



Volume 9 | Issue 4



CONCEPT

Smart technologies

Mireille Hildebrandt *Vrije Universiteit Brussel* mireille.hildebrandt@vub.be



OPEN
ACCESS

DOI: <https://doi.org/10.14763/2020.4.1531>

Published: 17 December 2020

Received: 12 September 2020 **Accepted:** 13 November 2020



PEER
REVIEWED

Competing Interests: The author has declared that no competing interests exist that have influenced the text.

Licence: This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License (Germany) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. <https://creativecommons.org/licenses/by/3.0/de/deed.en>
Copyright remains with the author(s).

Citation: Hildebrandt, M. (2020). Smart technologies. *Internet Policy Review*, 9(4).
<https://doi.org/10.14763/2020.4.1531>

Keywords: Artificial intelligence, Cybernetics, Manipulation, Nudge theory, Machine learning

Abstract: Speaking of 'smart' technologies we may avoid the mysticism of terms like 'artificial intelligence' (AI). To situate 'smartness' I nevertheless explore the origins of smart technologies in the research domains of AI and cybernetics. Based in postphenomenological philosophy of technology and embodied cognition rather than media studies and science and technology studies (STS), the article entails a relational and ecological understanding of the constitutive relationship between humans and technologies, requiring us to take seriously their affordances as well as the research domain of computer science. To this end I distinguish three levels of smartness, depending on the extent to which they can respond to their environment without human intervention: logic-based, grounded in machine learning or in multi-agent systems. I discuss these levels of smartness in terms of machine agency to distinguish the nature of their behaviour from both human agency and from technologies considered dumb. Finally, I discuss the political economy of smart technologies in light of the manipulation they enable when those targeted cannot foresee how they are being profiled.

This article belongs to **Concepts of the digital society**, a special section of *Internet Policy Review* guest-edited by Christian Katzenbach and Thomas Christian Bächle.

Introduction: the rise of smart ‘anywares’

A wide range of products and services is currently discussed under the heading of ‘smart’, loosely overlapping with the technologies of (cloud) robotics, Internet of Things (IoT) and more generally to hardwired applications of AI.¹ More concretely we can think of semi-autonomous vehicles (smart cars), cyberphysical infrastructures such as homes animated with sensor technologies that enable surreptitious adaptation of temperature, light, and all kinds of services (smart homes), or energy networks that afford real-time demand-side energy supply systems (smart grids) and myriad online architectures that require responsive and adaptive interaction, such as collaborative platforms (smart office) and e-learning (smart education). Most of these systems will be hybrid between online and offline, as exemplified by crime mapping (smart policing), social security fraud detection (smart forensics), the sharing economy (smart services), remote healthcare (smart healthcare) and initiatives in the realm of computational law that employ the alluring notion of *smart law* to sell their services. We seem to have arrived in the era of smart *every-ware* (Greenfield, 2006), confronted with myriad smart *anywares*.

In this article I first ground the concept of *smartness* in the history of artificial intelligence and cybernetics (sections 1 and 2), highlighting the different strands in the research domains that enabled the development of smart technologies. Though this will bring out the smartness of those selling the idea of computational systems as cognitive engines, there is more to smart technologies than some social constructivists may wish to believe.² Grounded in postphenomenological philosophy and embodied enaction I will develop a layered approach to the kind of machine agency involved in smart technologies (section 3) that allows me to simultaneously reject confounding human and machine agency while nevertheless acknowledging the fact that their computational inferences are used to anticipate our behaviours, thus introducing agential affordances to our environments. Coupled with nudge theory such affordances have resulted in attempts to manipulate

1. One could write a history of these loosely overlapping concepts, for instance noting the rise of previous concepts such as *ubiquitous computing* (Weiser, 1991), *ambient intelligence* (Berg, 2010) that preceded the term *Internet of Things* (ITU, 2005).
2. A good pointer to how we might steer free from both technological determinism and social constructivism may still be Latour (1992), though my own position is rooted in e.g., Winch (1958), Ihde (1990), Gibson (2014), and Varela, Thompson, and Rosch (1991)

consumers (behavioural advertising, credit rating) and further manipulation in the political realm that has been reconfigured as a marketplace under the influence of neoliberal economics (section 4). In the final part I address the need to confront the affordances of smart environments, taking their machine agency seriously without buying into either marketing speak or social constructivist relativism (section 5).

