Deep Neural Networks, Explanations, and Rationality

Edward A. Lee[0000−0002−5663−0584]

University of California, Berkeley, CA 94720, USA
eal@berkeley.edu
https://ptolemy.berkeley.edu/∼eal/

Abstract. "Rationality" is the principle that humans make decisions on the basis of step-by-step (algorithmic) reasoning using systematic rules of logic. An ideal "explanation" for a decision is a chronicle of the steps used to arrive at the decision. Herb Simon's "bounded rationality" is the observation that the ability of a human brain to handle algorithmic complexity and data is limited. As a consequence, human decision-making in complex cases mixes some rationality with a great deal of intuition, relying more on Daniel Kahneman's "System 1" than "System 2." A DNN-based AI, similarly, does not arrive at a decision through a rational process in this sense. An understanding of the mechanisms of the DNN yields little or no insight into any rational explanation for its decisions. The DNN is also operating in a manner more like System 1 than System 2. Humans, however, are quite good at constructing post hoc rationalizations of their intuitive decisions. If we demand rational explanations for AI decisions, engineers will inevitably develop AIs that are very effective at constructing such post hoc rationalizations. With their ability to handle vast amounts of data, the AIs will learn to build rationalizations using many more precedents than any human could, thereby constructing rationalizations for *any* decision that will become very hard to refute. The demand for explanations, therefore, could backfire, resulting in effectively ceding to the AIs much more power.

Keywords: deep neural networks · explainable AI · rationality.

1 Imperfect Intelligence

The momentous AI earthquake that surfaced in the form of ChatGPT in late 2022 surprised even experts in the field. ChatGPT is based on GPT-3.5, a large language model (LLM) from OpenAI. Other examples that emerged around the same time include Google's Bard and Microsoft's Sydney (attached to the Bing search engine). As expressed in Kissinger et al., "[t]he ability of large language models to generate humanlike text was an almost accidental discovery. Further, it turns out that the models also have the *unexpected ability* to create highly articulate paragraphs, articles, and in time perhaps books" (emphasis added) [13]. Everyone was surprised, and even top experts continue to watch with fascination as the machines perform in unexpected ways [3].

The AI researchers who are developing these tools are seeing a relatively gradual evolution of capabilities [10], but even they have been surprised by the outcomes. Because of their expertise in the technology, they are less surprised to be surprised; they have gradually come to expect surprises, but the rest of us were caught off guard. The public witnessed an explosive revelation that contorted expectations.

Many experts have tried to downplay the phenomenon. They argue that the AIs do not understand like we do, they make things up, they plagiarize content from the internet, and they make errors. For example, Chomsky et al., in "The False Promise of ChatGPT," state, "we know from the science of linguistics and the philosophy of knowledge that they differ profoundly from how humans reason and use language" [4]. It is indisputable that the mechanisms of the AIs differ markedly from those of humans, but these authors seem much more confident about the state of the "science of linguistics and the philosophy of knowledge" than might be justified. It is possible that we can learn about human cognition from observing the AIs.

Consider the fact that the AIs make mistakes. As pointed out by Bubek et al. [3], the LLMs acquired the ability to do arithmetic and perform mathematical reasoning by training a language prediction engine. They make no direct use (today) of the arithmetic capability of their machines (which do not make mistakes on arithmetic) nor computer algebra systems such as Maple and Mathematica. It is astonishing to see the emergence of this capability from a token prediction engine. On mathematical problems, my own empirical experimentation reveals that OpenAI's GPT-2 makes the sort of mistakes a child would make, GPT-3.5 makes the sort of mistakes a smart high-school student could make, and GPT-4 makes the sort of mistakes a Berkeley graduate student might make. Could these machines teach us something about how humans reason?

Some people hope that the scope of the AIs will be limited, for example just giving us better search engines. It is not clear, however, where the limits are, or even whether there are any. For example, a previously prevalent presumption that AI would be incapable of creativity was also shattered in 2022 by text-to-image generators such as DALL-E-2 from OpenAI, Stable Diffusion from Stability AI, and Midjourney from the research lab with the same name. These text-to-image tools showed how AIs could absorb stylistic influences, as all human artists do, and then synthesize original works informed by these styles (see Figure 1). Together with the LLMs, these technology releases have led to a massive cultural shift in the public understanding of the role that AI will have in our society and have spurred a gold rush to develop more AI tools.

