The cybernetic future of competition law: information exchange, entropy and artificial intelligence

David Foster

Frontier Economics, London

Stafford Beer was one of those classic British eccentrics. A long professorial beard, a monocle, a habitual cigar smoker who would sip whisky from a hip flask while debating complex theory late into the night. His obituary in 2002 described him as 'a cross between Orson Welles and Socrates'. He was a pathfinder in the field of cybernetics: the study of information and control in complex systems. Beer applied those ideas to a particular sort of complex system: firms and the markets they operate in.

One concept sits centrally at the core of cybernetics, or what we might now call systems theory, and that concept is *information*. What I'll argue in this article is that we are going to need these ideas to interpret and apply competition law, specifically Article 101 of the Treaty on the Functioning of the European Union (TFEU), as artificial intelligence (AI) advances and increasingly begins to organise economic activity. I'm building here on the work of Nicholas Petit and Thibault Schrepel, who have both recognised the importance of information and complexity theory in antitrust analysis.[1]

This vision of the future is not yet here, but it may be a lot closer than we think.

1 Nicolas Petit and Thibault Schrepel, 'Complexity-Minded Antitrust' (2023) 30 Journal of Evolutionary Economics 541.

An agreement is legal fiction, all that really matters is information

Competition law prohibits firms from entering into anti-competitive agreements. But what is an agreement? Imagine Bob meets Alice at an industry conference and they talk about aligning their future pricing strategies. That is definitely an infringement. But have they really reached an agreement? In any meaningful economic sense, they have not. There is no enforceable contract, no structural link between their firms.

What has actually happened is a simple exchange of information. Economists sometimes refer to this as 'cheap talk'. Bob and Alice may have communicated their intentions, but there is nothing to stop them from leaving the room and doing the exact opposite.

From an economic point of view, there is therefore no qualitative difference between a 'hardcore' agreement on the one hand and 'information exchange' on the other. The difference is one of degree. The information Bob and Alice have exchanged is highly targeted, clear and effective in establishing coordinated behaviour. Calling this an 'agreement' is, in that sense, a form of *legal fiction*.

This legal fiction plays a much greater role in the United States, where the law requires the establishment of an actual agreement between parties. In the EU, however, the concept of a *concerted practice* aligns much more closely with the underlying economics. It captures forms of coordination that rely on wider forms of information exchange, and don't depend on an explicit agreement. Ultimately, all agreements, whether formal or informal, are forms of information exchange. They differ only in their information content. We can therefore think of Article 101 as defining the set of information exchanges between economic entities that are prohibited; those that create an impermissible degree of *mutual information* between independent firms.

Information is entropy

In a future where human interaction plays less of a role in economic life, we will not be able to rely on a lawyer's human intuition to interpret what counts as a prohibited exchange of information. The currency of information exchange between machines is not words, but bits: 1s and 0s.

Fortunately, we already have a rigorous definition of what information actually is, and this definition turns out to be fundamental to our understanding of everything from the laws of thermodynamics to modern telecoms.

Information, it turns out, is *entropy*.

Entropy is the measure of disorder in a system. It is the concept that underpins the Second Law of Thermodynamics: that disorder in the universe always increases.

Entropy also defines the arrow of time because of its irreversibility: once an egg has been scrambled, it cannot be put back together. In the past, there was an egg; in the future there will only ever be breakfast. If you think about it, information is also like this. Revealing information is also an irreversible process: you cannot simply 'undo' information exchange.

The connection between information and entropy was discovered independently by two people during the 1940s: Norbert Wiener, the godfather of cybernetics and inspiration for Beer, and Claude Shannon, a mathematician at Bell Labs. Both were involved in the Second World War effort through their work with the National Defense Research Committee in the US. It was by working on problems like anti-aircraft missile targeting and cryptography that they realised that the same mathematical principles could describe how information is transmitted and how systems maintain order. In the end, it was mostly Shannon who took the credit. His 1948 paper, 'A Mathematical Theory of Communication', was described by *Scientific American* as 'the Magna Carta of the Information Age'. Together, he and Alan Turing, with whom Shannon also collaborated during the war, provided the theoretical foundations of our modern digital world.

A simple way to understand why information is entropy is to think of information as a *surprise*. If an event is 100 per cent certain, it conveys no information when it happens. Conversely, if something unexpected occurs, it teaches us a lot about the state of the world. When the world is fully predictable, entropy is zero. When it is totally random, entropy is maximal.

This insight also helps us to see that information exchange is not about the volume of data sent but about the *reduction of uncertainty*. A single signal can transmit vast information if it resolves large uncertainty. This matters deeply for competition law: a single exchange of a surprising, targeted message between rivals might have far more anti-competitive impact than a thousand pages of routine disclosures.

