum physics may be relevant for this discussion as well because it stresses the relativity of space and time. This has led to speculations over the existence of black holes and the possibility of time travel. If these phenomena cannot be logically excluded, and some experts suggest that they can't, that would change our understanding of cause and effect radically, because it would be possible for the effect to affect its cause, leading to a logical paradox. More down to earth, what is verified time and again is that particles move differently when they are observed, leading some to conclude that there is a relationship between the observer (subject) and the observed (object), as if the object 'knows' that it is observed, and as if subjects cause static objects to behave differently.

In psychology, experiments have been run with remarkable results. For example, in a setting where participants were shown pictures to measure their stress response (eg flowers, car crash, candy, blood), participants showed stress response when seeing the shocking pictures even before they were shown. Tests were run again to exclude the effect that participants can derive from the researcher's response which picture they will be shown, a quantum device was used to create a fully random order of pictures and tests were run to see whether before seeing 'happy' pictures, participants also showed a stress response. But even when excluding all potential explanations, a 'presentiment' was measured that could not be logically explained. These experiments seem to suggest that the effect comes before the cause.

All this ties into the larger philosophical debate between mind and matter, for which there are roughly three positions.² The first and most commonly popular one is that both mind and matter have an independent existence, but it has never been satisfactorily explained whether these two entities interact and if so, how and what is cause and what is effect. That is why, as a second position, some have suggested that everything is matter; our feelings and thoughts can simply be explained by activities of neurons and particles in our brain. People having undergone brain surgery sometimes change music taste or even personality radically overnight.³ The third position, to the other extreme, suggests that everything is mind, either leading to solipsism (everything

1 See further eg: <<https://plato.stanford.edu/entries/kant-hume-causality/>>.

2 <<https://plato.stanford.edu/entries/dualism/>>.

3 <<https://plato.stanford.edu/entries/materialism-eliminative/>>.

what an individual holds to be objective is produced by their individual brain and is thus subjective)⁴ or to pan-psychism (everyone is connected to a larger psyche, from which everything that is sprouts).⁵

Also, we all know how intuitively difficult it is to assign causal relationships, if only because there is often a chain of events leading up to a certain effect. If you ask your spouse to check the mail before they go, and later they get into a car crash, we all know the feeling of: if I only I hadn't asked, then they would not have been there. To attribute accurately and exhaustively what caused a person to be then and there where the crash happened, you would need to describe life, not only of that of you and your spouse, but also that of the person that caused the accident, of the other persons on the road, etc. The other way around, one tiny element can set a series of changes in motion. This is commonly known as the butterfly effect: the wings of one butterfly at one part of the world can, theoretically at least, trigger a chain of effects that causes a hurricane at the other end of the world.⁶ This feeds into what is known as the chaos theory.

Statistical correlations at their turn are not fully antithetical to causality. Many times, of course, causal relations and correlations align. At other times, the correlations do not show direct causal relationships, but indirect ones. There may be no direct causal relationship between the prevalence of sunburns and the number of ice creams sold, but both have an indirect and partial causal relationship to the abundance of sunlight. Still other correlations that make no sense to us (yet) may in fact uncover a complex chain of cause and effect, eg the nonsensical correlation between the number of hurricanes in India and the rise in peanut butter consumption in the States may in fact be explained through the butterfly effect, or by something so down to earth as that hurricanes in India jam normal shipments of rice from South-East Asia to the States, leading to products produced domestically (eg peanut production) having a lower relative price, spiking their domestic consumption.

In addition, the stark contrast between the past (causation) and the present (correlation) is theoretical in nature only. Obviously, organisations have always used correlations. States have, from early on, held census to estimate how many people lived in certain areas as well as its population's age, gender, religious and cultural background and profession. Insurance companies, banks and other financial institutions have always relied on statistical correlations for setting prices, for example estimating the likelihood a potential customer would use an insurance product and if so, what would be the costs involved. The opportunities offered by data-driven technologies, Big Data and AI have not unleashed a radically new phenomenon, but only present more potent means to effectuate existing practices. Similarly, although predictive policing sys-

4 <<https://plato.stanford.edu/entries/other-minds/>>.