Consider the fact that LLMs hallucinate, stating as fact things that are not true. This is also a property of human cognitive minds. In fact, it is a property that we highly value when we call it "creative writing." There are currently many disputes over whether images like those in Figure 1 violate the copyrights on the images used to train the AIs, but there is no question that these images are unique. Humans also routinely copy the styles of other artists to produce new creative works.

Fig. 1. Untitled winner of the Sony world photography award, 2023, in the Open Creative Category, generated by an unidentified AI prompted by Boris Eldagsen and winner of the 2022 Colorado State Fair Grand Prize, entitled "Théâtre D'opéra Spatial," generated by Midjourney prompted by Jason Allen.

Chomsky et al. continue, "These differences place significant limitations on what these programs can do, encoding them with ineradicable defects." But the defects they cite strikingly resemble defects in humans. We make stuff up. We parrot lies. We take morally inconsistent positions or weasel our way out of taking a position at all. And we copy material from others.

Chomsky et al. assert that, "the human mind is a surprisingly efficient and even elegant system that operates with small amounts of information." I have studied (and taught) information theory, and any measure I can imagine for the information provided to human brain during its 20 or so years of development into an educated, rational being is not small. They speak of the "miniscule data" and "minimal exposure to information" that lead a child to be able to distinguish between a grammatically well-formed sentence and one that is not. They then cite the "consciously and laboriously ... explicit version of the grammar" constructed by (adult, highly-educated) linguists as evidence that "child's operating system is completely different from that of a machine learning program." It is true that humans can learn from exposure to far less information, but it seems that the ability of the LLMs to learn complex grammar at all is evidence to the contrary. A child learns from examples, exactly like the LLMs. The child is not able to synthesize the explanations that the adult linguists have laboriously constructed. Interestingly, the LLMs *can* synthesize these explanations, but only

because they have "read" all the works of these adult linguists. Leave those texts out of the training data, and their sentences would be no less grammatical.

Chomsky et al. speculate, "[b]ecause these programs cannot explain the rules of English syntax, for example, they may well predict, incorrectly, that 'John is too stubborn to talk to' means that John is so stubborn that he will not talk to someone or other (rather than that he is too stubborn to be reasoned with)." I asked ChatGPT what their sentence means (on March 9, 2023), and it responded:

> The sentence "John is too stubborn to talk to" means that John is so unyielding and inflexible in his opinions, ideas, or behavior that it is difficult or impossible to have a productive conversation with him. His stubbornness makes it challenging to reason with him or persuade him to consider alternative viewpoints or options. The use of the word "too" implies that his stubbornness is excessive and beyond what is reasonable or tolerable in a conversation or interaction.

I then asked for exactly the contrast that Chomsky et al. speculated would foil the machine, for the meaning of the sentence, "John is too stubborn to talk to Jane." The answer was:

> The sentence "John is too stubborn to talk to Jane" means that John's stubbornness or strong refusal to change his thoughts or opinions is preventing him from engaging in a conversation with Jane. In other words, John's inflexibility or obstinacy is causing him to avoid or resist communicating with Jane.

The subtle grammatical differences between these sentences do not elude ChatGPT. Moreover, I seriously doubt that the reason that humans can distinguish the meanings of these sentences is because we can explain the rules of English syntax. We use intuition, not deductive reasoning.

Chomsky et al. observe that the programmers of AIs have struggled to ensure that they steer clear of morally objectionable content to be acceptable to most of their users. What they fail to observe is that humans also struggle to learn to apply appropriate filters to their own thoughts and feelings in order to be acceptable in society. Perhaps the LLMs can teach us something about how morally objectionable thoughts form in humans and how cultural pressures teach us to suppress them. Given the poor behavior of many humans in online forums, we could certainly benefit from new insights into how such behavior emerges.

In a reference to Jorge Luis Borges, Chomsky et al. conclude, "[g]iven the amorality, faux science and linguistic incompetence of these systems, we can only laugh or cry at their popularity." When Borges talks about experiencing both tragedy and comedy, he reflects on the complex superposition of human foibles and rationality. Rather than reject these machines, and rather than replacing ourselves with them, we should reflect on what they can teach us about ourselves. They are, after all, images of humanity as reflected through the internet.