For the purposes of what I am arguing in this article, the important thing is that entropy, and therefore mutual information between two entities, can be precisely defined in mathematical terms. Just for fun (and you don't need to know, understand or remember this), here is the formal definition of mutual information between two entities X and Y:

$$I(X;Y) = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}.$$

Returning to Article 101, recall that we can think of it as prohibiting situations in which mutual information *I(X;Y)* between firms *X* and *Y* is excessively high. It is rather useful that this can, at least, in principle, be measured mathematically, something that may prove essential when the flow of data between two AI-driven systems becomes the only observable evidence of coordination.

Black boxes and POSIWID

This brings us back to Beer, and how he extended Shannon and Wiener's wartime ideas into a framework for understanding firm behaviour.

Cybernetics can sound like an outdated science fiction term, something from the era of tin foil robots and Cold War laboratories. And in a sense, it was. The field became unfashionable partly because of its association with Soviet planning in the 1960s. Understanding how information can be used to control complex systems is, after all, rather useful if you want to run a large command economy and a totalitarian state.

Today, the same ideas have been rebranded as *systems theory* or *complexity science*, and they are once again gaining relevance. They remain the right toolkit whenever we need to understand the behaviour of a *black box*.

When we refer to a black box, we mean a system that cannot be understood by opening it up and inspecting its parts. The human mind is one example. If you want to understand a person's behaviour, dissecting their brain will not help. In other words, a black box is a complex system whose behaviour amounts to more than the sum of its parts.

Current pricing algorithms may be complicated, but most are not yet truly complex. They are still deterministic programs that can be understood by looking at their code. In such systems, you can still answer the question 'what has this algorithm been designed to do?' The ability to do that will disappear as AI systems become more autonomous and self-learning. Soon it will no longer make sense to ask what the system was 'designed' to do because it will not have been designed by anyone in the traditional sense. It will simply be a black box.

This has important implications for competition law. The popular idea of *compliance by design*, once advocated by former Vice-President Vestager, is appealing but flawed.[2] You cannot 'design in' compliance when the system itself is not designed by anyone. Asking what the machine 'was trying to do' will make as little sense as asking what the weather was trying to do.

Beer's answer to this problem of complexity was a principle he called *POSIWID*: the *purpose of a system is what it does*. The only way to understand a black box is by

2 See https://ec.europa.eu/newsroom/comp/items/55994/en accessed 17 November 2025.

observing its inputs and outputs: the information it receives and the information it emits. In information-theoretic terms, the purpose of a system can be understood through mutual information between the system and its environment. What information theory tells us is that, in fact, there is actually nothing more to know. All information – all entropy – is stored at the boundary of a system, where its surface meets the outside world.[3] By observing these information flows, we can identify the *emergent properties* of the system: the dominant patterns of behaviour that appear when it interacts with the world.

So, in short, you can observe what information signals flow into the AI system and you can observe what information signals flow out of the AI system, and by doing that, you can identify what that AI system is really up to.

This is not so alien. It is exactly how we understand human behaviour. You have never seen inside Bob's mind and no amount of dawn raids or cross-examination will reveal his inner state. All the information that exists about Bob is found at his surface: how he acts in response to what he perceives.

This guy called Bob

This connection was captured beautifully by former acting Chair of the US Federal Trade Commission, Maureen Ohlhausen, when she proposed what has become known as the 'Bob Test':[4] 'Everywhere the word "algorithm" appears, please just insert the words "a guy named Bob". Is it ok for a guy named Bob to collect confidential price strategy information from all the participants in a market, and then tell everybody how they should price? If it isn't ok for Bob to do it, then it probably isn't ok for an algorithm to do it either.'

Whether deliberate or not, the choice of 'Bob' is apt: in information theory, 'Bob' and 'Alice' are the canonical parties used to illustrate communication, for example, when talking about how to send encrypted messages between two parties.

One interpretation of the Bob Test is that we do not need to change anything; that at some fundamental level, competition law remains fully capable of handling the algorithmic future. There is truth in that. In the end, our understanding of an AI system will rest on the same principle as our understanding of Bob: POSIWID.

But the difficult question will not be 'should Bob have done this?' but 'what did Bob actually do?'

---

3 As an aside, this principle, known as the 'holographic principle', is so fundamental to the universe that it is also extremely important to our understanding of black holes, and this may be the key to solving the ongoing mystery of how to reconcile quantum mechanics with Einstein's theory of general relativity. See, eg, Vlako Vedral, *Decoding Reality* (OUP, 2010).

4 See www.wsgr.com/a/web/9s4qwEHMgYtoz9aVZPcAQg/what-about-bob.pdf accessed 17 November 2025.

We have spent millennia developing an intuitive model of human behaviour. We know, roughly, what humans mean when they act. Our legal system is built on our ability to interpret their actions. In formal terms, we and our legal system usually have an instinctive grasp of the *mutual information* created between the 'Bob system' and the 'Alice system' when they interact.

That understanding will vanish when AI takes over the operation of markets. We will observe the Bob-a-tron 3000 and have no intuitive sense of what it is doing or why. The challenge will not be applying the Bob Test, it will be knowing how to frame it in the first place.

Mutual information and the future of competition law

This is why more rigorous ways of defining and measuring mutual information will ultimately become necessary. And they will imply big shifts in the way competition authorities go about enforcing the law.