5 <<https://plato.stanford.edu/entries/panpsychism/>>.

6 Edward Lorenz, 'The butterfly effect' (World Scientific Series on Nonlinear Science Series A, 2000) 39, 91-94.

tems aim at foreseeing where and when crime will take place based on historical data, police agents have always done the same, though on a smaller (personal) scale.

In the legal domain, it is clear how difficult it is to establish causality. For example, there is a complex web of theories that has been developed for determining causality. The *conditio sine qua non* theorists focus on the condition without which the effect would not have materialised, but this may be neither be the most important cause nor the most proximate. The *causa proxima* method suggests looking at the cause most directly related to the effect (either based on *on spatio* or on temporal proximity), while the *causa remota* model locates the determinative factor back to the cause that set the chain of events in motion. What if a person causes a non-lethal accident, but in the treatment of the person injured, the doctors make a mistake, leading to the patient's death? The theory of *causa causae est caussa causati* would find that the cause that caused the cause that caused the person to die is where the burden should be placed, thus differing from the *causa proxima* paradigm. The notion of *culpa in causa* suggests that if a person is to blame for a situation in which he caused something for which they might otherwise be excused, such as a person who commits hate speech when they were highly intoxicated, can be held accountable for putting themselves in a state in which they knew or should have known that they might exhibit behaviour they could not account for. And so, the list of theories concerning culpability goes on.

But at least, one could argue, the legal domain has always exclusively focused on causality instead of correlation. Neither that is fully true. For example, laws often work with presumed causality, which is in fact derived from correlation, to circumvent complicated discussions and evidentiary complications. In addition, it is clear how important statistical data have become with the rise of DNA-evidence (with judges often-times making interpretative errors). Another area where correlation is increasingly becoming important is that of environmental law, where it is often difficult to prove or substantiate that pollution caused harmful effects on an individual claimant. Did a person get cancer because of the smog of the nearby factory, and if so, what percentage was attributable to this cause (eg if a person also is a smoker)? Generally, it can be proven that on a population level, people that live in the vicinity of a polluting factory have a higher risk of getting a certain disease, but that does not exclude the possibility that in an individual case, the cause of a person living in the vicinity of a polluting factor getting that disease was not the smog, but their unhealthy lifestyle choices or hereditary predispositions. That is why courts increasingly rely on statistical evidence to determine presumed causality and culpability.

Another debate relates to the use of biological factors in law. For example, there is a correlation between gender (male) and crime, which law enforcement authorities sometimes use, but should they be allowed to? With the statistical correlation between ethnicity and crime, that would be absolutely prohibited, but with gender, there seems to be more lenience. Is that because implicitly, we believe there is an indirect causality between male gender and testosterone levels and between testosterone levels and proneness to crime? However, though there is a correlation between gender and testos-

terone levels, there are men that through a genetic defect have low testosterone levels and women that have high levels.⁷ Some people have consequently problematised the use in law of gender altogether, as these are legal constructs based on correlations (eg between reproduction organs and testosterone and estrogen levels), while reality shows sliding scales rather than binary distinctions.

A final legal debate that is worth noting builds on the question over determinism. Though law typically works with causality, attributability and culpability, many have doubted the extent to which free will exist or, less extreme, to what extent other factors than our free will factor in when making decisions. For example, there are people that believe that our behaviour is to a large extent determined by our DNA and genetic makeup, others propagate a form of cultural determinism, and still others believe in a combination of both deterministic effects. Psychologists will tell you that the majority of the decisions we make are informed by subconscious processes; we make decisions and later rationalise them, coming up post-hoc reasons for our conduct. What all these theories have in common is that they suggest that the causal relationship between an individual's decision and their action (and thus their responsibility for it) is a legal construct rather than reality.