Other critics say the LLMs perform a glorified form of plagiarism, stealing content created by humans. It is easily shown, however, that the data stored in an LLM cannot possibly contain verbatim more that a minuscule subset of the internet. The LLMs somehow encode the concepts and then resynthesize the expression "in their own words" (or pictures) much like humans do. Most human expression is also a reworking of concepts, texts, and images that have been seen before.

Many of these criticisms are implicitly comparing the AIs to ideal forms of intelligence and creativity that are fictions. In these fictions, an intelligence works with true facts and with logic (what Kant called "pure reason"), and creativity produces truly novel artifacts. But we have no precedents for such intelligence or creativity. It does not exist in humans nor in anything humans have created. Perhaps the LLMs have in fact achieved human-level intelligence, which works not with true facts but rather with preconceptions [14], not with logic as much as with intuition [11], and rarely produces anything truly novel (and when it does, the results are ignored as culturally irrelevant). Could it be that these AIs tell us more about humans than about machines?

Janelle Shane, an AI researcher, writes in her book, *You Look Like a Thing and I Love You*, that training an AI is more like educating a child than like writing a computer program [21]. Computer programs, at their lowest level, specify algorithms operating on formal symbols. The symbols are devoid of meaning, except in the mind of human observers, and the operations follow clearly defined rules of logic. Deep neural networks (DNNs), however, exhibit behaviors that are not usefully explained in terms of the operations of these algorithms [15]. An LLM is implemented on computers that perform billions of logic operations per second, but even a detailed knowledge of those operations gives little insight into the behaviors of the DNNs. By analogy, even if we had a perfect model of a human neuron and structure of neuron interconnections in a brain, we would still not be able to explain human behavior [18]. Given this situation, regulatory calls for "algorithmic transparency" are unlikely to be effective.

2 Explainable AI

A hallmark of human intelligence is our ability to explain things. DARPA's XAI program [8] sought to develop a foundation for explainable AI and yielded some useful results. For example, in image classification algorithms, it has become routine to identify portions of an image that most influence the classification. This can sometimes reveal interesting defects in the classification mechanisms, such as a classifier that distinguishes a wolf from a husky based on whether there is snow in the background [20]. For the most part, however, explaining the output of the LLMs remains elusive.

In contrast, humans are good at providing explanations for our decisions, but our explanations are often wrong or at least incomplete. They are often post hoc rationalizations, offering as explanations factors that do not or cannot account for the decisions we make. This fact about humans is well explained by Kah-

neman, whose Nobel-prize winning work on "prospect theory" challenged utility theory, a popular theory in economics at the time. In prospect theory, decisions are driven more by gains and losses rather than the value of the outcome. Humans, in other words, will make irrational decisions that deliver less value to them in the end. In *Thinking Fast and Slow* [11], Kahneman offers a wealth of evidence that our decisions are biased by factors that have nothing to do with rationality and do not appear in any explanation of the decision.

Kahneman reports, for example, a study of the decisions of parole judges in Israel by [5]. The study found that these judges, on average, granted about 65 percent of parole requests when they were reviewing the case right after a food break, and that their grant rate dropped steadily to near zero during the time until the next break. The grant rate would then abruptly rise to 65 percent again after the break. In Kahneman's words,

> The authors carefully checked many alternative explanations. The best possible account of the data provides bad news: tired and hungry judges tend to fall back on the easier default position of denying requests for parole. Both fatigue and hunger probably play a role. [11]

And yet, I'm sure that every one of these judges would have no difficulty coming up with a plausible explanation for their decision for each case. That explanation would not include any reference to the time since the last break.

Taleb, in his book *The Black Swan*, cites the propensity that humans have, after some event has occurred, to "concoct explanations for its occurrence after the fact, making it explainable and predictable" [25]. For example, the news media always seems to have some explanation for movements in the stock market, sometimes using the same explanation for both a rise and a fall in prices.

Taleb reports on psychology experiments where subjects are asked to choose among twelve pairs of nylon stockings the one they like best. After they had made their choice, the researchers asked them for reasons for their choices. Typical reasons included color, texture, and feel, but in fact, all twelve pairs were identical. Taleb concludes,

> Our minds are wonderful explanation machines, capable of making sense out of almost anything, capable of mounting explanations for all manner of phenomena, and generally incapable of accepting the idea of unpredictability. [25]

Demanding explanations from AIs could yield convincing explanations for anything, leading us to trust their decisions too much. Explanations for the inexplicable, no matter how plausible, are simply misleading.