1. Smart-arse technologies: avoiding the AI debate

Why would one call a technology smart? The Cambridge English Dictionary (2020) tells us that smart is about 'having a clean, tidy, and stylish appearance (mainly in the UK)', and/or refers to someone who is 'intelligent, or able to think quickly or intelligently in difficult situations (mainly in the US)'. The geographical difference raises some interesting questions about cultural inclinations, but that may be merely a smart-arse move on my side, taking note that a smart-arse is defined in the same dictionary as 'someone who is always trying to seem more clever than other people in a way that is annoying'. The answer must be situated in the computer-related adjective, which is defined as '[a] smart machine, weapon, etc. uses computers to make it work so that it is able to act in an independent way: Until the advent of smart weapons, repeated attacks were needed to ensure the destruction of targets' (Cambridge Dictionary 2020). This must be what my article here aims to describe: technologies with a mindless mind of their own, 'getting on with things' such that we have time to do more important things. Let's not forget, however, that the term normally refers to something aesthetic (perhaps to sleek interfaces that intuitively guide our usage based on a clean, tidy and stylish interface) and that it otherwise refers to a clever type of acuity, or intelligence. Acting in an independent way comes close, though one wonders in what sense smart weapons act independently, considering the fact that they are designed by humans with specific computational skills in mind and are also used by humans to destroy very precise targets. The independence seems to only refer to how these smart technologies achieve the goals set for them, even if one could say that they may be able to develop subgoals if they 'think' it will help to achieve our goals (Boulanin & Verbruggen, 2017). One could ask who the smart-arse is here: the human or the machine.

McCarthy, the AI pioneer who coined the term artificial intelligence (AI) in the famous 1956 Dartmouth conference (Leslie, 2019), certainly was smart. One of the other founding fathers of what we now refer to as AI, Herbert Simon, preferred to stick to the term 'complex information processing' (Leslie, 2019 at 32), because it

supposedly better described the operations they were conceptualising. McCarthy, however, had the luminous insight that nobody would invest much in something so dry as ‘complex information processing’, whereas the mysterious idea of ‘artificial intelligence’ might trigger some real interest (Leslie, 2019 at 32) – let’s acknowledge McCarthy’s skills in the realm of public relations (PR). One of the reasons why I used the term *smart technologies* in one of my books, was to avoid the endless debates about whether machines can be intelligent, what is meant with intelligence and whether and, if so, how machine intelligence differs from both animal and human intelligence. Perhaps the more intriguing question is why it was not simply called machine or computational intelligence. I would actually claim that human intelligence itself is deeply artificial (Hildebrandt, 2020b; Plessner & Bernstein 2019). Speaking of machine or computational intelligence would probably highlight the difference that McCarthy wished to ignore, as becomes clear in the (cl)aim of those early pioneers who professed that (Moor, 2006):

[t]he study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.

Against such overreach, I dare say that the advantage of the term *smart technologies* is first and foremost that it avoids metaphysical debates on (1) whether machines ‘think’, (2) whether it suffices to answer that question in terms of observable behaviour only, and perhaps also (3) the debate on whether intelligence requires ‘thinking’ in the first place, considering that viruses and reptiles do very well in terms of independent navigation – even though we may think that they don’t think.³ Speaking of *smart* instead of *AI* technologies allows us to be more pragmatic about their added value, steering free from the pseudo-religious (and thereby pseudo-scientific) claims around ‘artificial intelligence’ (Natale & Ballatore, 2020).

This may, however, result in those advocating smart technologies getting away with even bolder claims about their added value, precisely because the threat of a competing intelligence is sidetracked. Adopting the term may therefore be a smart

3. One may suggest that viruses don’t independently navigate their environment. I use the term *navigate* in a broader sense than intended physical movement (even humans navigate their environment in a broader sense, e.g. their institutional environment, and let’s note that much physical movement is induced by our autonomic nervous system rather than the result of conscious intent). The 2020 pandemic has shown the extraordinary intelligence of a virus’ global navigation, see on the viral communication that informs their navigation e.g. (Dolgin, 2019).

move to downplay the extent to which smart technologies are capable of reconfiguring our choice architecture. This reconfiguration is often meant to manipulate our intent, as the entanglement of nudge theory and machine learning demonstrates in the realm of behavioural targeting (for advertising, recruiting, insurance, admission to college or parole decisions). Such entanglement hides the choices made at the backend of computing systems by those who stand to gain from their employment, or those intent on serving our best interests (though preferably behind our backs, see e.g. Yeung, 2017).

This confronts us with the question of ‘smart for whom?’, which is a political question. Ignoring it will not make such backend choices neutral. Perhaps we should coin these technologies, smart-arse technologies, highlighting their ability to manipulate our environment and our selves, without, however, attributing thoughts or intent to technologies that can only follow their masters’ script. Though their masters may actually be at a loss, due to the interacting network effects and emergent behaviours they generate (known as the Goodhart effect) (Strathern, 1997).