All in all, the supposed causality-correlation divide may be infinitely more complex than sometimes suggested. Also when causality is used in law, this oftentimes comes down to an attribution of responsibility, a presumption or correlation that is taken to imply causation.

Let me now proudly turn to this edition. Proud, because we have two opinions on the future of data protection law by the newly elected Chair of the European Data Protection Board Anu Talus, and by Maja Brkan, a Court of Justice judge and long serving member on the Board of EDPL. This theme is also reflected in the articles section. To celebrate the five-year anniversary of the European Union's General Data Protection Regulation, we have invited authors not to look back, but ahead. In five- or ten-years' time from now, what would the data landscape look like, what new challenges will have emerged, what questions will need to be tackled? We are deeply honoured to feature articles by Paul M. Schwartz and Anupam Chander, Alexander Dix, Dara Hallinan and Yannick Alexander Vogel. As you can imagine, their reflections are highly interesting and valuable.

As always, the reports section, led by Mark Cole and Christina Etteldorf, is packed with interesting reflections as well. Hugo Lami points us to the French DPA's decision on cookies, Sven Braun discusses the German Data Retention Law that got stricken down, Giorgia Bincoletto evaluates the Italian DPA's standpoint on information obligations in the health domain and Mark Cole and Katharina Kollmann discuss the Norwegian DPA's use of the urgency procedure with respect to personalised advertisements. In the

7 See eg: <<https://www.theguardian.com/sport/2023/jul/11/caster-semenya-discriminated-against-by-testosterone-levels-rules-echr>>.

case notes section, led by Maria Tzanou, two CJEU cases are discussed. Ana-Maria Hriscu reflects on the CJEU's interpretation of the relation between data protection and competition law in a highly debated matter. Shweta Reddy Degalahal discusses a case that has not gotten the attention it perhaps deserves, as the CJEU has provided some very important cues on the notion of identifiability. In the book review section, Gloria González Fuster picks up on that same theme when evaluating the book authored by Suzanne Vergnolle.

Finally, two Board members have decided to step down to focus on new areas of expertise. Marc Rotenberg was and is one of the key figures in the privacy and data protection field and has as such enriched EDPL with his insights and writings from the very start. As the director of the Center for AI and Digital Policy, he will move his focus to the area of Artificial Intelligence. Alessandro Spina equally has been of great value for EDPL, bringing in expertise from the health domain and from the EU's inner workings. I want to thank both Alessandro and Marc for their invaluable contribution to the journal. That counts as well for Jakob McKernan, who served as the executive editor of EDPL. Jakob was the spider in the web, managing about everything that matters, from laying contact with authors and managing the website to organising EDPL panels, proof reading and getting the journal to print in time. I will miss Jakob as much as I will Alessandro and Marc.

At the same time, I'm happy and proud that Nelly Stratieva will return to the position of executive editor, and given her experience as the first and founding executive editor of EDPL, she will have no trouble filling Jakob's shoes. The two vacant positions on the Editorial Board have been filled by two professors whom I have admired for years. Bilyana Petkova and Susanna Lindroos-Hovinheimo. Bilyana is a professor at Middlesex University and is the author of the much cited 'Privacy as Europe's first Amendment', among many other articles and books. Susanna is professor in Helsinki, a regular contributor to EDPL and, among many articles and books, the author of 'Private Selves: Legal Personhood in European Privacy Protection'.

For those interested in submitting an article, report, case note or book review, please e-mail our Executive Editor Nelly Stratieva (<stratieva@lexxion.eu>) and keep in mind the following deadlines:

- Issue 1/2024: 15 January 2024;
- Issue 2/2024: 30 April 2024;
- Issue 3/2024: 15 July 2024.

Bart van der Sloot
Tilburg Institute for Law, Technology, and Society (TILT)
Tilburg University, Netherlands