Vrije Universiteit Brussel

VRIJE
VUB UNIVERSITEIT
BRUSSEL

The Artificial Intelligence of European Union Law

Hildebrandt, Mireille

The Artificial Intelligence of European Union Law

Mireille Hildebrandt*

A. Introduction

In this Article, I take my chance to briefly introduce the key ideas of two German philosophers whose work is highly relevant for the rule of law in the age of machine intelligence. The current predominance of Anglo-American moral and legal philosophy, with its emphasis on either utilitarian or a specific type of neo-Kantian moral philosophy calls for some countervailing thinking, and the *German Law Journal* seems the right place to dare such a thing. The recent launch of an English translation of biologist and philosopher Helmuth Plessner's seminal *Levels of Organic Life and the Human* (1928)[1] invites a fundamental reflection on the difference between human and machine intelligence, including a penetrating criticism of the kind of behaviorism that underpins personalized micro targeting.[2] The core findings of Plessner, based on what he calls the ex-centric positionality of human beings, connect well with key insights of lawyer and legal philosopher Gustav Radbruch, taken from his *Legal Philosophy* (1932),[3] notably the idea that law is defined by antinomian goals.

AI usually stands for artificial intelligence, referring to a rather vague notion upon which no agreement exists, neither amongst experts nor amongst those affected by its supposedly disruptive character. Therefore, AI is better understood as referring to automated inferences and is better described as machine intelligence. Based on Plessner, I will argue that current machine intelligence is radically different from human intelligence. My point will be that it is precisely human intelligence that is deeply artificial, whereas machine intelligence is merely automated. This relates to the importance of recognizing, appreciating, and protecting the artificial nature of law and the specific intelligence it affords human society. Finally, I will argue that a proper understanding of the "mode of existence" of machinic agency will be one of the major challenges for the EU in the 2020s. If we get it right, we should be able to avoid the quest for certainty[4] that informs both informational capitalism[5] and state-centered surveillance.[6] Both are premised on mistaken visions

*Mireille Hildebrandt is Research Professor of 'Interfacing Law and Technology', appointed by the Research Council at Vrije Universiteit Brussels, at the research group of Law Science Technology & Society, Faculty of Law and Criminolog. She also has a Chair on Smart Environments, Data Protection & the Rule of Law at the institute of Computing and Information Sciences (iCIS), Science Faculty, Radboud University Nijmegen. In 2018 she was awarded an ERC Advanced Grant for the 5-year Research Project on 'Counting as a Human Being in the Era of Computational Law' (COHUCICOL), grant nr. ERC-2017-ADG No 788734. See www.cohubicol.com.

[1]Helmuth Plessner & J. M. Bernstein, Levels of Organic Life and the Human: An Introduction to Philosophical Anthropology (Millay Hyatt trans., 2019). *See* Gesa Lindemann, *Editorial*, 42 Hum. Stud. 1–12 (2019) (introducing the special issue on Plessner's work).

[2]F. J. J. Buytendijk & H. Plessner, *Die Physiologische Erklärung Des Verhaltens*, 1 Acta Biotheoretica 151–72 (1936).

[3]Gustav Radbruch, *Legal Philosphy*, in The Legal Philosophies of Lask, Radbruch and Dabin 107–12 (Kurt Wilk Trans., 2014). *See also* Mireille Hildebrandt, *Radbruch's Rechtsstaat and Schmitt's Legal Order: Legalism, Legality, and the Institution of Law*, 2 Critical Analysis of L. 42 (2015).

[4]*See* Stephen Toulmin, Cosmopolis: The Hidden Agenda of Modernity (1992).

[5]*See* Julie E. Cohen, Between Truth and Power: The Legal Constructions of Informational Capitalism (forthcoming 2019).

[6]*See* Richard H. Thaler & Cass R. Sunstein, Nudge: improving decisions about health, wealth, and happiness (2008).

of total control.[7] By avoiding the pitfalls of algorithmic overdetermination EU law should keep our future open in ways that empower us, instead of treating us as manipulatable pawns.

B. Human Intelligence is Deeply Artificial

In his *Levels*, philosophical anthropologist Plessner discusses three constitutional laws of the human condition. First, he demonstrates in what way our nature is deeply artificial. Second, he explains how our cognition depends on a mediated immediacy. Third, he shows that our ex-centric positionality generates a utopian point of view. The latter refers to an observer's position, which enables an external, third-person perspective that is nevertheless dependent on primordial first- and second-person perspectives. In essence, these constitutional laws describe the fact that we are never at one with our self. According to Plessner, we are incapable of a natural, unmediated access to the world we perceive, the outer world, or to the world we institute, the shared world, or even to the self we experience, the inner world. Our self is not identical with itself. The self is not a substance but a first-person perspective that depends on our ability to take a second-person perspective, anticipating how others will understand us, and on a third-person perspective that situates our embodied self in an objectified space.