2. Smart technologies: embracing the cybernetic approach

Being at a loss is deeply human. Machines are never at a loss, because they lack the mind for it. But machines or technologies can help humans to gain control, if they somehow get smart about how to achieve their goals (I am glad to leave ‘they’ undefined here). The aim of increased control, often achieved by developing remote control, is key to another branch of what became AI, namely cybernetics (Wiener, 1948) or operations research (Pickering, 2002). Cybernetics is focused on developing the right kind of feedback loops, such that a technology is capable of adapting its behaviour to changing circumstances, based on its perception. This fits well with the idea that perception is triggered by the need to anticipate how one’s actions will affect one’s environment, and how this will in turn affect one’s action-potential, familiar in phenomenological research and affordance theory (Kopjar, 2016; Gibson, 2014). A narrower, machinic understanding often describes cybernetics as a specific type of regulation, entailing standard-setting, behaviour monitoring and behaviour modification. This obviously accords with computational systems, which require formalisation and thrive on disambiguated instructions. One could say that these systems require formalised regulation in order to regulate their environment, even if their approach is data-driven and based on machine learning. Even in that case input must be formalised (digitised, labelled or clustered) and the patterns mined are based on mathematical hypotheses (Blockeel,

2010). *Cyber* is Greek for steering (Hayles, 1999 at 8), and though steering can be done in many ways, computational systems necessarily reduce the art of steering to computational ‘algorithmics’. One could surmise that the notion of smart technologies is connected with the idea of remote control over an environment, for instance in the case of hardware failure detection combined with automated interventions once failure is detected.

In point of fact, SMART technologies originally referred to the acronym of self-monitoring, analysis and reporting technology (SMART), mainly used for failure prediction in hardware (Gaber et al., 2017). This is interesting because such SMART-ness remains focused on monitoring and reporting, without necessarily including an automated or even autonomous response in order to repair. In that sense the current use of the term *smart* is different as it refers to systems capable of responding to changes in their environment based on input data, and it is precisely the notion of control and the ability to steer that are central to what we now call smart technologies, both within the context of online and offline systems. Smart technologies are digital technologies of control, and the most important question will be whose control they facilitate. We will return to this question in section 4.

3. Levels of being smart: the diversity of non-human agents

The term *smart* is not a technical term within either computer science, law, the life sciences or social science. It may have been initiated as a marketing term, hoping to lure investors and potential users, offering an intuitive way to speak about computational systems that offer real world interventions supposedly achieving efficiency, convenience, or novel products and services. Though ‘stuff’ may be sold under the heading of *smart* even if it has no adaptive qualities, the core idea is that smart technologies do have some level of agency, in the practical sense of being able to respond to input data with behaviours that are conducive to our goals (who “our” refers to is the key question, see the final section).

Smartness, however, is not a categorical attribute of a technology, as there are different levels of being smart. Much depends on the extent to which its adaptive behaviours have been premeditated by the developer, which corresponds with the extent to which a technology can be said to express its own agency. Simultaneously, much depends on the complexity of the environment, taking note of the fact that environments are often stabilised or even closed, to ensure safe functioning of a smart technology. In robotics the environment of the robot, called its ‘envelop’,

is usually developed in tandem with the robot, aligning its ‘degrees of freedom’ with the physical properties of the envelop (Floridi, 2014a). This way developers can more easily foresee the scenarios they must code for, preventing undefined behaviours of the device. Many roboticists foresee that self-driving cars will only hit the road if the road infrastructure is locked off from e.g. pedestrians, making sure the car runs within a relatively predictable environment, meeting only other self-driving cars on connected roads (Brooks, 2018).

If we acknowledge the fact that a technology may exhibit a specific type of agency, we can move away from discussions of whether they ‘have’ humanesque agency or not. My take is that they don’t, because they are embodied otherwise and that has consequences (Pfeifer & Bongard, 2007). Also, why invest in human-like agents if we already have wonderful ways to reproduce? Aficionados such as Kurzweil (2005) will claim that artificial agents will, however, soon outperform us and reach a singularity where humanity becomes a historical footnote to an intelligence too advanced for us to even conceive (though he did, nevertheless). I would say that here we enter the realm of pseudo-religious devotion if not hubris, observing that most believers in such singularity come from outside the domain of computer science, uninformed by the limits imposed by physics, complexity and much more (Walsh, 2017). It makes more sense to acknowledge that plants, animals and smart technologies all have different but nevertheless pivotal types of agency, and thereby differ from stones, screwdrivers and volcanoes. This will allow us to focus on figuring out what kind of agency is typical for smart technologies and raise the crucial question of how it may affect human agency. In this section I will discuss three levels of *smartness*, in the next I will probe into the implications for human agency.