It is a frustrating result of the recent successes in deep neural nets that people have been unable to provide explanations for many of the decisions that these systems make [16, Chapter 6]. In May 2018 a new European Union regulation called the General Data Protection Regulation (GDPR) went into effect with a controversial provision that provides a right "to obtain an explanation of the decision reached" when a decision is solely based on automated processing.

Legal scholars, however, argue that this regulation is neither valid nor enforceable [27]. In fact, it may not even be desirable. I conjecture that sometime in the near future, someone will figure out how to train a DNN to provide a convincing explanation for *any* decision. This could start with a generative-adversarial network (GAN) that learns to provide explanations that appear to be generated by humans.

Kahneman identifies two distinct human styles of thinking, a fast style (System 1) and a slow style (System 2) [11]. The slow style is capable of algorithmic reasoning, but the fast style, which is more intuitive, is responsible for many of the decisions humans make. It turns out that many of today's AIs more closely resemble System 1 than System 2. Even though they are realized on computers, they do not reach decisions by algorithmic reasoning.

Given that humans have written the computer programs that realize the AIs, and humans have designed the computers that execute these programs, why is it that the behavior of the programs proves inexplicable? The reason is that what the programs do is not well described as algorithmic reasoning, in the same sense that an outbreak of war is not well described by the interactions of protons and electrons. Explaining the implementation does not explain the decision.

Before the explosive renaissance of AI during the past two decades, AI was dominated by attempts to encode algorithmic reasoning directly through symbolic processing. What is now called "good old-fashioned AI" (GOFAI) encodes knowledge as production rules, if-then-else statements representing the logical steps in algorithmic reasoning [9]. GOFAI led to the creation of so-called "expert systems," which were sharply criticized by Dreyfus and Dreyfus in their book, *Mind Over Machine* [6]. They pointed out, quite simply, that following explicit rules is what novices do, not what experts do.

DNNs work primarily from examples, training data, rather than rules. The explosion of data that became available as everything went online catalyzed the resurgence of statistical and optimization techniques that had been originally developed in the 1960s through 1980s but lay dormant through the AI winter before exploding onto the scene around 2010. The techniques behind today's AI renaissance are nothing like the production rules of GOFAI.

There have been attempts to use machine learning techniques to learn *algorithmic* reasoning, where the result of the training phase is a set of explicable production rules, but these have proven to underperform neural networks. Wilson et al. created a program that could write programs to play old Atari video games credibly well [28]. Their program generated random mutations of production rules, and then simulated natural selection. Their technique was based on earlier work that evolved programs to develop certain image processing functions [19]. The Atari game-playing programs that emerge, however, are far less effective than programs based on DNNs. Wilson et al. admit this, saying that the main advantage of their technique is that the resulting programs are more explainable [28]. The learned production rules provide the explanations.

In contrast, once a DNN has been trained, even a deep understanding of the computer programs that make its decisions does not help in providing an

explanation for those decisions. Exactly the same program, with slightly different training, would yield different decisions. So the explanation for the decisions must be in the data that results from the training. But those data take of the form of millions of numbers that have been iteratively refined by backpropagation. The numbers bear no resemblance to the training data and have no simple mapping onto symbols representing inputs and possible decisions. Even a deep understanding of backpropagation does little to explain how the particular set of numbers came about and why they lead to the decisions that they do. Fundamentally, the decisions are not a consequence of algorithmic reasoning that could constitute an explanation.

In a previous paper, I study more deeply the relationship between explanations and algorithms [15]. The well-known work on "bounded rationality" of Herb Simon provides a useful framework for what we mean by an explanation [23]. What we seek is a description of a rational process that arrives at a decision, where a rational process is a sequence of logical deductions that reaches a conclusion. Simon's key insight, for which he got the Nobel Prize in economics, was that enconomic agents (individuals and organizations) do not have the capability to make the kinds of rational decisions that economists assumed they would. In his words:

> Theories that incorporate constraints on the information-processing capacities of the actor may be called theories of bounded rationality. [23]

He identified three human limitations: uncertainty about the consequences that would follow from alternative decisions, incomplete information about the set of alternatives, and complexity preventing the necessary computations from being carried out. He argued that "these three categories tend to merge," using the game of chess as an example and saying that the first and second, like the third, are fundamentally an inability to carry out computation with more than very limited complexity:

> What we refer to as "uncertainty" in chess or theorem proving, therefore, is uncertainty introduced into a perfectly certain environment by inability — computational inability — to ascertain the structure of that environment. [23]

Three decades later, he reaffirmed this focus on the process of reasoning:

> When rationality is associated with reasoning processes, and not just with its products, limits on the abilities of Homo sapiens to reason cannot be ignored. [24]

Reasoning and rationality as algorithmic, terminating sequences of logical deductions, are central to his theory, and he argued that economists' assumptions that agents would maximize expected utility was unrealistic in part because that maximization is intractable to a human mind. It requires too many steps.

An explanation, therefore, needs to be not just a description of a finite sequence of logical deductions, but also a *very short* sequence. Our human minds

cannot handle it otherwise. It turns out that such short sequences are not good descriptions of human decision making, and neither are they good descriptions of neural network decision making. Thus, any "explanation" of an AI decision (especially one provided by an AI) should be taken with a grain of salt. It may just be a post hoc rationalization.

3 Fear

Rapid change breeds fear. With its spectacular rise from the ashes in the last 15 years or so, we fear that AI may replace most white collar jobs [7]; that it will learn to iteratively improve itself into a superintelligence that leaves humans in the dust [1, 2, 26]; that it will fragment information so that humans divide into islands with disjoint sets of truths [16]; that it will supplant human decision making in health care, finance, and politics [12]; that it will cement authoritarian powers, tracking every move of their citizens and shaping their thoughts [17]; that the surveillance capitalists' monopolies, which depend on AI, will destroy small business and swamp entrepreneurship [29]; that it "may trigger a resurgence in mystic religiosity" [13]; and that it will "alter the fabric of reality itself" [13].

With this backdrop of fear, it is particularly disconcerting if the behavior of the AIs is inexplicable. We will have to learn to deal with them not so much as tools we control, but rather as partners with whom we coevolve [16]. As long as we can keep the coevolution symbiotic, we should be able to benefit. However, there are real risks.

As of 2023, the LLMs such as ChatGPT have been trained on mostly human-written data. It seems inevitable, however, that the LLMs will be generating a fair amount of the text that will end up on the internet in the future. The next generation of LLMs, then, will be trained on a mix of human-generated and machine-generated text. What happens as the percentage of machine-generated text increases? Feedback systems are complicated and unpredictable. Shumailov, et al. [22], show that such feedback learning leads to a kind of "model collapse," where original content (the human-written content) is forgotten. As that occurs, we will be left in the dust, possibly becoming unable to understand much of the generated content. The machines will be speaking to each other, not to us.

4 Conclusions

Humans have bounded rationality. We are unable to follow more than a few steps of logical deduction. A useful explanation for any decision, therefore, has to comprise just a few steps. The decisions made by a neural network may not be explicable with just a few steps. Demanding an explanation for an AI-generated decision may therefore be like demanding a post hoc rationalization for a human-generated decision. The decision may not have been arrived at by the steps that constitute the rationalization, but rather may be based much more on intuition or even biochemical factors such as mood or comfort. If we demand rational explanations for AI decisions, engineers will inevitably develop AIs that

are effective at constructing such post hoc rationalizations. With their ability to handle vast amounts of data, the AIs will learn to build rationalizations using many more precedents than any human could, thereby constructing rationalizations for *any* decision that will become very hard to refute. The demand for explanations, therefore, could backfire, resulting in effectively ceding to the AIs much more power.