In other words, Plessner highlights that our self is constituted by taking an ex-centric position.[8] Though some might think this is a bug, it is actually a feature. It is exactly the *incongruence of the self with the self* that generates productive misunderstandings and creative leaps. It is our ability to imagine and solidify a world of institutions that affords novelty while safeguarding legitimate mutual expectations that allow us sufficient certainty to act into the future.

At the same time, obviously, the *incongruence of the self with the self* that defines us creates an existential uncertainty that cannot be resolved. This uncertainty may cause fear, mental pain, and anxiety that may in turn invite false promises about a safe and simple life. This is the lure of populism, fake news and other attempts to ignore and deny the schism that grounds us. Our artificial nature is only a feature if we face the turbulence it implies and resist easy resolutions that would turn the feature into a bug. Facing the fragility as well as the resilience of the shared institutional world we create and sustain is core to human intelligence, requiring vigilance and care as well as autonomy and adaptation.

According to lawyer and legal philosopher Radbruch, in a constitutional democracy law is core to the sustainability of the shared world, because law integrates three antinomian values that define the goals of the law. First, legal certainty grounds the reasonable foreseeability of the consequences of one's actions. Without such foreseeability meaningful action is not possible. Second, justice mediates both distributive and proportional equality, requiring that a government treats similar cases equality to the extent of their similarity. Third, instrumentality contributes to achieving objectives determined by a democratic legislator.[9] Law operates at the nexus of the second-person perspective, which enables people to address each other based on legitimate, enforceable, and contestable norms, and a third-person perspective, which takes into account relevant, contestable facts. This way law protects the indeterminate first-person perspectives, both the "I" and the "we," that depend on the combination of foreseeability and contestability inherent in the human condition.

---

[7]*See* Frank Pasquale, *Tech Platforms and the Knowledge Problem*, 11 AM. AFF. J.   3–16 (2018), https://americanaffairsjournal.org/2018/05/tech-platforms-and-the-knowledge-problem/.

[8]Others have come to similar conclusions about the nature of self, mind, and society. *See e.g.*, GEORGE HERBERT MEAD & CHARLES WILLIAM MORRIS, MIND, SELF, AND SOCIETY FROM THE STANDPOINT OF A SOCIAL BEHAVIORIST (1962); PAUL RICOEUR, ONESELF AS ANOTHER (1992); RAF VANDERSTRAETEN, *Parsons, Luhmann and the Theorem of Double Contingency*, 2 J. Classical Soc. 77–92 (2007).

[9]*See* Radbruch, *supra* note 3.

C. The Machinic Nature of Machine Intelligence

In *Smart Technologies and the End(s) of Law*[10] I have argued that it is crucial to recognize and acknowledge the "agency" of smart technologies. In alignment with the concept of agency developed in the context of systems theory[11] and robotics,[12] I defined agency as the ability to perceive an environment in terms of actionability, coupled with the ability to act on the world. This highlights the relational nature of agency and the equiprimordiality of perception and action; agents perceive what their environment affords them as to the actions they can take.[13] This does not imply that all agents have intentions or an inner life. Agents need not be conscious to survive and flourish in their environment. In fact, this definition would include a thermostat, which perceives temperature in terms of the action it can take, as its perception of whether a certain threshold has been reached exists in function of the decision whether or not to actuate its heating or cooling capacities. It also includes plants, which have no central nervous system and do not engage in deliberation before adapting to their environment. Usually, to qualify as an agent, systems theory and robotics require more than action and perception, notably the ability to reconfigure the own system to achieve set goals, based on continuous feedback loops. Moreover, human agency includes the capability to set one's own goals, both as a "I" and as a "we". That is, as a person and as society.

There are two reasons for introducing a high-level concept of agency that includes non-human and even non-organic agents. First, it confronts us with the fact that we are currently surrounded by machinic agents that adapt their behavior to our behavioral data with the goal of thus modifying our behaviors. Their adaptation is based on a behaviorist, cybernetic understanding of human society, looping standard setting with monitoring and behavior modification.[14] By acknowledging that non-human and even inorganic entities may have some kind of agency, we can learn to anticipate and respond to the way these agents try to predict and preempt us. Ignoring their agency may leave us in the dark and thus manipulable. Second, this high-level concept of agency also confronts us with the fact that although non-human agents may be anticipating us and may learn how to affect us, they do not necessarily share our type of agency. This goes for plants and animals but also for machines.