Based on a series of seminal definitions of agency (Floridi & Sanders, 2004; Pfeifer & Bongard, 2007; Russell et al., 2010; Steels, 1995; Varela et al., 1991) I have compiled a definition that depicts agency as (Hildebrandt, 2015 at 22):

the capability:

- to act autonomously,
- to perceive and respond to changes in the environment,
- to endure as a unity of action over a prolonged period of time and
- to adapt behaviours to cope with new circumstances.

Let's note that the term *autonomous* here does not refer to moral autonomy and does not assume human *mindness*, it simply refers to an entity capable of goal-oriented behaviour without intervention of its creator, designer or engineer. Let's also note that the most crucial aspect of agency is the combination of perception and action; a simple thermostat perceives temperature and if a threshold is met it acts by triggering the heating system. Though nobody would imply that a thermostat can think, is autonomous in the sense of deciding which temperature we prefer or is in need of human rights to protect their dignity, a thermostat differs substantially from a water tap that only provides water after we open it. The thermostat has a minimal type of agency. Interestingly, we would not call such a thermostat smart, because it is not learning from our behaviour, it does not adapt to the way we set its temperature.

A smart thermostat, however, responds to our behaviour by adapting its own behaviour in line with how it profiles us. This may suit us, or suit those who profit from the data it provides about our behaviour. The smart thermostat may not be designed to endure as a unity of action over a prolonged period of time, which would imply self assessment and self repair. Similarly, it may not be designed to reconfigure its architecture to survive in a changing environment. And this brings me back to the first aspect of agency, being autonomy.

Here we need to distinguish between automation and autonomy. A simple and a smart thermostat both display automation, but their level of autonomy differs. Though one could claim that both are capable of 'goal-oriented behaviour without intervention of its creator, designer or engineer', the smart thermostat has a broader palette of goal-oriented behaviour due to its learning capacity. The difference between automation (which does not imply agency) and autonomy (which does) is usually sought in the ability of a system to respond to its environment in a way that sustains its own endurance and prevents or repairs failure modes, notably in the face of uncertainty (Endsley, 2017). This means that autonomy here refers to a technology capable of behaving as an adaptive unity of action over the course of time, highlighting the overlap and the interaction with the different aspects of machine agency or *smartness*.

This has led me to distinguish three levels of machine agency (Hildebrandt, 2015, chapter 2):

1. **Agency of logic-based** systems that are by definition deterministic and explainable, though in real life their interaction with unforeseen events may cause unforeseeable behaviour and the more complex their decision-tree and the more complex their environment, the more difficult it may be

to reconstruct a correct explanation. A simple thermostat is a fine example of such a system, but so is a fraud detection system based entirely on a complex but preconceived decision tree. This is the type of agency closely related to good old fashioned artificial intelligence (GOF AI), grounded in knowledge representation and symbol manipulation (Haugeland, 1997, at 16).

2. **Agency of machine learning systems** that are capable of perceiving their environment (in terms of data), and of improving their performance of a specified task based on such perception. Such systems thrive on iterant feedback loops that are validated in terms of pattern recognition, which is ultimately grounded in statistical correlations expressed in mathematical functions. Examples of these types of systems abound: from e-learning to fintech and from load balancing in smart energy networks to facial recognition technology, machine translation as well as personal digital assistants such as Siri or Alexa. This type of agency is closely related to cybernetics and to *artificial intelligence the modern approach* (AIMA), grounded in inductive inferences and predictive analytics (Russell, Norvig, & Davis, 2010).
3. **Agency of multi-agent systems (MASs)** where a large number of agents interact, each based on their own goal or scripts, which results in emergent behaviours of the overall system due to network effects and complex dynamics (Vázquez-Salceda, Dignum, & Dignum, 2005). The agents can be logic-based or based on machine learning and the environment that a MAS navigates and regulates can be closed or open, predictable or unpredictable. Interestingly, a MAS will often be a distributed system that in itself constitutes an environment for humans and other types of agents, thus blurring the borders between an agent, an agent system and an environment. Whereas a smart fridge may be an identifiable agent in my home environment, my smart home that consists of interacting agents working together to better serve my needs (or those of the provider) is an environment rather than an agent. This is precisely where things become complicated, opaque and potentially turbulent. Some have termed this situation, where the environment itself takes on agency, an *onlife world* (Floridi, 2014b). Others have highlighted how the continuous and reflexive updating of one's environment may reshape our own agency, though some do and others don't make a difference between the nature of expectations between human agents and those between a human and a smart environment (Dittrich, 2002; Esposito, 2017).