References

1. Barrat, J.: Our Final Invention: Artificial Intelligence and the End of the Human Era. St. Martin's Press (2013)
2. Bostrom, N.: Superintelligence: Paths, Dangers, Strategies. Oxford University Press, Oxford, UK (2014)
3. Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y.T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M.T., Zhang, Y.: Sparks of artificial general intelligence: Early experiments with GPT-4. arXiv:2303.12712v1 [cs.CL] (March 22 2023). https://doi.org/10.48550/arXiv.2303.12712
4. Chomsky, N., Roberts, I., Watumull, J.: The false promise of ChatGPT. The New York Times (March 8 2023), https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html
5. Danziger, S., Levav, J., Avnaim-Pesso, L.: Extraneous factors in judicial decisions. Proceedings of the National Academy of Sciences of the United States of America **108**(17), 6889–6892 (April 26 2011). https://doi.org/10.1073/pnas.1018033108
6. Dreyfus, H.L., Dreyfus, S.E.: Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer. Free Press, New York (1986). https://doi.org/10.1016/0160-791X(84)90034-4
7. Ford, M.: Rise of the Robots — Technology and the Threat of a Jobless Future. Basic Books, New York (2015)
8. Gunning, D., Vorm, E., Wang, J.Y., Turek, M.: DARPA's explainable AI (XAI) program: A retrospective. Applied AI Letters **2:e61** (2021). https://doi.org/10.1002/ail2.61
9. Haugeland, J.: Artificial Intelligence: The Very Idea. MIT Press, Cambridge, Mass (1985)
10. Heaven, W.D.: The inside story of how ChatGPT was built from the people who made it. MIT Technology Review (March 3 2023), https://www.technologyreview.com/2023/03/03/1069311/inside-story-oral-history-how-chatgpt-built-openai/
11. Kahneman, D.: Thinking Fast and Slow. Farrar, Straus and Giroux, New York (2011)
12. Kelly, K.: The Inevitable: Understanding the 12 Technological Forces That Will Shape Our Future. Penguin Books, New York (2016)
13. Kissinger, H.A., Schmidt, E., Huttenlocher, D.: ChatGPT heralds an intellectual revolution. The Wall Street Journal (February 24 2023)
14. Kuhn, T.S.: The Structure of Scientific Revolutions. University of Chicago Press, Chicago, IL (1962)
15. Lee, E.A.: What can deep neural networks teach us about embodied bounded rationality. Frontiers in Psychology **25** (April 2022). https://doi.org/10.3389/fpsyg.2022.761808

16. Lee, E.A.: The Coevolution: The Entwined Futures of Humans and Machines. MIT Press, Cambridge, MA (2020)
17. Lee, K.F.: Super-Powers: China, Silicon Valley, and the New World Order. Houghton Mifflin Harcourt Publishing Company, New York (2018)
18. Lichtman, J.W., Pfister, H., Shavit, N.: The big data challenges of connectomics. Nature Neuroscience **17**, 1448–1454 (October 2014). https://doi.org/10.1038/nn.3837
19. Miller, J.F., Thomson, f.: Cartesian genetic programming. In: European Conference on Genetic Programming. vol. LNCS Vol. 10802, pp. 121–132. Springer (2000)
20. Ribeiro, M.T., Singh, S., Guestrin, C.: Why should I trust you? explaining the predictions of any classifier. In: International Conference on Knowledge Discovery and Data Mining. pp. 1135–1144. ACM (2016). https://doi.org/10.1145/2939672.2939778
21. Shane, J.: You look like a thing and I love you. Hachette, United Kingdom (2019)
22. Shumailov, I., Shumaylov, Z., Zhao, Y., Gal, Y., Papernot, N., Anderson, R.: The curse of recursion: Training on generated data makes models forget. arXiv:2305.17493v2 [cs.LG] (May 31 2023). https://doi.org/10.48550/arXiv.2305.17493
23. Simon, H.A.: Theories of bounded rationality. In: McGuire, C.B., Radner, R. (eds.) Decision and Organization, pp. 161–176. North-Holland Publishing Company, Amsterdam (1972)
24. Simon, H.A.: Bounded rationality in social science: Today and tomorrow. Mind & Society **1**, 25–39 (2000)
25. Taleb, N.N.: The Black Swan. Random House (2010)
26. Tegmark, M.: Life 3.0: Being Human in the Age of Artificial Intelligence. Alfred A. Knopf, New York (2017)
27. Wachter, S., Mittelstadt, B., Floridi, L.: Why a right to explanation of automated decision-making does not exist in the general data protection regulation. International Data Privacy Law, Available at SSRN (January 24 2017). https://doi.org/10.2139/ssrn.2903469, https://ssrn.com/abstract=2903469
28. Wilson, D.G., Cussat-Blanc, S., Luga, H., Miller, J.F.: Evolving simple programs for playing Atari games. In: The Genetic and Evolutionary Computation Conference (GECCO) (June 15-19 2018). https://doi.org/10.1145/3205455.3205578
29. Zuboff, S.: The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. PublicAffairs , Hatchette Book Group (2019)