Machines capable of automated inferences (AI) have a specific type of agency that can best be defined as data- and code-driven. They are data-driven since they can only perceive their environment in the form of data. Human beings perceive color, sound, contours, smells, tastes, touch, while our perception is always already mediated by language and interpretation.[15] AI machines can only perceive any of this as data. This implies an act of translation or an environment that consists of data, for example, an online environment, virtual reality, or an IoT environment. AI machines are code-driven because data does not speak for itself. To make inferences these machines require code, for instance based on machine learning research designs that seek to compress big data into a mathematical function, the so-called target function, that defines the data in view of a specific machine-readable task. This, in turn, requires developing a so-called hypothesis space that consists of potentially relevant mathematical functions. These functions serve as hypotheses for the so-called target function that supposedly underlies regularities in the outer, the inner, or the shared world. Because algorithms cannot be trained on future data, it is never

---

[10] Mireille Hildebrandt, Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology (2015).

[11] See H.R. Maturana & F.J. Varela, Autopoiesis and Cognition: The Realization of the Living (1991).

[12] See Luc Steels, *When Are Robots Intelligent Autonomous Agents?*, 15 Robotics & Autonomous Systems 3–9 (1995). *See also* Luciano Floridi & J.W. Sanders, *On the Morality of Artificial Agents*, 14 Minds & Machines 349–79 (2004).

[13] See F.J. Varela et al., The Embodied Mind: Cognitive Science and Human Experience (1991).

[14] On the cybernetic approach to influencing a population, see, for example, Mark Leiser & Andrew Murray, *The Role of Non-State Actors and Institutions in the Governance of New and Emerging Digital Technologies*, in The Oxford Handbook of Law, Regulation and Technology 671 (Roger Brownsword et al. eds., 2017).

[15] On the crucial role of language in the constitution of our eccentric nature, see, for example, Paul Ricoeur, *The Model of the Text: Meaningful Action Considered as a Text*, 5 New Literary Hist. 91–117 (1973). *Cf.* T Cheung, *The Language Monopoly: Plessner on Apes, Humans and Expressions*, 26 Language & Comm. 316–30 (2006).

certain that the target function is sufficiently approximated.[16] To measure the approximation, machine learning employs so-called objective functions that aim to either minimize error or to maximize likelihood. These objective functions thus optimize the solution to the mathematical problem of identifying the best possible approximation of the target function. In other words, machine intelligence is not based on *meaning* but on *mathematics*. It is not capable of taking a first- or second-person perspective. Instead, it is built on *our* ability to take a third-person perspective. The nature of machine learning is automation, not artificiality in Plessner's sense. Therefore, machinic agency has no ability to develop an ex-centric position as it is only ever executing code.

Human beings understand their environments in multiple ways, often by way of automated inferences, which we call intuitions. These inferences are not based on machine readable data but on the ex-centric positionality that invites us to "read" our own actions from the perspective of others. This enables us to foresee how others will respond to our actions, and what kind of effects our actions will have in the institutional environment we inhabit. This institutional environment—the shared world—thus shapes the action-potential of human agents, building on speech acts[17] that constitute and stabilize mutual expectations. These expectations hinge on the *meaning* of human discourse that in turn depends on the performative nature of speech acts, which can be spoken or written. For instance, when a civil servant declares two people husband and wife they are not *describing* a marriage but *instituting* it. The declaration does what it says. Human language use is performative in the sense that what counts as a marriage, but also as bread, a home or public transport, is shaped by the way these terms are used.

This highlights the normative, though not moral, character of language usage. As the norms that guide the use of language are in turn constituted by the usage of a language,[18] human discourse suffers and celebrates a fundamental uncertainty and concomitant openness. The fact that the meaning of terms is dependent on actual usage vouches for the open texture of both the vocabulary and the grammar of our languages and the institutional context they enable. This affords ambiguity, creativity, and a permanent reinvention of the shared world of institutional artefacts, such as, marriage, money, contracts, property, the state, or a university. For this reason, human intelligence is deeply entwined with the performative nature of human discourse and the artificial environment it creates, sustains, and transforms.

Machine intelligence is not based on the ex-centric position of human agents, though the creation of machine intelligence is. Machines cannot do anything but execute programs developed by humans, even if those programs enable the machine to reconfigure its program in view of specified machine-readable tasks, and even if humans may develop programs that build new programs. Though the latter will generate a potentially unforeseeable dynamic, the machine is still merely executing code, in response to the data it feeds on. Machine agency has no inner world, is not grounded in a shared world, and does not thrive on the productive ambiguity of meaning. Machinic agency does not suffer the incongruence of a self with itself.

D. The Artificial Intelligence of Human Law

Modern positive law is an artificial construct, not in the naïve sense of social constructivism or neoliberal voluntarism, but in the sense that it is a text-driven artifact contingent upon the performative nature of human discourse. This implies that legal effect is not a matter of brute force or mechanical application, but a matter of ensuring what use of language *counts as* having what

---

[16]*See* David H. Wolpert, *Ubiquity Symposium: Evolutionary Computation and the Processes of Life: What the No Free Lunch Theorems Really Mean: How to Improve Search Algorithms*, 2013 UBIQUITY 2 (2013).