Obviously, this tripartite categorisation is not the ultimate way to frame different levels of smartness, while noting that these levels interact and morph where smart technologies meet. This is where the difference between smart technologies and smart infrastructures becomes pivotal, especially since so many smart infrastructures are data-driven and surreptitiously reconfigure our choice architecture and

our life world on the basis of our behavioural data. As indicated, this has been framed under the headings of Ambient Intelligence, the Internet of Things, cloud robotics, smart cities, connected cars or smart energy grids. All these headings refer to what computer science now refers to as cyberphysical infrastructures (Suh et al., 2014). As such infrastructure increasingly configures our environments, it becomes critical infrastructure. This will in turn augment our dependence on networked smart technologies and the energy requirements they involve, while also raising a number of issues of distributive justice, both with regard to access and with regard to potential unfair distribution of risks.

Although there seems to be an urge to discuss these societal implications in terms of ethics, I prefer to end this article with a discussion of the political economy of smart technologies, including some pointers to the importance of democracy and the rule of law. The main reason for this preference is the need to counter attempts to engage in 'ethics-washing' and 'ethics shopping' (Wagner, 2018) by teasing out the power relationships that are at stake. This is not to suggest that ethics is not relevant here (Hildebrandt, 2020a, chapter 11).

4. A political economy of smart technologies

For many decades economic analysis has been in the thralls of neo-classical economics, originating from the Chicago school of e.g. Milton Friedman, Gary Becker and Richard Posner, influenced by the Austrian-British philosopher and economist Hayek (2007), whose 1940 *The Road to Serfdom* celebrated an unconstrained market libertarianism. Becker's rational choice theory inspired the Chicago school of political economy, applying the atomistic type of methodological individualism that grounds rational choice theory when used in for instance crime control and antitrust policy. Those unimpressed by the somewhat metaphysical and unsubstantiated claims of libertarian market fundamentalism usually address the conglomerate of neo-classical economic policies as 'neoliberalism' (Blakely, 2020). Under that heading I would also include the application of Kahneman and Tversky's (1979) models of cognitive science to economics, for which they won the Nobel Prize. Their particular strand of neoliberalism identifies as behavioural economics, popularised as *nudge theory* by Thaler and Sunstein (2008). Nudge theory, in turn, fits very well with machine learning applications that aim to manipulate 'users' of smart technologies into specific behaviours, whether purchasing, voting or eating. Both nudge theory and machine learning thrive on an atomistic variant of methodological individualism that is closely aligned with utilitarianism (Hildebrandt, 2020a, chapter 11). Nudge theory's alliance with behavioural economics has result-

ed in paternalistic libertarianism (Sunstein, 2016), though not everyone buys into the sneaky tyranny it professes (Yeung, 2017). Meanwhile most of our free services feed on behavioural advertising, which is firmly grounded on the pseudo-science of behavioural nudging by way of real-time-bidding systems, ultimately built on the quicksand of the same contested branch of cognitive science mentioned above (see Gigerenzer (2018) and Lepenies and Małecka (2019) on the pitfalls of nudge theory). The unwieldy marriage between machine learning and nudge theory has nevertheless been a major success for the rather deep pockets of big tech companies (Frederik & Martijn, 2019; Lomas, 2019a, 2019b).

To properly understand the political economy of smart technologies it may be more interesting to study the work of Karl Polanyi (2012), whose seminal *The Great Transformation* is inspiring a new generation of scholars. For instance, Benkler (2018), Britton-Purdy et al. (2020) and Cohen (2019) are intent on a better understanding of how law and the rule of law are implicated in the rise of economic superpowers such as big tech. They demonstrate the need to reinvent the countervailing powers of Montesquieu's checks and balances, or calling for new versions of Polanyi's counter-movements to put a halt to predatory capitalism. It is interesting that some of these scholars have been studying the advent of smart technologies for decades, notably investigating how the smart aspects of big tech products and services invalidate previous forms of legal protection, playing into the hands of those who control the backend systems.

In my own *Smart Technologies and the End(s) of Law* (Hildebrandt, 2015), I argue that we must learn to interact with smart systems, instead of perpetuating the illusion that we are using them whereas these systems are often using us (as data engines for their data-driven nudging). Moving from usage to interaction will require keen attention to effective and actionable transparency (Schraefel et al., 2020) and what some have called 'throttling mechanisms' that slow down high-frequency profiling by big tech platforms (Ohm, 2020). It seems clear that the kind of transparency and the human timescale that is needed to reorient cybernetic control back to individual human beings will require legislation, supervisory oversight and dedicated case law to rebalance current power relationships (Hildebrandt, 2018). Though developers of smart technologies often speak of human centred design, the current economic incentive structure reduces such humane approaches to PR about 'the human in the loop'. Instead, we need an approach that puts the machine back in its proper place: smart technologies should be 'in the loop', while human beings navigate an animated technological landscape that serves them instead of spying on them (Zuboff, 2019).