For these purposes, I want us to assume that we have already entered a future where AI has brought about systemic change in how economic activity is organised. That is, we are no longer talking about replacing what humans do with some AI equivalent, but rather that AI has transformed how things work in such a way that they no longer resemble how a human would have done it.

Let me give a simple example of what I mean. Take, for instance, the way digital advertising works today. A merchant has a product and lists it for sale on the company website at a certain price. The merchant buys digital advertising that might list the product, alongside rival products, for customers who are searching for that type of product, or display adverts to customers who are likely to be interested.

A first-order impact of AI, what Ajay Agrawal, Joshua Gans and Avi Goldfarb refer to as 'substitution',[5] might be that a customer instead interacts with an AI tool that helps it crawl through the websites of available products and make a decision about which one is best. That technology already exists today.

But once AI technology causes systemic transformation, the world will not look like this at all. There will be no website; why would that be needed if no one ever looks at it? One version of that future might be that the buyer's AI simply exchanges information with seller AIs, negotiating prices in real time across multiple competing merchants. Information, in the form of 1s and 0s, will flow multilaterally, in vast amounts, and in near-instantaneous unison across all market participants; a hub with billions of spokes, spinning millions of times a second.

---

5 Ajay Agrawal, Joshua Gans and Avi Goldfarb, *Power and Prediction: The Disruptive Economics of Artificial Intelligence* (Harvard Business Review Press, 2022).

This potential future raises another interesting question: what does it mean to 'exchange' information in this world? The concept of exchange has meaning between human communicators, and while it might be ambiguous sometimes, we are nonetheless pretty clear what we mean when we say it. Bob has a message in his mind, he communicates through some medium (eg, his words), and the message is heard and understood by Alice, who communicates a response through her chosen medium. We can use our mental model of the human being to construct for ourselves what constitutes a response, as opposed to other non-response behaviours. That won't do when the responder isn't human. We need to define what a response is for a machine. This is where we have to turn to the concept of mutual information.

A response in this context is something that creates mutual information between two systems. The purpose of Alice responding is so that Bob knows something about Alice in such a way that Bob's future states will now correlate to Alice's future states. This is actually a lot like what is known as 'entanglement' in the field of quantum physics, and is incidentally the mechanism by which quantum computers, such as Google's Willow, work their magic. Data flows that create mutual information lead to a situation where the Bob system and the Alice system become entangled with each other. This state of entanglement is something that can readily be measured through statistics, provided there is enough of the right data, of course. The methodologies to do this are perfectly well known.

So who ya gonna call?

Ghostbusters, obviously. And in a sense, that is precisely what we will be searching for: what philosophers have called the 'ghost in the machine';[6] that is, the emergent behaviours of AI that mimic human behaviour and give it apparent intent, such that we can meaningfully say that the AI has broken competition law in just the same way that Bob did.

To make this at all possible in the future, it stands to reason that we will need four things:

1. to be able to *process vast data* on very large numbers of multilateral machine-to-machine communications;
2. to understand the extent to which those interactions *create mutual information* between those machines, which will rely on statistical analysis;
3. to build new technology for analysis that will rely on the concepts of entropy and *Shannon information*; and
4. to define new *limiting principles* for when interacting systems have created too much mutual information, crossing the line from legal to illegal.

6 Gilbert Ryle, *The Concept of Mind* (University of Chicago Press, 1949).

I want to finish by saying something more about what it might mean to define those limiting principles for competition law once we are dealing only with the statistical analysis of machine-to-machine interactions.

Mutual information is really about your ability to predict the behaviour of one system once you know something about another. But unlike conventional econometric analysis, which seeks causal relationships between one system and another, information theory has us looking for surprises, for example, signals from Bob to Alice that are statistically very unlikely to occur, but reveal, with high probability, the likely future state of Bob. A series of such signals, flowing in both directions, could then be shown analytically to place the Bob and Alice systems in a position to predict each other's future states with a high degree of probability.

Having already said that compliance by design is a flawed concept, you might ask whether I am saying that nothing can be done by the owners of these systems to remain compliant once some 'standard' has been set for what amount of mutual information is too much to be legal. I would say that the kind of compliance that emerges in this world is actually 'compliance by compliance' rather than 'compliance by design', although of course, I will admit, this type of compliance still has to be, by definition, designed by someone or something. Systems that are tasked with obeying competition law will be able to model the creation of mutual information in real time and adjust their data flows accordingly. This isn't the same thing as designing a set of compliant versus non-compliant behaviours up front, rather it is about monitoring, as a good in-house competition lawyer might, live behaviour and adapting to the world as it evolves.

Conclusion

Do all of this and we will once again be in a position to apply the Bob Test. None of this will really be new, as such. It is, in fact, exactly what we do today when we analyse human behaviour. The difference is that, today, we do all this information processing unknowingly, through instinct and established societal norms. In other words, today, lawyers can do it on their own. Tomorrow, they will need data scientists.