[17]*See* J.L. AUSTIN, HOW TO DO THINGS WITH WORDS (2nd ed. 1975); NEIL MACCORMICK, INSTITUTIONS OF LAW: AN ESSAY IN LEGAL THEORY (2007).

[18]*See* Ricoeur, *supra* note 15; Charles Taylor, *To Follow a Rule, in* PHILOSOPHICAL ARGUMENTS 165–81 (Charles Taylor ed., 1995).

effect. The effect is not causal but performative. Therefore, violating the criminal law does not result in punishment but in punishability, which sets free the legitimate use of violence if specified conditions are met. This integrates legal instrumentality with legal protection, in that the criminal law hopes to reduce crime *in a way that* protects against arbitrary use of state violence.

The artificial nature of human law acts upon the shared world, as positive law grants permissions, attributes property, imposes obligations to pay compensation or duties to refrain from interfering with public goods or private interests. This often requires a balancing act, for instance, where two or more fundamental rights are incompatible in a concrete case, or where a fundamental right is incompatible in a concrete case with a prevailing public good such as national security or privacy. The incompatibility at the level of concrete cases, however, does not imply incompatibility per se. Decisions on how to balance competing antinomian goals in practice depends on respecting their constitutive force. Paraphrasing Radbruch, modern positive law co-determines and protects the shared world by sustaining the antinomian character of the key incentives that drive law in a constitutional democracy: legal certainty, instrumentality, and justice.[19]

Without legal certainty, the rule of law would either collapse into ethics and come to depend on the ethical inclinations of those in power and authority, or collapse into arbitrary rule by law, undoing the checks and balances secured by an independent judiciary. Without instrumentality, law would collapse into morality. This would force policy makers to find other instruments to make their policies operational, for example, techno-regulation. Without justice, positive law collapses into the kind of legalism of formal positivism, which separates law entirely from morality, or into a rule by men that delivers us to the whims of whoever is in power or authority. This implies that attempts to resolve the antinomian nature of law could end up with the destruction of the shared world, as the shared world in a constitutional democracy is grounded in modern, positive law. Legal positivism reduces law to a rigid type of legalism. Natural law, which equates legal philosophy with moral philosophy, reduces law to ethics. Finally, the "Real Politik" of a cynical political science, or even critical legal studies, reduces law to a distribution of risks and opportunities entirely dependent on power play.

The antinomian character of modern positive law accords with the ex-centric positionality of human nature. Both highlight the underdetermined nature of human agency and the shared world it draws on and nourishes. The intelligence of human individuals and human collectives thrives on this constitutive undecidability; it can therefore be simulated but not instituted by machinic agents.

E. The Challenge for the European Legislator: Getting AI Right

Once we begin to value automated inferencing (AI) for what it is, the playground for building reliable AI applications can be expanded. The incredible challenges we face as a result of climate change may require extensive use of reliable AI. For instance, we may need AI tools when developing a sustainable approach to energy and water shortages, droughts, turbulent weather conditions with potentially catastrophic consequences, disaster management, and reconfiguration of labor markets due to economic migration. Reliable AI can only be developed if it is based on a sound and contestable research design anchored in the core tenets of reproducible open science.[20] It cannot be based on the core tenets of seductive marketing strategies grounded in the manipulative assumptions of behaviorist nudge theory.[21]

---

[19]Radbruch, *supra* note 3. *See* Hildebrandt, *supra* note 3. *See also* ANTHONY CARTY, PHILOSOPHY OF INTERNATIONAL LAW 241–45 (1st ed. 2007), one of the very few lawyers engaging with the relevance of Plessner's investigation of the human condition for the law, referring to HELMUTH PLESSNER, THE LIMITS OF COMMUNITY: A CRITIQUE OF SOCIAL RADICALISM (Andrew Wallace trans., 1999).

[20]*See* Xavier Bouthillier et al., *Unreproducible Research is Reproducible*, *in* INTERNATIONAL CONFERENCE ON MACHINE LEARNING 725–34 (2019).

[21]*See* Nick Seaver, *Captivating Algorithms: Recommender Systems As Traps*, J. MATERIAL CULTURE (2018).

This is where European law will have to assert and reinvent its artificial intelligence as anchored in the checks and balances of the rule of law. Plessner's anthropology may help in highlighting how the ex-centric positionality that is core to human intelligence informs the antinomian character of law in a constitutional democracy that was highlighted by Radbruch. This has foundational implications for the employment of automated inferences, machinic agency, and machine intelligence. Jurisdictions that understand the limitations of machine inferences that feed on machine readable human behaviors will gain a competitive advantage compared to jurisdictions that fail to take into account the ex-centric positionality of human agency.