5. Conclusion: human acuity and smart machines

Speaking of smart technologies could have the advantage of steering free from the snake oil narratives of AI, while still hinting at the cybernetics that inspires their design. But as it is not a technical term it can mean anything to anybody, and this becomes a drawback when it is used to sell systems based on unsubstantiated claims, as often happens under the heading of AI or machine learning (Giles, 2018; Lipton & Steinhardt, 2018).

Users and those forced to live with these systems (in smart cities, under surveillance of smart policing, at the mercy of smart recruiting) should ask themselves “smart for whom?”, “smart in what sense?” and “smart compared to what?”. Though many so-called *smart* applications are sold as ‘outperforming humans’ there is much to say about this (Cabitza et al., 2017), as such bombastic claims are based on performance metrics that refer to accuracy in the context of a data set, which may say little about the real-world performance (e.g. Caruana et al., 2015). Humans do not live in a data set, they navigate the material and institutional fabric of a real world, and they flourish based on an acuity that is not within the remit of smart machines.

Some will say the latter is an open question, others are certain that artificial general intelligence is around the corner. Still others focus their attention on how smart technologies affect human agency, taking a more pragmatic and simultaneously concerned approach. Referring to technologies as *smart* hopefully avoids metaphysical beliefs in technologies that cannot but follow a script, even if that script allows them to reorganise their internal structure to perform better on a given task. What matters is that human beings must be able to anticipate how the machinic agency of clever engineering feats profiles them, noting the difference between the fragile beauty of human acuity and the precocious foresight of machines that result from human ingenuity.

References

Benkler, Y. (2018). The Role of Technology in Political Economy: Part 1 [Blog post]. *The Law and Political Economy Project*. <https://lpeproject.org/blog/the-role-of-technology-in-political-economy-part-1/>

Blakely, J. (2020). How Economics Becomes Ideology: The Uses and Abuses of Rational Choice Theory. In P. Róna & L. Zsolnai (Eds.), *Agency and Causal Explanation in Economics* (pp. 37–52). Springer International Publishing. https://doi.org/10.1007/978-3-030-26114-6_3

Blockeel, H. (2010). Hypothesis Space. In C. Sammut & G. I. Webb (Eds.), *Encyclopedia of Machine Learning* (pp. 511–13). Springer US. https://doi.org/10.1007/978-0-387-30164-8_373

Boulanin, V., & Verbruggen, M. (2017). *Article 36 Reviews: Dealing with the Challenges posed by Emerging Technologies* [Report]. Stockholm International Peace Institute. <https://www.sipri.org/publications/2017/other-publications/article-36-reviews-dealing-challenges-posed-emerging-technologies>

Britton-Purdy, J., Singh Grewal, D., Kaczynski, A., & Rahman, K. S. (2020). Building a Law-and-Political-Economy Framework: Beyond the Twentieth-Century Synthesis. *The Yale Law Journal*, 129(6), 1784–1835. <https://www.yalelawjournal.org/feature/building-a-law-and-political-economy-framework>

Brooks, R. (2018, January 1). My Dated Predictions [Blog post]. *Robots, AI, and Other Stuff*. <https://rodneybrooks.com/my-dated-predictions/>

Cabitza, F., Rasoini, R., & Gensini, G. F. (2017). Unintended Consequences of Machine Learning in Medicine. *JAMA*, 318(6), 517–518. <https://doi.org/10.1001/jama.2017.7797>

Cambridge English Dictionary. (2020). Smart. In *Cambridge English Dictionary Online*. <https://dictionary.cambridge.org/dictionary/english/smart>

Caruana, R., Lou, Y., Gehrke, J., Koch, P., Sturm, M., & Elhadad, N. (2015). Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-day Readmission. *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1721–1730. <https://doi.org/10.1145/2783258.2788613>

Cohen, J. E. (2019). *Between Truth and Power: The Legal Constructions of Informational Capitalism*. Oxford University Press. <https://doi.org/10.1093/oso/9780190246693.001.0001>

Dittrich, P. K. (2002). On the Scalability of Social Order—Modeling the Problem of Double and Multi Contingency Following Luhmann. *Journal of Artificial Societies and Social Simulation*, 6(1). <http://jass.soc.surrey.ac.uk/6/1/3.html>

Dolgin, E. (2019). The Secret Social Lives of Viruses. *Nature*, 570(7761), 290–92. <https://doi.org/10.1038/d41586-019-01880-6>

Endsley, M. R. (2017). From Here to Autonomy: Lessons Learned From Human–Automation Research. *Human Factors*, 59(1), 5–27. <https://doi.org/10.1177/0018720816681350>

Esposito, E. (2017). Artificial Communication? The Production of Contingency by Algorithms. *Zeitschrift Für Soziologie*, 46(4), 249–265. <https://doi.org/10.1515/zfsoz-2017-1014>

Floridi, L. (2014a). Agency: Enveloping the World. In *The Fourth Revolution: How the infosphere is reshaping human reality*. Oxford University Press.

Floridi, L. (2014b). *The Onlife Manifesto—Being Human in a Hyperconnected Era*. Springer. <https://doi.org/10.1007/978-3-319-04093-6>

Floridi, L., & Sanders, J. W. (2004). On the Morality of Artificial Agents. *Minds and Machines*, 14(3), 349–379. <https://doi.org/10.1023/B:MIND.0000035461.63578.9d>

Frederik, J., & Martijn, M. (2019, November 6). The new dot com bubble is here: It's called online advertising. *The Correspondent*. <https://thecorrespondent.com/100/the-new-dot-com-bubble-is-here-its-called-online-advertising/13228924500-22d5fd24>

- Gaber, S., Ben-Harush, O., & Savir, A. (2017). Predicting HDD failures from compound SMART attributes. *Proceedings of the 10th ACM International Systems and Storage Conference*, 1. <https://doi.org/10.1145/3078468.3081875>
- Gibson, J. J. (2014). *The Ecological Approach to Visual Perception* (1st ed.). Routledge. <https://doi.org/10.4324/9781315740218>
- Gigerenzer, G. (2018). The Bias Bias in Behavioral Economics. *Review of Behavioral Economics*, 5(3–4), 303–336. <https://doi.org/10.1561/105.00000092>
- Giles, M. (2018, September 13). Artificial intelligence is often overhyped—And here's why that's dangerous. *MIT Technology Review*. <https://www.technologyreview.com/2018/09/13/240156/artificial-intelligence-is-often-overhypedand-heres-why-thats-dangerous/>
- Greenfield, A. (2006). *Everyware. The dawning age of ubiquitous computing*.
- Haugeland, J. (Ed.). (1997). *Mind Design II: Philosophy, Psychology, and Artificial Intelligence* (2nd Revised, Enlarged ed.). The MIT Press.
- Hayek, F. A. (2007). *The Road to Serfdom: Text and Documents – The Definitive Edition* (B. Caldwell, Ed.). University of Chicago Press.
- Hayles, K. N. (1999). *How we became posthuman. Virtual bodies in cybernetics, literature, and informatics*. University of Chicago Press.
- Hildebrandt, M. (2015). *Smart Technologies and the End(s) of Law*. Edward Elgar. <https://doi.org/10.4337/9781849808774>
- Hildebrandt, M. (2018). Primitives of Legal Protection in the Era of Data-Driven Platforms. *Georgetown Law Technology Review*, 2(2), 252–273. <https://georgetownlawtechreview.org/primitives-of-legal-protection-in-the-era-of-data-driven-platforms/GLTR-07-2018/>
- Hildebrandt, M. (2020a). *Law for Computer Scientists and Other Folk*. Oxford University Press. <https://doi.org/10.1093/oso/9780198860877.001.0001>
- Hildebrandt, M. (2020b). The Artificial Intelligence of European Union Law. *German Law Journal*, 21(1), 74–79. <https://doi.org/10.1017/glj.2019.99>
- Ihde, D. (1990). *Technology and the Lifeworld: From Garden to Earth*. Indiana University Press.
- I.T.U. (2005). *The Internet of Things*. International Telecommunications Union.
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263–292. <https://doi.org/10.2307/1914185>
- Kopljar, S. (2016). *How to think about a place not yet: Studies of affordance and site-based methods for the exploration of design professionals' expectations in urban development processes* [PhD Thesis, Lund University]. <http://lup.lub.lu.se/record/4bdb64d0-8551-44c7-aabc-26888e0cfbb4>
- Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. Viking.
- Latour, B. (1992). Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts. In W. E. Bijker & J. Law (Eds.), *Shaping Technology / Building Society* (pp. 225–258). MIT Press.
- Lepenies, R., & Matecka, M. (2019). Behaviour Change: Extralegal, Apolitical, Scientific? In H. Straßheim & S. Beck (Eds.), *Handbook of Behavioural Change and Public Policy* (pp. 344–360). Edward Elgar Publishing. <https://doi.org/10.4337/9781785367854.00032>

- Leslie, D. (2019). Raging robots, hapless humans: The AI dystopia. *Nature*, 574(7776), 32–33. <http://doi.org/10.1038/d41586-019-02939-0>
- Lipton, Z. C., & Steinhardt, J. (2018). Troubling Trends in Machine Learning Scholarship. *ArXiv*. <http://arxiv.org/abs/1807.03341>
- Lomas, N. (2019a). The case against behavioral advertising is stacking up. In *TechCrunch*. <https://techcrunch.com/2019/01/20/dont-be-creepy/>
- Lomas, N. (2019b, May 31). Targeted ads offer little extra value for online publishers, study suggests. *TechCrunch*. <https://social.techcrunch.com/2019/05/31/targeted-ads-offer-little-extra-value-for-online-publishers-study-suggests/>
- Moor, J. (2006). The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years. *AI Magazine*, 27(4), 87–87. <https://doi.org/10.1609/aimag.v27i4.1911>
- Natale, S., & Ballatore, A. (2020). Imagining the thinking machine: Technological myths and the rise of artificial intelligence. *Convergence*, 26(1), 3–18. <https://doi.org/10.1177/1354856517715164>
- Ohm, P. (2020). Throttling machine learning. In M. Hildebrandt & K. O'Hara (Eds.), *Life and the Law in the Era of Data-Driven Agency* (pp. 214–229). Edward Elgar. <https://doi.org/10.4337/9781788972000.00019>
- Pfeifer, R., & Bongard, J. (2007). *How the Body Shapes the Way We Think. A New View of Intelligence*. MIT Press.
- Pickering, A. (2002). Cybernetics and the Mangle: Ashby, Beer and Pask. *Social Studies of Science*, 32(3), 413–437. <https://doi.org/10.1177/0306312702032003003>
- Plessner, H., & Bernstein, J. M. (2019). *Levels of Organic Life and the Human: An Introduction to Philosophical Anthropology* (M. Hyatt, Trans.). Fordham University Press. <https://doi.org/10.5422/fordham/9780823283996.001.0001>
- Polanyi, K. (2012). *The Great Transformation: The Political and Economic Origins of Our Time*. Amereon Ltd.
- Russell, S. J., Norvig, P., & Davis, E. (2010). *Artificial intelligence: A modern approach*. Prentice Hall.
- schraefel, m. c., Gomer, R., Gerding, E., & Maple, C. (2020). Rethinking transparency for the Internet of Things. In M. Hildebrandt & K. O'Hara (Eds.), *Life and the Law in the Era of Data-Driven Agency* (pp. 100–116). Edward Elgar. <https://doi.org/10.4337/9781788972000.00012>
- Steels, L. (1995). When are robots intelligent autonomous agents? *Robotics and Autonomous Systems*, 15, 3–9. [https://doi.org/10.1016/0921-8890\(95\)00011-4](https://doi.org/10.1016/0921-8890(95)00011-4)
- Strathern, M. (1997). "Improving Ratings": Audit in the British University System'. *European Review*, 5(3), 305–321.
- Suh, S. C., Tanik, U. J., Carbone, J. N., & Eroglu, A. (Eds.). (2014). *Applied Cyber-Physical Systems*. Springer. <https://doi.org/10.1007/978-1-4614-7336-7>
- Sunstein, C. R. (2016). *The Ethics of Influence: Government in the Age of Behavioral Science*. Cambridge University Press.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

- van den Berg, B. (2010). *The Situated Self: Identity in a World of Ambient Intelligence*. Wolf Legal Publishers.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind. Cognitive Science and Human Experience*. MIT Press.
- Vázquez-Salceda, J., Dignum, V., & Dignum, F. (2005). Organizing Multiagent Systems. *Autonomous Agents and Multi-Agent Systems*, 11(3), 307–60. <https://doi.org/10.1007/s10458-005-1673-9>
- Wagner, B. (2018). Ethics as an Escape from Regulation. From "Ethics-Washing" to Ethics-Shopping? In E. Bayamlioglu, I. Baraliuc, L. Janssens, & M. Hildebrandt (Eds.), *Being Profiled: Cogitas Ergo Sum: 10 Years of Profiling the European Citizen* (pp. 84–87). Amsterdam University Press. <https://doi.org/10.2307/j.ctvhrd092.18>
- Walsh, T. (2017). The Singularity May Never Be Near. *AI Magazine*, 38(3), 58–62. <https://doi.org/10.1609/aimag.v38i3.2702>
- Weiser, M. (1991). The Computer for the 21st Century. *Scientific American*, 265(3), 94–104. <https://doi.org/10.1145/329124.329126>
- Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. MIT Press.
- Winch, P. (1958). *The Idea of a Social Science*. Routledge & Kegan Paul.
- Yeung, K. (2017). 'Hypernudge': Big Data as a mode of regulation by design. *Information, Communication & Society*, 20(1), 118–136. <https://doi.org/10.1080/1369118X.2016.1186713>
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile Books.

Published by



ALEXANDER VON HUMBOLDT
INSTITUTE FOR INTERNET
AND SOCIETY

in cooperation with



CREATE

centre
— internet
et societe



R&I IN3
Internet
interdisciplinary
Institute

Universitat Oberta de Catalunya